Die Theorie der Collectivgegenstände.

Von

Gottl. Friedr. Lipps.

Die folgenden Untersuchungen wurden durch die Collectivmaßlehre Fechner's und durch die Contributions to the mathematical theory of evolution, die Pearson in den Philosophical Transactions of the Royal Society of London veröffentlicht hat, angeregt. Sie gründen sich jedoch weder auf das Werk Fechner's noch auf die Forschungen Pearson's, sondern geben eine neue Theorie der Collectivgegenstände, zu der mich die Erkenntniss führte, dass die von Fechner und Pearson ausgebildeten Methoden zwar weit über den von Quetelet vertretenen Standpunkt hinausführen, aber doch ihrem Wesen nach mit Mängeln behaftet sind, die auch bei der Anwendung der zuerst von Bruns empfohlenen und für die Collectivgegenstände nutzbar gemachten Methoden zur Darstellung willkürlich gegebener Functionen nicht überwunden werden können (vergl. II § 4).

Als wesentliche Stütze habe ich bloß das Theorem Jakob Bernoulli's in der ursprünglichen Form, welche die Ars conjectandi gibt, und die Theoria combinationis observationum erroribus minimis obnoxiae von Gauß zu nennen. Die Bedeutung des Bernoullischen Theorems erhellt aus den Erörterungen über die Methoden der Wahrscheinlichkeitsbestimmung (I § 6). Bezüglich des Zusammenhangs der im zweiten Capitel entwickelten Methode der Mittelwerthe mit dem von Gauß in der erwähnten Abhandlung aufgestellten Princip des mittleren Fehlers verweise ich auf II § 5.

I. Die inductive Wahrscheinlichkeitserkenntniss.

§ 1. Denkgegenstand und Begriff.

Ein beliebiger, aber bestimmter Gegenstand des Denkens, der durch A bezeichnet werden möge, sei gegeben.

Da A als gegeben angenommen wird, bleibt seine Herkunft unerörtert. Es ist jedoch klar, dass A einerseits in der ursprünglich und schlechthin gegebenen Welt des Bewusstseins wurzelt und anderseits der Bethätigung des Denkens sein Dasein verdankt. Denn niemand wird daran zweifeln, dass es weder aus Nichts erschaffene, noch von vorn herein vorhandene Gegenstände des Denkens gibt. Die Annahme von A setzt demnach voraus, dass an irgend einem, aus Empfindungen und Gefühlen bestehenden, räumlich und zeitlich geformten Inhalte des Bewusstseins das Denken sich irgendwie bethätigt habe und so zu A gelangt sei.

Ein Bewusstseinsinhalt liegt somit zu Grunde. Er gibt dem Denkgegenstande die reale, psychische Existenz. Es ist jedoch gleichgültig, von welcher Art er ist, und ob der Inhalt in seiner räumlichen und zeitlichen Wirklichkeit, oder ob nur ein durch das Denken abgesonderter und hervorgehobener Bestandtheil jenes Inhalts zum Gegenstande des Denkens gemacht wird.

Wie dem auch sein mag: in jedem Falle muss eine angebbare Beschaffenheit vorhanden sein. Denn A ist ein bestimmter, von anderen unterscheidbarer Gegenstand des Denkens. Die Bestimmtheit beruht auf dem Denken, durch das A in seiner Besonderheit erfasst und unterschieden wird. Da nun das Denken im Urtheilen sich vollzieht, so muss es möglich sein, die Beschaffenheit von A durch Urtheile, deren logisches Subject A ist, anzugeben. Jedes Urtheil aber, das nicht bloß das Vorhandensein von A behauptet, das also nicht ein bloßes Existentialurtheil ist, setzt A zu anderen Denkgegenständen in Beziehung. In solchen Beziehungen besteht die logische Bestimmtheit von A. In ihnen bietet sich das von A Erkannte oder Begriffene — der Begriff von A dar. Mit der Annahme eines bestimmten A wird daher zugleich ein Begriff von A vorausgesetzt. Der Begriff bestimmt oder definirt A.

Die zum Erfassen und Begreifen von A führende Denkarbeit lässt den zu Grunde liegenden Bewusstseinsinhalt unverändert. Es haften nur die Denkacte an jenem Inhalte, so dass Denkgegenstand und Begriff sich bilden. Der Vollzug dieses Denkprocesses ist durch die Bedürfnisse und die Fähigkeiten des nach Erkenntniss strebenden Geistes bedingt. Dabei kann eine größere oder geringere Vollständigkeit erreicht werden. Es werden darum im allgemeinen Besonderheiten des vorliegenden Bewusstseinsinhaltes unberücksichtigt bleiben; einestheils, weil das Denken an ihnen kein Interesse hat, anderntheils, weil das Denken ihre Bedeutung nicht zu erkennen vermag. Dies wäre nur dann ausgeschlossen, wenn der zu Grunde liegende Bewusstseinsinhalt in allen denkbaren Beziehungen klargestellt wäre, so dass er selbst in seiner Individualität den Denkgegenstand A bildete, und zugleich der Begriff des individuellen Bewusstseinsinhaltes gewonnen würde. So lange dies nicht der Fall ist, dürfen die außer Acht gelassenen Besonderheiten als abstufbar und veränderlich vorausgesetzt werden, so dass andere Bewusstseinsinhalte an Stelle des vorhandenen treten können, die alle in gleicher Weise das Erfassen und Begreifen von A gestatten. Jeder einzelne der eine begrenzte oder unbegrenzte Mannigfaltigkeit bildenden Bewusstseinsinhalte fällt unter den Begriff von A; und man kann diese Mannigfaltigkeit als den Umfang des Begriffs bezeichnen, während unter dem Inhalt des Begriffs die Gesammtheit der A bestimmenden Beziehungen zu verstehen ist.

Wird den im vorliegenden Begriff zusammengefassten Beziehungen eine neue beigesellt, so verändert sich mit dem Begriff zugleich der Denkgegenstand. Sein psychisches Substrat bleibt allerdings dasselbe; aber seine logische Bestimmtheit wird anders und somit ist er selbst nicht mehr der alte. Weil er aus dem vorhandenen A entsteht, möge er A' heißen. Da A die Voraussetzung für A' ist, so ist klar, dass jeder Bewusstseinsinhalt, der das Erfassen und Begreifen von A' gestattet, nothwendig auch das Erzeugen von A ermöglicht. Es kann aber sehr wohl ein Bewusstseinsinhalt mit den für A geltenden Beziehungen behaftet sein, ohne zugleich die neue, A' charakterisirende Beziehung zu zeigen. Der Begriff von A' hat darum keinen größeren Umfang als der Begriff von A; er kann aber einen ebenso großen oder kleineren Umfang haben. Ist der Begriffsumfang von A gleich demjenigen von A', so dass jeder Bewusstseinsinhalt, der als A gedacht werden kann, auch die zu A' führende Beziehung aufweist, so hat A' als der im Vergleich zu A vollkommener bestimmte Denkgegenstand zu gelten: der Begriff von A' stellt sich als eine Ergänzung des Begriffs von A dar. Dann ist kein Anlass, den Unterschied von A' und A in der sprachlichen Bezeichnung hervorzuheben, sondern dasselbe Wort wird im allgemeinen sowohl A' als auch A bezeichnen. Gibt es hingegen Bewusstseinsinhalte, die unter den Begriff von A fallen, ohne jener neuen Beziehung theilhaftig zu sein, so ist der Begriffsumfang von A' kleiner als derjenige von A. Der Uebergang von A zu A' führt alsdann zu einer Verengerung des Begriffs von A und es ist nothwendig, auch sprachlich A' von A zu unterscheiden. In dem einen wie in dem andern Falle kann A' das durch die neue Beziehung näher bestimmte A genannt werden.

Folgendes Beispiel diene zur Erläuterung:

Ein Blatt Papier, auf dem drei sich schneidende gerade Linien ausgezogen sind, kann den Bewusstseinsinhalt bilden, an dem das Denken sich bethätigt und durch alleiniges Beachten der Form des allseitig umgrenzten Ebenentheils das Dreieck als Denkgegenstand A erzeugt. Das deutliche Erfassen dieses A hat zur Voraussetzung, dass die Dreiecksform von anderen geometrischen Formen unterschieden und durch Besonderheiten, die nur dem Dreieck zukommen, gekennzeichnet wird. Der so zu Stande kommende Begriff des Dreiecks findet in der Definition: »das Dreieck ist ein von drei Geraden vollständig umgrenzter Ebenentheil« seinen Ausdruck. »Ebenentheil« »vollständige Umgrenzung«, »gerade Linie«, Anzahl »drei« sind somit die Denkobjecte, mit denen das Denkobject »Dreieck« in Beziehung steht. Diese Beziehungen sind enthalten in den Urtheilen: das Dreieck ist ein Ebenentheil, ist allseitig begrenzt, hat geradlinige Grenzen, hat drei Grenzlinien. Wird das Dreieck in dieser Weise begriffen, so sieht man ein, dass nicht nur der vorliegende, sondern jeder andere, durch drei Gerade vollständig umgrenzte Ebenentheil ebenso bestimmt ist und unter den Begriff des Dreiecks fällt. Denn die Lage und Größe der Seiten und Winkel, durch welche sich die einzelnen Dreiecke unterscheiden, bleibt unbeachtet. Jede die Lage und Größe der Seiten und Winkel betreffende Bestimmung verändert daher den angegebenen Begriff des Dreiecks. Wird z. B. festgesetzt, dass zwei Seiten einander gleich oder senkrecht zu einander gerichtet sein sollen, so wird der Begriffsumfang verengert. Denn von der ursprünglichen Mannigfaltigkeit der Dreiecke werden alle diejenigen, welche drei verschieden große Seiten oder keinen rechten Winkel haben, ausgeschieden. Wird hingegen erkannt, dass die Winkelsumme des Dreiecks gleich zwei Rechten sei, so tritt keine Verengerung, wohl aber eine Ergänzung des Begriffs vom Dreieck ein. Der Begriff wird nicht verengert, weil alsdann die drei Winkel eines jeden, von drei Geraden vollständig umgrenzten Ebenentheils zusammen zwei Rechte betragen. Der Begriff wird hingegen ergänzt, weil die urprüngliche Definition des Dreiecks, wie die Nicht-Euklid'sche Geometrie lehrt, sich auch mit der Annahme, dass die Summe der Dreieckswinkel kleiner oder größer als zwei Rechte sei, verträgt. Denn der Satz: »die Winkelsumme ist gleich zwei Rechten« steht und fällt mit der Voraussetzung, dass es zu einer gegebenen Geraden durch einen gegebenen Punkt eine und nur eine Parallele gibt.

Es zeigt sich so, dass Denkgegenstand und Begriff unlöslich mit einander verbundene Erzeugnisse des Denkens sind. Der Denkgegenstand ist das an einem zu Grunde liegenden Bewusstseinsinhalt durch das Denken Erfasste und Hervorgehobene Der Begriff ist die Definition eines Denkgegenstandes, d. h. die einen Denkgegenstand bestimmende Erkenntniss.

Diese enge Zusammengehörigkeit kann Anlass sein, Denkgegenstand und Begriff mit einander zu vermengen. Die Beachtung des Unterschieds dient aber, wie ich glaube, wesentlich zur Klärung und Sicherung der mannigfachen Schwankungen[1]) unterworfenen Auffas-

1) Sigwart (Logik, I., 2. Aufl. 1889, S. 316 ff.) unterscheidet eine dreifache Bedeutung des Wortes »Begriff«: eine empirische, für welche der Begriff ein durch ein Wort bezeichnetes natürliches psychologisches Erzeugniss in der Gestalt einer allgemeinen Vorstellung ist; eine ideale, metaphysische, wonach der Begriff den Zielpunkt des Erkenntnissstrebens angibt und der Ausdruck des Wesens der Dinge ist; eine zwischen der empirischen und metaphysischen Bedeutung liegende logische, welche Bestimmtheit und allgemeine Gültigkeit der Vorstellungen ohne Rücksicht auf die Uebereinstimmung zwischen Gedachtem und Seiendem fordert. — Wundt (Logik, I., 2. Aufl. 1893, S. 94 ff.) bezeichnet als logischen Begriff jeden Denkinhalt, der aus einem Urtheil durch Zergliederung desselben gewonnen werden kann, und als wissenschaftlichen Begriff das aus einer Reihe von Urtheilen sich ergebende Resultat einer Erkenntniss, wonach der logische und der wissenschaftliche Begriff als Anfang und Abschluss in der Ent-

sungsweise des Begriffs. — So ist ja auch ein räumlich wahrgenommener Gegenstand nothwendig räumlich bestimmt oder umgrenzt. Man wird aber den räumlich umgrenzten Gegenstand nicht mit der räumlichen Umgrenzung des Gegenstandes verwechseln. Ueberdies muss schon aus dem Grunde Denkobject und Begriff unterschieden werden, weil der Begriff selbst zum Denkobjecte werden kann. Denn wenn in der Geometrie das Dreieck Gegenstand der Untersuchung ist, so kann in der Logik und Erkenntnisslehre der Begriff des Dreiecks oder auch die Entstehung dieses Begriffs zum Forschungsobjecte werden.

Indessen ist hier nicht der Ort, auf die Bedeutung der hervorgehobenen Unterscheidung für die logische Lehre vom Urtheil und Begriff einzugehen. Die soeben angegebene Auffassungsweise des Begriffs muss aber beachtet werden, weil auf ihr die folgenden Darlegungen über deductives und inductives Erkennen im allgemeinen und über inductive Wahrscheinlichkeitserkenntniss insbesondere beruhen.

§ 2. Deduction und Induction.

Ich frage nun: wie kann die in dem Begriff des gegebenen A vorliegende Erkenntniss einen Zuwachs erhalten?

Diese Frage kann zunächst bestimmter gestellt werden. Denn jegliches Erkennen, das nicht in dem Erkennen des bloßen Vorhandenseins von Denkgegenständen besteht, ist ein Erkennen von Beziehungen zwischen verschiedenen Denkgegenständen. Ein Zuwachs zu der bezüglich A vorhandenen Erkenntniss erfolgt daher nur dann, wenn erkannt wird, dass ein bis jetzt noch nicht auf A bezogener

---

wicklung des Denkens sich gegenüberstehen. — Ohne auf solche Unterscheidungen einzugehen, nennt Th. Lipps (Grundzüge der Logik, 1893, S. 124) den Begriff die Bedeutungssphäre eines Wortes. Der Begriff ist »die Sphäre möglicher Bewusstseinsobjecte, die und sofern sie in einem sprachlichen Ausdruck ihren zusammenfassenden Mittelpunkt und damit zugleich ihre Abgrenzung gefunden haben«. — Schuppe hingegen (Erkenntnisstheoretische Logik, 1878, S. 121) sagt: »Der Begriff im eigentlichen Sinne kann überhaupt nur gedacht werden als das Zusammen eines Subjectes mit Prädicaten«. Er entsteht nicht nur, sondern er besteht aus Urtheilen; und er ist »nicht das abgesondert von dem hervorbringenden Urtheil existirende Resultat desselben, sondern er ist selbst dieses Urtheil, resp. eine Mehrheit solcher Urtheile«.

Denkgegenstand in einer Beziehung zu A steht. Dieser neue Denkgegenstand möge durch B und die Beziehung zwischen A und B durch $(A; B)$ bezeichnet werden. Die Frage, wie die Erkenntniss einen Zuwachs erhalten kann, darf daher durch die bestimmtere Frage ersetzt werden: wie kann die Zugehörigkeit eines bis jetzt noch nicht auf A bezogenen B zu A, oder wie kann die Beziehung $(A; B)$ erkannt werden?

Da diese Beziehung erst erkannt werden soll, so darf sie nicht als bekannt vorausgesetzt werden. Sie kann daher nicht den Inhalt eines unmittelbar vollziehbaren Urtheils bilden, sondern muss erschlossen werden. Soll aber $(A; B)$ erschlossen werden, so muss eine Reihe zusammenhängender Denkacte ausgeführt werden, die eine Verbindung zwischen A und B herstellen. Dies setzt voraus, dass insbesondere ein an A anknüpfender Denkact ausführbar sei. Es ist darum festzustellen, wie das Denken an A anknüpfen und demzufolge B mit A verknüpfen kann. Hierfür bestehen zwei Möglichkeiten.

a. Deduction.

A ist ein durch die Beziehungen, welche den Begriff von A bilden, bestimmter Denkgegenstand. Eine dieser Beziehungen kann daher dazu dienen, um eine zu B führende Reihe von Denkacten an A anzuknüpfen. Der hieraus sich ergebende Schluss auf das Bestehen der Beziehung $(A; B)$ gründet sich alsdann auf den Begriff von A: die Beziehung $(A; B)$ wird aus dem Begriff von A abgeleitet. Eine solche Ableitung nenne ich Deduction.

Bezeichnet $(A; C)$ jene dem Begriff von A angehörende Beziehung, so muss demnach der Denkgegenstand C mittelbar oder unmittelbar mit B verknüpft sein, so dass die Beziehung $(C; B)$ als bekannt vorausgesetzt werden darf. Es folgt dann aus dem Zusammenbestehen von $(A; C)$ und $(C; B)$ die Beziehung $(A; B)$, d. h. die Beziehungen $(A; C)$ und $(C; B)$ bilden die Prämissen eines Syllogismus, dessen Schlusssatz die zwischen A und B bestehende, erschlossene Beziehung zum Ausdruck bringt.

Da $(A; B)$ aus dem Begriff von A folgt, so bleibt dieser Begriff unverändert. Denn das Erkennen einer Folge kann den Grund, aus dem die Folge sich ergibt, nicht ändern. Der Begriff von A hat darum nach dem Vollzug der Deduction den nämlichen Inhalt und Umfang wie vorher, so dass die neu erschlossene Beziehung ohne weiteres von jedem Bewusstseinsinhalt gilt, der unter den Begriff von A fällt.

Es wird aber ein Zuwachs an Erkenntniss erzielt. Sonach gibt es außer den Beziehungen, welche einen gegebenen Gegenstand des Denkens bestimmen und den Begriff desselben festlegen, noch andere Beziehungen, die mittelbar (auf Grund des Begriffs) dem Denkgegenstande angehören. Deshalb scheint es zweckmäßig, den Begriff und den logischen Bereich eines Denkgegenstandes zu unterscheiden. Der Begriff erhält die Beziehungen, welche dem Denkgegenstande die Bestimmtheit für das Denken verleihen; der logische Bereich umfasst alle auf Grund des Begriffs erschließbaren Beziehungen.

Jede Deduction aus A führt alsdann zu einer dem logischen Bereich von A angehörenden Beziehung. Und da jede derartige Beziehung durch den Begriff von A an A geknüpft ist, so kann sie stets auf deductivem Wege erkannt werden. Die Deduction wird mithin dadurch gekennzeichnet, dass die erschlossene Beziehung aus dem vorliegenden Begriffe folgt, somit für den ganzen Umfang des Begriffs gilt und den Inhalt des Begriffs nicht verändert.

Ist z. B. ein Parallelogramm als Denkgegenstand A gegeben, und wird der Begriff desselben durch die Definition: »das Parallelogramm ist ein Viereck, an dem je zwei gegenüberliegende Seiten parallel sind« festgelegt, so folgt aus dem Parallelismus je zweier gegenüberliegender Seiten die Congruenz der beiden Dreiecke, in die das Parallelogramm durch eine Diagonale zerlegt wird; und aus der Congruenz dieser Dreiecke folgt weiterhin die Gleichheit je zweier gegenüberliegender Seiten des Parallelogramms. Diese Gleichheit ist hier der Denkgegenstand B, der durch den Begriff von A an A geknüpft ist. Die so gewonnene Erkenntniss gilt für alle Parallelogramme, da sie alle in gleicher Weise als Vierecke mit je zwei parallelen Seiten begriffen werden. Sie lässt ferner den anfänglichen Begriff des Parallelogramms wie er ist, da die Beziehung zwischen dem Parallelogramm und der Gleichheit gegenüberliegender Seiten aus der Beziehung zwischen dem Parallelogramm und dem Parallelismus gegenüberliegender Seiten folgt.

b. Induction.

Ist $(A; B)$ nicht deductiv erkennbar, so kann diese Beziehung nicht zum logischen Bereiche von A gehören: sie ist dem Begriff von A in der vorliegenden Form fremd. Die Erkenntniss von $(A; B)$ setzt daher eine veränderte Auffassung von A voraus, so dass an dem aus A entstandenen A' die fragliche Beziehung, die somit in Wahrheit eine Beziehung $(A'; B)$ ist, haftet. Ist nun A' an Stelle von A getreten, so gehört die Beziehung $(A'; B)$ entweder dem Begriff oder dem logischen Bereich von A' an. Sie gehört zum Begriff von A', wenn sie unmittelbar zur Bestimmung von A' dient. Sie gehört zum logischen Bereich von A', wenn A' durch eine andere neue Beziehung $(A'; C)$, die gleichfalls dem Begriff von A in der vorliegenden Gestalt fremd ist, bestimmt wird, und sodann $(A'; B)$ aus $(A'; C)$ deducirt werden kann. Es ist aber A' nichts anderes, als das durch die Bezugnahme auf B oder C näher bestimmte A. Man kann darum A' das mit B unmittelbar oder mittelbar in Beziehung gesetzte A nennen und sagen, dass die Beziehung $(A; B)$ den bereits vorhandenen Beziehungen beigesellt und in den Begriff von A oder in den logischen Bereich von A eingeführt wird. Eine solche Einführung nenne ich Induction.

Es fragt sich nun, wie die veränderte Auffassung von A oder der Fortgang von A zu A' oder die nähere Bestimmung von A durch Induction möglich ist. Um dies einzusehen, ist die Entstehung der Denkgegenstände zu berücksichtigen.

Ein Denkgegenstand entsteht, indem das Denken an dem ursprünglich und schlechthin gegebenen Inhalte des Bewusstseins sich bethätigt. A kann daher ohne einen zu Grunde liegenden Bewusstseinsinhalt und die an letzterem vollführte Denkarbeit nicht bestehen. Der Bewusstseinsinhalt und die Denkarbeit bilden die unumgängliche Voraussetzung für A. Ich will sie die Unterlage von A nennen. Demzufolge ist es möglich, statt auf den Begriff von A vielmehr auf die Unterlage von A sich zu stützen, um den Zusammenhang von A und B zu erschließen. Die Unterlage von A kann nämlich zugleich die Unterlage von B oder von einem in erkennbarer Beziehung zu B stehenden C sein, so dass, wenn A durch das Denken erfasst wird, zugleich B oder C zum Gegenstande des Denkens wird.

Dann ist das Zusammensein von A und B oder von A und C eine Thatsache, die bemerkt und im Gedächtniss behalten wird, so dass die Reproduction von A die Erinnerung an B oder C wachruft. Es entsteht so die Neigung, auch bei einer veränderten Unterlage das Zusammen von A und B oder von A und C zu erwarten. Und diese Neigung wird um so stärker, je öfter die Erwartung bestätigt wurde.

Die mehr oder minder feste Association von Denkobjecten, wobei ihre jeweilige gemeinsame Unterlage außer Acht bleiben kann, berechtigt aber keineswegs, diese Denkobjecte ohne Rücksicht auf ihre Unterlage als zusammengehörig aufzufassen, so dass A durch die Bezugnahme auf B oder C eine größere Bestimmtheit für das Denken erhielte und das an dieser oder jener Unterlage thatsächlich beobachtete Zusammensein zu einer für das Denken verbindlichen Zusammengehörigkeit würde. Und doch muss dies geschehen, wenn eine Induction stattfinden soll.

Die Herbeiführung dieses Erfolges hat zur Voraussetzung, dass das Denken befähigt ist, Objecte auf einander zu beziehen und mit einander zu verknüpfen, d. h. einen thatsächlich gegebenen Zusammenhang als logischen Zusammenhang aufzufassen; und dass ferner dem menschlichen Geiste der Trieb innewohnt, auf Grund dieser Befähigung das gesammte Sein und Geschehen in einen durchgängigen und widerspruchslosen Zusammenhang zu bringen. Es wird sonach allerdings stets das Streben vorhanden sein, ein mit A associirtes B oder C als zugehörig zu A anzusehen.

Dem bezeichneten Trieb zufolge müsste aber jedes thatsächliche Zusammensein von Denkobjecten in gleicher Weise und in gleichem Maße für das Denken als Zusammengehörigkeit gelten. Dies mag für die Anfänge des Erkennens in der That zutreffen. So wird wohl einem unreifen Denken jede in die Augen fallende Eigenschaft und jeder der Aufmerksamkeit sich aufdrängende Zustand eines Dinges ohne Kritik für wesentlich gelten und das Behaftetsein mit jenen Eigenschaften und Zuständen als zum Begriff des Dinges gehörig erscheinen. Diese naive Auffassungsweise hält aber nicht Stand, wenn ein Vorrath von Erkenntnissen sich sammelt und jede neue Beziehung in ein System bereits erkannter Beziehungen sich widerspruchslos einfügen lassen muss.

Es ist darum einestheils die im Begriff von A bereits vorliegende

Erkenntniss in Rücksicht zu ziehen, und es sind anderntheils die aus der neuen Beziehung sich ergebenden Folgerungen zu beachten, wenn man aus der Thatsache, dass an dieser oder jener Unterlage A mit B oder mit C verbunden ist, mit Fug und Recht den Schluss ziehen will, dass diese Denkobjecte ohne Rücksicht auf jene Unterlage zusammengehören. Denn die Beziehung (A; B) oder (A; C) muss in dem mannigfach zusammenhängenden Gewebe von Erkenntnissen ihre Stelle behaupten. Auch kann erst durch die Beachtung dieses Zusammenhanges entschieden werden, ob die neue Beziehung für den ganzen Umfang des Begriffs von A oder nur für einen Theil desselben gilt, ob — mit andern Worten — die nähere Bestimmung des vorhandenen Begriffs von A eine Ergänzung dieses Begriffs ist, oder zu einer Verengerung des Begriffs führt.

Hiernach ist der Trieb nach Erkenntniss die Vorbedingung, und die gemeinsame Unterlage der Denkobjecte die Grundlage und der Ausgangspunkt der Induction. Ihre Ausführung besteht aber in dem Schluss, dass die an dieser oder jener Unterlage thatsächlich verbundenen Denkobjecte A und B oder A und C in einer logischen, für das Denken verbindlichen Zusammengehörigkeit stehen.

Es kann aber kein Kriterium angegeben werden, wonach von vorn herein und in jedem Falle entschieden werden könnte, ob die Beziehung (A; B) oder (A; C) dem Begriff von A in seinem ganzen Umfang oder nur mit Einschränkung zukommt. Der Erfolg der Induction kann darum ebensowohl eine Begriffsergänzung, wie eine Begriffsverengerung sein.

§ 3. Scheinbare und wirkliche Induction.

Es ist denkbar, dass der vorliegende Begriff von A hinreicht, um die als Thatsache gegebene Verbindung mit B zu verstehen. Die Beziehung zwischen A und B kann dann nicht dem Begriff selbst, sie muss vielmehr dem logischen Bereiche von A angehören. Denn mit A ist der Begriff von A vorauszusetzen (nach § 1). Es müsste somit auch die Beziehung zwischen A und B, wenn sie zu dem Begriff von A gehörte, schon bekannt sein, und die Constatirung der Thatsache, dass A mit B wirklich verbunden ist, wäre für das Denken ohne Bedeutung. Gehört aber die Beziehung dem logischen Bereiche von A an, so muss es möglich sein, dieselbe aus dem Begriff von A zu deduciren, da die Erforschung des logischen Bereichs Aufgabe der Deduction ist.

Der Ausführung der Deduction können indessen Schwierigkeiten gegenüberstehen, die zur Zeit unüberwindlich sind. Die den Zusammenhang von A und B vermittelnden Denkobjecte wollen sich vielleicht nicht finden lassen, so dass man den Weg nicht kennt, auf dem die Deduction ans Ziel gelangen kann. Dann ist es gewiss sehr werthvoll, die Verbindung von B mit A in einem besonderen Fall als Thatsache vor Augen zu haben, da man hierdurch möglicherweise auf den rechten Weg geleitet wird. Eine Induction, eine Einführung in den logischen Bereich von A findet jedoch nicht statt; eben weil die Beziehung — wenn sie auch nicht erkannt war — doch dem logischen Bereich von A bereits angehört.

Ist z. B. ein rechtwinkliges Dreieck als Denkobject A gegeben, so kann man sich ohne weiteres darüber klar werden, dass eine Seite, etwa die Hypotenuse a, eine Function der beiden andern Seiten, der Katheten b und c, ist. Denn durch die beiden Katheten ist das rechtwinklige Dreieck und somit auch die Hypotenuse a vollständig bestimmt. Diese Einsicht ermöglicht aber keine Bestimmung der Form der Function; und man kann nur vorläufig auf Grund der Bemerkung, dass $a = b$, wenn $c = 0$, und $a = c$, wenn $b = 0$, etwa die Form

$$a = b + c + b \cdot c \cdot \varphi\,(bc)$$

oder allgemein die Form

$$a^n = b^n + c^n + b \cdot c \cdot \varphi\,(bc)$$

voraussetzen, wo $\varphi\,(bc)$ eine Function sein muss, die für $b = 0$ und $c = 0$ nicht unendlich groß wird, damit für $b = 0$ und für $c = 0$

$$b \cdot c \cdot \varphi\,(bc) = 0\,.$$

Dann wird es in der That durchaus zweckmäßig sein, die Seiten eines construirten rechtwinkligen Dreiecks sorgfältig zu messen. Findet man beispielsweise für a, b, c der Reihe nach 5, 3, 4 Längeneinheiten, oder eine solche Annäherung an diese Werthe, dass die Abweichungen als bloße Messungsfehler aufgefasst werden können, so wird man leicht bemerken, dass

$$5^2 = 3^2 + 4^2\,.$$

Man wird so zu der Vermuthung geführt, dass für die vorausgesetzte Form der Function $n = 2$ und $\varphi(bc) = 0$ und somit

$$a^2 = b^2 + c^2$$

zu setzen sei. Die frühere Rathlosigkeit über den Weg, auf dem man zu der Bestimmung der Abhängigkeit zwischen den Seiten a, b und c gelangen könne, ist nun beseitigt: die empirische Feststellung der Werthe $a = 5$, $b = 3$, $c = 4$ dient als Wegweiser. Denn nun wird man die Quadrate über den Seiten a, b, c construiren und durch Vergleichen und geeignetes Zerlegen derselben zu dem Beweise des Pythagoreischen Lehrsatzes gelangen. Man findet so in der Gleichheit zwischen dem Quadrate über der Hypotenuse und der Summe der Quadrate über den beiden Katheten das Denkobject B, das mit dem als Denkobject A gegebenen rechtwinkligen Dreieck in logischem Zusammenhange steht. Die Beziehung zwischen A und B gehört aber dem logischen Bereiche von A an, da sie auf Grund des vorhandenen Begriffs von A erkannt werden kann.

Eine Induction in dem oben angegebenen Sinne läge auch dann nicht vor, wenn die Form der Function weniger einfach wäre und die empirische Feststellung der Werthe von a, b und c in einem einzigen Falle nicht genügen würde, um auf den rechten Weg zu leiten. Selbst wenn hundert oder tausend solcher empirischer Feststellungen nothwendig wären, so können sie doch nur als Wegweiser für die Ausführung der Deduction dienen. Und der bereits vorhandene Begriff müsste ausreichen, um die fragliche Beziehung zu erkennen; denn andernfalls würde nicht eine geometrische, für das Dreieck als solches gültige Wahrheit gefunden werden.

Um auch ein dem arithmetischen Gebiete angehörendes Beispiel von scheinbarer Induction anzuführen, möge der Begriff der Potenz, d. h. eines aus lauter gleichen Factoren bestehenden Productes als bekannt vorausgesetzt werden und die allgemeine Potenz von $a + b$ als Denkobject A gegeben sein. Es besteht dann offenbar eine Beziehung zwischen den Potenzen von $a + b$ und den Potenzen von a und b, die vorläufig in der Form

$$(a + b)^n = a^n + c_{1,n}a^{n-1}b + c_{2,n}a^{n-2}b^2 + \ldots c_{n-1,n}ab^{n-1} + b^n$$

darstellbar ist. Sind nun die zur Bestimmung der Coefficienten c dienenden Sätze aus der Combinatorik unbekannt, und ist man nicht im Stande, dieselben abzuleiten, um so aus dem Begriff der Potenz jene Beziehung ohne weiteres zu deduciren, so kann man aus der Berechnung von

$$(a+b)^2 = a^2 + 2ab + b^2\,,$$
$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3\,, \text{ u. s. w.}$$

die Vermuthung gewinnen, dass allgemein

$$(a+b)^n = a^n + \frac{n}{1}a^{n-1}b + \frac{n(n-1)}{1\cdot 2}a^{n-2}b^2 + \ldots + b^n\,.$$

Die Richtigkeit dieser Vermuthung wird sodann bewiesen, indem man zeigt, dass nun auch

$$(a+b)^{n+1} = a^{n+1} + \frac{n+1}{1}a^n b + \frac{(n+1)n}{1\cdot 2}a^{n-1}b^2 + \ldots + b^{n+1}\,.$$

Denn da die Formel für $n = 2$ gilt — wie die unmittelbare Berechnung lehrt —, so gilt sie auch für $n = 3$ und somit auch für $n = 4$, folglich auch für $n = 5$ u. s. w. in infinitum.

Hier sind es nicht einmal empirisch festgestellte Thatsachen wie im vorigen Beispiele, sondern Berechnungen, d. h. auf dem Begriffe der Zahl und insbesondere der Potenz beruhende Deductionen, die als Wegweiser zum Auffinden der gesuchten Beziehung dienen. Diese vorläufigen Berechnungen haben aber mit dem Beweise selbst gar nichts zu thun. Denn der Beweis besteht lediglich darin, dass ein zunächst bloß vermuthetes Gesetz als richtig erwiesen wird, indem durch den Schluss von n auf $n+1$ und durch die Erprobung für $n = 2$ gezeigt wird, dass es für jeden einzelnen der unendlich vielen Fälle $n = 2, 3, 4 \ldots$ in infinitum thatsächlich gilt. Man sollte darum das gewöhnlich als vollständige Induction bezeichnete mathematische Beweisverfahren vielmehr einen Beweis durch eine **unbegrenzte Reihe von Deductionen** nennen.

Die so erkannte Beziehung gehört aber in der That dem logischen Bereiche der Potenz von $a + b$ an; sie enthielte ja sonst keine arithmetische, auf dem Begriffe der Zahl beruhende Erkenntniss.

Der Schein einer Induction entsteht demgemäß dann, wenn die Beobachtung einzelner Thatsachen oder eine für specielle Fälle deductiv gewonnene Einsicht zu einer Vermuthung führt, die es ermöglicht, aus dem vorhandenen Begriffe eines Denkgegenstandes eine dem logischen Bereiche dieses Denkgegenstandes angehörige Beziehung zu deduciren.

Eine wirkliche Induction kann hingegen nur dann stattfinden, wenn es nicht möglich ist, aus der im Begriffe des gegebenen A vorliegenden Erkenntniss die Beziehung zwischen A und B zu erschließen. Man muss daher über den vorliegenden Begriff hinausgehen. Dies Hinausgehen ist aber nur in der Weise möglich, dass den im vorhandenen Begriffe zusammengefassten Beziehungen eine neue Beziehung beigesellt wird. Darum wurde oben die Einführung einer solchen neuen Beziehung in den vorhandenen Begriff, die entweder zu einer Begriffsergänzung oder zu einer Begriffsverengerung führt, als das Wesen der Induction hervorgehoben.

Diese Unterscheidung zwischen scheinbarer und wirklicher Induction lässt sich nicht durchführen, wenn man in dem Aufsteigen vom Einzelnen zum Allgemeinen oder in der Verallgemeinerung überhaupt das wesentliche Kennzeichen der Induction erblickt. Denn auch in den angeführten Beispielen scheinbarer Induction ist anfänglich nur eine beschränkte Einsicht und zuletzt die für den ganzen Begriffsumfang gültige Erkenntniss vorhanden. Es findet daher auch hier ein Aufsteigen vom Einzelnen zum Allgemeinen statt, wennschon das Einzelne nur als Nothbehelf dient und schließlich nicht das Einzelne sondern der vorhandene Begriff den Fortschritt in der Erkenntniss ermöglicht. Da nun den obigen Darlegungen zufolge die scheinbare Induction sich als eine Deduction darstellt, bei der die Mittelglieder nicht von vorn herein bekannt sind und unmittelbar vor Augen stehen, sondern als Vermuthungen durch die Betrachtung specieller Fälle nahe gelegt werden, so muss ich zugeben, dass ich zu der üblichen Auffassung [1]) in einen gewissen Gegensatz trete. In-

---

1) Die Induction als Verallgemeinerung aufzufassen, wurde durch Aristoteles veranlasst, der die Induction oder Epagoge im Gegensatz zum Syllogismus als das Hinaufsteigen vom Einzelnen zum Allgemeinen bezeichnete (ἐπαγωγὴ ἡ ἀπὸ τῶν καθ᾽ἕκαστον ἐπὶ τὰ καθόλου ἔφοδος). Angaben über die Entwicklung der Lehre von der Induction findet man unter anderen bei Apelt, Theorie der Induction, 1854, S. 129 ff. und bei Sigwart, Logik, II., 2. Aufl., S. 403 ff. — Sigwart selbst definirt das Inductionsverfahren »als Methode der Gewinnung allgemeiner Sätze aus einzelnen Wahrnehmungen«. — Aehnlich sagt Wundt (Logik, II., 1. Abth., 2. Aufl., S. 25): »Als das Resultat der Induction ergibt sich stets ein allgemeiner Satz, welcher die einzelnen Thatsachen der Erfahrung,

dessen wird wohl aus den folgenden Erörterungen die Triftigkeit der hier vertretenen Auffassungsweise erhellen.

§ 4. Axiom und Hypothese.

Da eine wirkliche Induction nicht auf vorhandene Begriffe sich gründen kann, so muss man von der Unterlage der Denkobjecte ausgehen, um die Zusammengehörigkeit von A und B zu erschließen. Nimmt man der Einfachheit wegen an, dass B selbst an A durch die gemeinsame Unterlage geknüpft sei; sieht man also davon ab, dass auch ein anderer, mit B in erkennbarer Beziehung stehender Denkgegenstand C die nämliche Unterlage wie A haben und dass auf diese Weise die Erkenntniss von (A; B) durch die inductiv erkannte Beziehung (A; C) in Verbindung mit der Beziehung (C; B) gewonnen werden kann — so ist die Thatsache, dass die Denkobjecte A und B irgendwann und irgendwo verbunden sind, die einzige Quelle, aus der die Erkenntniss der Zusammengehörigkeit von A und B geschöpft werden kann.

Der Zusammenhang von Naturerscheinungen, die Verknüpfung psychischer Erlebnisse, die Vereinigung verschiedener Merkmale an einem und demselben physischen oder psychischen Träger sind solche Verbindungen. Denn indem die Naturerscheinung A, oder das Erlebniss A, oder das Merkmal A Gegenstand des Denkens wird, drängt sich zugleich die mit A zusammenhängende Naturerscheinung B, oder das mit A verknüpfte Erlebniss B, oder das mit A vereinigte Merkmal B der Beachtung auf, so dass nicht nur A sondern auch B und zwar zusammen mit A (an der nämlichen Unterlage) bemerkt wird.

In diesen beispielsweise angeführten Fällen liegt es unzweifelhaft an der Beschaffenheit des Gegebenen, das Gegenstand des Bewusstseins wird, wenn A und B zusammen bemerkt werden. Die dem In-

---

die zu seiner Ableitung gedient haben, als specielle Fälle in sich enthält«. — Schuppe nennt (Erkenntnisstheoretische Logik, S. 130) Induction »das Verfahren, Ursache und Wirkung zu erkennen und so allgemeine Sätze aufzustellen«. — Th. Lipps versteht (Grundzüge der Logik, S. 170) »unter Induction allgemein die Verallgemeinerung einzelner Urtheile« und bezeichnet als die Voraussetzung der Induction das Denkgesetz, demzufolge jedes Erfahrungs- oder Erinnerungsurtheil ein allgemeines Urtheil von Haus aus unmittelbar in sich schließe.

ductionsprocess zu Grunde liegende Thatsache ist alsdann eine **Erfahrungsthatsache.**

Weniger zweifelsfrei erscheint die Grundlage der Induction in folgendem Beispiele: Das Dreieck wird bekanntlich als ein von drei Geraden vollständig umgrenzter Ebenentheil definirt. Nun ist dieses A dadurch ausgezeichnet, dass die Summe der drei Winkel gleich zwei Rechten ist. Diese Eigenschaft ist das B, das thatsächlich mit A verbunden ist. Die Verbindung von A und B ist aber nicht aus dem Begriff von A deducirbar; sie wird somit nicht scheinbar sondern wirklich auf inductivem Wege erkannt. Denn alle Versuche, den Satz von der Winkelsumme des Dreiecks zu beweisen, schlugen fehl; und die Unmöglichkeit des Beweises wurde dargethan, als es gelang, ein folgerichtiges System geometrischer Wahrheiten unter der Voraussetzung, dass die Winkelsumme nicht gleich zwei Rechten sei, oder unter Preisgabe des Parallelenaxioms zu entwickeln. Es schien darum die Thatsache, dass im Dreieck die Winkelsumme gleich zwei Rechten ist, eine Erfahrungsthatsache zu sein; und dem entsprechend wurde die Geometrie als eine auf Erfahrung gegründete Wissenschaft bezeichnet.

Es ist aber zu beachten, dass die Denkobjecte zwar in dem ursprünglich und schlechthin Gegebenen wurzeln, und somit eine als Thatsache vorliegende Verbindung allerdings in der Beschaffenheit des Gegebenen ihren Ursprung haben kann, dass aber anderseits die Denkobjecte auch der Bethätigung des Denkens ihr Dasein verdanken und darum Spuren der Denkarbeit an sich tragen, die nicht von der Beschaffenheit des Gegebenen sondern von der Art und Weise des Denkens herrühren.

Nun bethätigt sich das Denken auch im räumlichen und zeitlichen Ordnen des Gegebenen oder in der anschaulichen **Auffassung** des Gegebenen. Demnach kann nur die Erforschung der Beschaffenheit des Denkens zu einer Entscheidung darüber führen, ob der Satz von der Winkelsumme des Dreiecks und in solidarischem Zusammenhange damit die im Parallelenaxiom ausgesagte Eigenschaft der Ebene der Erfahrung entstammt, oder vielmehr auf die dem menschlichen Geiste eigenthümliche Weise des räumlichen Ordnens sich gründet. Je nachdem das erstere oder das letztere der Fall ist, wird die

Geometrie mit Recht oder mit Unrecht eine Erfahrungswissenschaft genannt.

Man wird hierdurch darauf aufmerksam, dass eine als Thatsache vorliegende Verbindung von Denkobjecten nicht nur in der Beschaffenheit des Gegebenen oder in der Erfahrung, sondern auch in der Art und Weise des Denkens oder in den Denkformen ihren Ursprung haben kann. Und es ist offenbar unmöglich, beide Arten von Thatsachen in jedem Falle unzweideutig zu unterscheiden. Denn die Ansichten über den Ursprung der geometrischen Erkenntniss sind ja in der That getheilt, indem die grundlegenden Sätze von den einen als a priori gültige, von den andern als empirisch gefundene Wahrheiten aufgefasst werden.

Demzufolge ist es nicht durchführbar, nur Erfahrungsthatsachen als Grundlage des Inductionsprocesses vorauszusetzen. Vielmehr sind auch die auf den Denkformen beruhenden Thatsachen zu berücksichtigen, wenn man das ganze Gebiet inductiver, d. h. auf Thatsachen gegründeter und nicht aus vorhandenen Begriffen deducirbarer Erkenntniss übersehen will. Und darum wurde oben nicht nur der dem Denkgegenstande zu Grunde liegende Inhalt des Bewusstseins, sondern auch die den Denkgegenstand erzeugende Denkarbeit als die Unterlage des Denkgegenstandes bezeichnet, von der die Induction ihren Ausgang nehmen müsse.

Beruht nun das Zusammen der Denkobjecte A und B auf der Art und Weise des Denkens, so kann offenbar A nicht anders als in Verbindung mit B gedacht werden. Dann wird, welches auch der dem A zu Grunde liegende Bewusstseinsinhalt sein möge, immer und überall, sobald A erzeugt wird, auch B mit A verbunden sein. Die Beziehung zwischen A und B gehört somit dem Begriffe von A in seinem ganzen Umfange an. — Da ferner B nothwendig, nämlich auf Grund der Formen, an die das Denken gebunden ist, zu A gehört, so kann man nicht sagen, dass A in den bis jetzt beachteten Fällen, oder soweit die Erfahrung reicht, mit B verbunden sei; es ist vielmehr gar nicht denkbar, dass A ohne B zum Gegenstande des Denkens werden könne. Die Zusammengehörigkeit von A und B ist daher unmittelbar und unbedingt, d. h. unabhängig von aller Erfahrung oder a priori gewiss. — Wenn aber auch B nothwendig und durchweg mit A verbunden ist, so muss diese Verbindung doch erst erkannt werden. Vorbedingung hierfür ist, dass B neben A und verbunden mit A bemerkt wird und nicht nur unbemerkt mit A verknüpft ist. Aber erst die Reflexion über die Bethätigungsweise des Denkens ermöglicht die Erkenntniss, dass B nicht bloß an diesem oder jenem Bewusstseinsinhalte, der zur Erzeugung von A Gelegenheit gab, sich thatsächlich vorfindet, sondern immer sich vorfinden muss, wenn A Gegenstand des Denkens werden soll. Nunmehr wird die Beziehung zwischen A und B als ein denknothwendiger Bestandtheil in den Begriff von A eingeführt. Ist aber auf diese Weise die Zugehörigkeit von B zu A auf inductivem Wege erkannt worden, so fühlt man nicht das Bedürfniss, immer von neuem zu prüfen, ob denn auch wirklich in jedem Falle A mit B verbunden ist. Die Verbindung von A und B wird vielmehr unbedingt und unabhängig von der Erfahrung gefordert: sie ist ein Axiom.

In den Axiomen stellt sich mithin diejenige inductiv gewonnene Erkenntniss dar, die auf den in der Bethätigungsweise des Denkens begründeten Thatsachen beruht[1]).

So ist es z. B. eine in der Natur des Denkens begründete Thatsache, dass jedes Denkobject eine Einheit darstellt und zu anderen Denkobjecten in Beziehung steht. Denn das Denken ist nun einmal ein einheitlich erfassendes und Beziehungen setzendes Denken. Demnach haben die beiden Urtheile: »jedes Denkobject ist eine Einheit und steht zu anderen Denkobjecten in Beziehung« die Bedeutung von Axiomen. Sie können nicht deductiv erkannt werden, da das »einheitliche Erfassen« und »in Beziehung setzen« die Entstehung der

1) Es sind daher auch die Axiome der Mathematik inductiv erkannte Wahrheiten. Wenn hierin eine Bestätigung des von Wundt (Logik, II. 1. Abth., S. 118) ausgesprochenen Satzes liegt, dass alle mathematischen Axiome durch Induction entstanden seien, so wird doch keineswegs, wie aus den obigen Darlegungen hinreichend erhellt, zugestanden, dass die inductive Erkenntniss dieser Axiome auf Erfahrung sich gründe. Die Bethätigungsweise des Denkens ist vielmehr die Grundlage. — Die Entwicklung der Zahlenreihe aus der Reihenform des Denkens und die Ableitung der sogenannten allgemeinen Zahlen (der positiven und negativen, der beliebig theilbaren oder rationalen und irrationalen, der gewöhnlichen und höheren complexen Zahlen) aus bestimmten Formen des beziehenden, nach Grund und Folge verknüpfenden Denkens habe ich in meinen »Untersuchungen über die Grundlagen der Mathematik« Philosophische Studien, XI und XIV) gegeben.

Denkobjecte und Begriffe erst ermöglicht und die Voraussetzung für jede Deduction bildet. Sie beruhen aber auch nicht auf der Erfahrung, da es nicht am Gegebenen sondern am Denken liegt, wenn die Objecte einheitlich erfasst und auf andere Objecte bezogen werden.

Beruht hingegen das Zusammen der Denkobjecte A und B auf der Erfahrung und nicht auf der Art und Weise des Denkens, so ist es denkbar, dass die Unterlage von A nicht auch die Unterlage von B sei. Denn es könnte ja das Gegebene anders beschaffen sein, als es thatsächlich beschaffen ist. Die empirisch gegebene Verbindung von A und B ist darum nicht denknothwendig: sie ist kein Axiom. Dies könnte nur dann bestritten werden, wenn das Gegebene — seinem Wesen widersprechend — kein Gegebenes wäre, sondern eine freie Schöpfung des Denkens. Dann allerdings könnte das Gegebene nur die Formen des Denkens in objectiver Ausgestaltung zur Schau tragen und diese Ausgestaltungen wären so unwandelbar wie die Formen selbst, an die das Denken gebunden ist. Denn andere Formen als die, in welchen wir denken, sind für uns nicht denkbar.

Anderseits muss jedoch die Verbindung von A und B denkmöglich, d. h. mit der bereits erworbenen Erkenntniss und dem bis jetzt vorhandenen Begriff von A insbesondere verträglich sein. Denn der Trieb nach Erkenntniss gehört zum Wesen des Menschen. In Folge dieses Triebes fordert das menschliche Denken die Erkennbarkeit des gesammten Seins und Geschehens, das Gegenstand des Bewusstseins wird. Es besteht daher auch das Bestreben, die empirisch gegebene Verbindung von A und B in den Zusammenhang erkannter Beziehungen aufzunehmen. Erscheint dies nicht möglich und ist trotzdem jene Verbindung unleugbar vorhanden, so kann dieser dem denkenden Menschen unerträgliche Zwiespalt nur durch die Unvollkommenheit der bisherigen Erkenntniss veranlasst sein. Und das Denken kommt nicht zur Ruhe, bis ein System von Erkenntnissen gewonnen ist, dem sich auch die neue, in jener Verbindung von A und B bestehende Thatsache widerspruchslos einfügt.

Ist diese Bedingung erfüllt, ist also die Verbindung von A und B denkmöglich (ohne jedoch denknothwendig zu sein), so kann man nur annehmen oder vermuthen, dass auch die künftige Erfahrung zu keinen Widersprüchen führe. Die mit Rücksicht auf die vorhandene

Stufe des Erkennens mögliche und gebotene Auffassung und Verwerthung jener Thatsache kann daher nur als Vermuthung, als Hypothese sich behaupten, wofern sie durch andere, neue, bis jetzt unbekannte Thatsachen nicht widerlegt wird. Sie gilt somit nicht unabhängig von der Erfahrung, sondern auf Grund der Erfahrung, gefordert durch den Trieb nach der Erkenntniss des Gegebenen.

Den inductiv erkannten Axiomen der Lehre vom Denken im allgemeinen und der Mathematik insbesondere treten so die gleichfalls inductiv erkannten Hypothesen der auf Erfahrung gegründeten Natur- und Geisteswissenschaften zur Seite.

§ 5. Empirische Gewissheit und Wahrscheinlichkeit.

Ist das von der Erfahrung an einer gegebenen Unterlage dargebotene Zusammen von A und B als eine für das Begreifen von A wesentliche Zusammengehörigkeit aufzufassen, so tritt, da die Beziehung $(A; B)$ dem vorliegenden Begriffe von A nicht zugehört, der veränderte, aus A abgeleitete Denkgegenstand A' an die Stelle von A. A' wird durch die neue Beziehung im Vereine mit den für A bereits geltenden Beziehungen logisch bestimmt. Der Begriff von A' oder — wie man auch sagen kann — der Begriff des veränderten, näher bestimmten A hat nun, wie schon bemerkt wurde, entweder den nämlichen Umfang wie der vorliegende Begriff von A oder einen kleineren Umfang, so dass die Aufnahme von $(A; B)$ in den Kreis der schon erkannten Beziehungen entweder eine Ergänzung oder eine Verengerung des vorliegenden Begriffs herbeiführt.

Das Uebersehen der an zweiter Stelle genannten Möglichkeit bedingt einen wesentlichen Mangel in der herkömmlichen Auffassung des inductiven Erkennens. Es wird nämlich bloß der Fall der, von mir so genannten, Begriffsergänzung berücksichtigt, so dass es sich nur darum handelt, aus der Beobachtung aller oder einzelner Erscheinungsformen eines Denkgegenstandes oder — wie man zu sagen pflegt — eines Begriffs eine für den ganzen Begriffsumfang gültige nähere Bestimmung zu gewinnen.

Dies gilt zunächst von der sogenannten vollständigen Induction des Aristoteles, die übrigens streng genommen gar nicht als Induction in dem hier vertretenen Sinne in Anspruch genommen werden darf, sofern — wie Sigwart[1]) hervorhebt — »Aristoteles, wo er von Induction redet, kaum jemals daran denkt, aus der Beobachtung von Einzelfällen im eigentlichen Sinne einen allgemeinen Satz abzuleiten. Seine Beispiele beziehen sich meist auf die Speciesbegriffe, und er fasst nicht Einzelthatsachen zu einem untersten Begriffe, sondern speciellere Begriffe zu einem allgemeineren zusammen, beziehungsweise specielle Regeln zu einer allgemeinen«. Der Aristotelischen Induction ist es somit nicht wesentlich, an die Unterlage der Denkgegenstände anzuknüpfen; sie kann vielmehr auch in der Uebertragung einer aus den specielleren Begriffen deductiv gewonnenen Erkenntniss auf den allgemeinen Begriff bestehen. Man dürfte daher, wenn man dies Verfahren als wirkliche und wahre Induction bezeichnen wollte, nicht betonen, dass es Thatsachen sind, die der inductiven Erkenntniss zu Grunde liegen, sondern man müsste z. B. auch den Schluss, dass eine Eigenschaft, die vom spitzwinkligen, rechtwinkligen und stumpfwinkligen Dreieck nachweisbar gilt, nun auch vom allgemeinen Dreiecke gelte, als Induction anerkennen; dies würde dazu führen, eine Zusammenfassung von Deductionen als Induction zu bezeichnen. — Aber auch dann, wenn die vollständige Induction von der Unterlage der Denkobjecte ausgeht, handelt es sich stets um eine für den Begriff in seinem ganzen Umfange gültige Erkenntniss oder um eine Begriffsergänzung. Denn sie beruht auf dem Grundsatze, dass ein mit jeder Unterlage eines Denkgegenstandes verknüpftes Merkmal dem Denkgegenstande selbst oder dem Begriffe des Denkgegenstandes zugehöre.

Dies gilt nicht minder von der sogenannten unvollständigen Induction. Sie hat zwar als Induction in dem hier vertretenen Sinne zu gelten, da sie in der That an die Unterlage der Denkgegenstände anknüpft. Dass aber wiederum nur eine für den ganzen Begriffsumfang gültige Erkenntniss in Betracht kommt, geht schon aus der üblichen, schematischen Darstellung des Inductionsschlusses hervor, die beispielsweise B. Erdmann[2]) in folgender Form gibt:

---

1) Logik, II. S. 406.

2) Zur Theorie des Syllogismus und der Induction, Philosophische Aufsätze, Ed. Zeller gewidmet, 1887, S. 207. — Die Beschränkung auf Begriffsergänzung folgt auch aus der Regel, die Apelt (Theorie der Induction, 1854, S. 17) gibt:

$$\begin{array}{l} S_1 \text{ ist } P \\ S_2 \text{ ist } P \\ S_3 \text{ ist } P \\ \quad \cdot \quad \cdot \quad \cdot \\ \quad \cdot \quad \cdot \quad , \\ \quad \cdot \quad \cdot \quad \cdot \\ \hline \text{alle } S \text{ werden } P \text{ sein.} \end{array}$$

Denn hier sind S_1, S_2, S_3 ... oder alle S nichts anderes als die den Begriffsumfang bestimmenden Substrate des gegebenen Denkgegenstandes A, während P den auf A zu beziehenden Denkgegenstand B bezeichnet.

Es ist aber klar, dass eine vollständige Theorie der auf Erfahrung gegründeten Induction beide Möglichkeiten berücksichtigen muss: den Fall der Begriffsverengerung ebensowohl wie den Fall der Begriffsergänzung. Denn beide Fälle kommen vor, und jeder kann einen Fortschritt in der Erkenntniss des gegebenen A mit sich führen.

Der Erfolg der Induction besteht jedenfalls dann in einer Begriffsergänzung, wenn die in dem Wissensgebiete, dem A angehört, vorhandene Erkenntniss dazu nöthigt, eine Beziehung zwischen A und einem noch unbekannten Denkgegenstande vorauszusetzen, und es sich nur um die Feststellung dieses Denkgegenstandes handelt. Dann ist möglicher Weise eine einzige Beobachtung hinreichend, um das an irgend einer Unterlage von A thatsächlich gefundene Zusammen mit B als eine ohne Rücksicht auf die besondere Beschaffenheit der Unterlage gültige und somit für A schlechthin bestehende Zusammengehörigkeit aufzufassen. Fehlt eine solche besondere vorbereitende Erkenntniss, so muss ein allgemein gültiger Grundsatz der empirischen Forschung (z. B. Causalität, psychophysischer Parallelismus), der in

---

»Was von den Theilen einer Sphäre gilt, das gilt auch von dem Begriff selbst, in dessen Sphäre diese Theile stehen«, und ebenso aus der Definition Gneiße's (Deduction und Induction, eine Begriffsbestimmung, 1899, S. 23): »Induction ist das Denkverfahren, das die einzelnen Gegenstände eines Begriffs wie diesen selbst durch ein mit ihm nothwendig verknüpftes Merkmal näher bestimmt, indem es dieses Merkmal an allen beobachteten Erscheinungen desselben Begriffs als stetig nachweist«.

letzter Instanz in dem Triebe nach Erkenntniss wurzelt, dazu nöthigen, die einmal oder mehrmals gemachte Beobachtung als für A schlechthin gültig anzuerkennen.

Die so erkannte Beziehung (A; B) ist zwar eine auf Erfahrung gegründete Hypothese, nicht ein in der Art und Weise des Denkens begründetes Axiom. Sie wird aber durch den Trieb nach Erkenntniss gefordert, so dass es nicht im Belieben des denkenden Menschen steht, sie anzuerkennen oder zu verwerfen. Sie muss vielmehr anerkannt werden, wofern man nicht auf ein Erkennen des empirisch gegebenen Seins und Geschehens überhaupt verzichten will. Die Beziehung zwischen A und B ist darum gewiss, wenn schon die Gewissheit keine absolute, sondern eine relative, d. h. auf Erfahrung gegründete ist. Die Erkenntniss der empirischen Gewissheit, dass (A; B) besteht, bildet somit in diesem Falle das Ergebniss der Induction[1]).

Liegt aber keine Nöthigung vor, das in einem Falle oder in mehreren Fällen beobachtete Zusammen von A und B als eine für A schlechthin bestehende Zusammengehörigkeit aufzufassen, so hat man auch nicht das Recht, von einer für A ohne Rücksicht auf dessen Unterlage gültigen Beziehung zu sprechen. Man kann nur annehmen, dass es eine vorläufig noch unbekannte Beschaffenheit der Unterlage von A gibt, für welche die Beziehung zu B Geltung hat. Könnte man diese Beschaffenheit kennen lernen, so erhielte man hierdurch den durch die Bezugnahme auf B aus A entstehenden Denkgegenstand A' in voller Bestimmtheit; und wäre anderseits A' als Denk-

1) Die relative, auf Erfahrung beruhende oder empirische Gewissheit als Wahrscheinlichkeit zu bezeichnen, wäre mit Rücksicht auf den allgemein anerkannten Wahrscheinlichkeitsbegriff der Mathematik unzweckmäßig. Falls aber doch die Unterscheidung Apelt's (Theorie der Induction, S. 38) zwischen philosophischer und mathematischer Wahrscheinlichkeit Billigung finden sollte, so müsste man beachten, dass das, was dort philosophische Wahrscheinlichkeit bedeutet, hier empirische Gewissheit heißt. Denn Apelt sagt (S. 40): »Der philosophische Wahrscheinlichkeitsschluss schließt geradezu von der Vielheit der Fälle auf die Einheit und Allgemeingültigkeit der Regel« und »Beim philosophischen Wahrscheinlichkeitsschluss ist der Schluss vollkommen sicher, ohnerachtet seiner unvollständigen Begründung«. Die Berechtigung zum Schluss geben aber die Maximen oder Grundsätze der empirischen Forschung. — Der philosophische Wahrscheinlichkeitsschluss Apelt's führt somit in der That zu der nach obiger Bezeichnungsweise mit empirischer Gewissheit verknüpften Begriffsergänzung.

gegenstand gegeben, so würde für ihn die Beziehung zu B als empirisch gewiss zu gelten haben. Für A selbst kann aber keine Gewissheit erreicht werden. Denn es lässt sich nichts anderes feststellen, als dass A ebensowohl mit B verbunden sein als auch ohne B vorkommen kann. Das eine wie das andere ist als denkmöglich anzuerkennen: die Möglichkeit der Verbindung von A und B folgt aus der beobachteten Thatsache; die Möglichkeit des Vorkommens von A ohne B folgt aus dem mangelnden Zwang, A und B ohne Rücksicht auf die Unterlage von A verbunden zu denken. Die Erkenntniss der beiden Möglichkeiten ist hier das Ergebniss der Induction.

Hat man die beiden Möglichkeiten erkannt, ohne zugleich zu erkennen, wie die Unterlage von A für den Fall einer Verbindung mit B beschaffen sein muss, so weiß man nicht, ob in einem künftig eintretenden Falle bei einer bestimmten Unterlage von A auch B zu erwarten sei oder nicht. Denn nur dies ist gewiss, dass A entweder mit B verbunden sein oder ohne B vorkommen wird. Man sagt dann, dass eine gewisse Wahrscheinlichkeit für die Verbindung von A mit B bestehe, und man spricht von einem größeren oder geringeren Grade der Wahrscheinlichkeit, wenn man einen hinreichenden Grund zu haben glaubt, um eine größere oder geringere Häufigkeit des Auftretens der Verbindung von A und B vorauszusetzen.

Es erwächst so die weitere Aufgabe, die Wahrscheinlichkeit für das Auftreten der Verbindung von A und B zu bestimmen. Kann diese Bestimmung auf Grund der Erfahrung ausgeführt werden, so besteht der Erfolg der Induction in der Erkenntniss der empirischen Wahrscheinlichkeit für die Zusammengehörigkeit von A und B.

Bevor jedoch die Bestimmungsweise der empirischen Wahrscheinlichkeit erörtert werden soll, mögen noch folgende Beispiele den Gegensatz zwischen inductiver, auf Thatsachen der Erfahrung beruhender Gewissheitserkenntniss einerseits und Wahrscheinlichkeitserkenntniss anderseits erläutern.

Als Kepler den Weg der Erfahrung betrat, um zur Erkenntniss der Bahn der Planeten, zunächst derjenigen des Planeten Mars zu gelangen, war er im voraus gewiss, dass eine bestimmte, allen Planeten gemeinsame Gesetzmäßigkeit vorhanden sei. Die Erkenntniss, dass eine Beziehung zwischen Planet und Planetenbahn existire, war somit vorhanden; nur die Gestalt dieser Bahn war unbekannt und konnte nicht aus dem Begriff des Planeten deducirt werden. Es musste darum eine möglichst einfache Curve gesucht werden, auf der die beobachteten und berechneten Planetenörter liegen konnten. Nachdem in der Ellipse diese Curve gefunden war, blieb kein Zweifel, dass die Planeten nicht bloß für den Zeitraum, dem die Beobachtungen angehörten, sondern immer eine elliptische Bahn beschrieben haben und auch in Zukunft beschreiben werden, dass mithin die elliptische Gestalt der Bahn ein für jeden Planeten gültiges, den Begriff des Planeten ergänzendes Merkmal sei. — Wäre hingegen bei irgend einer Stufe des astronomischen Erkennens die Annahme denkbar, dass die Planeten lebende Organismen seien mit der Befähigung, ihren Lauf willkürlich zu bestimmen oder wenigstens innerhalb gewisser Grenzen abzuändern, so würde aus der Bestimmung, dass zu dieser oder jener Zeit die Planetenbahn eine Ellipse war, nur gefolgert werden können, dass unter anderen Bahncurven auch die elliptische möglich ist. Und auch eine oft wiederholte Feststellung der jeweiligen Bahncurve könnte nur erkennen lassen, ob die elliptische Bahn häufig oder selten vorkommt und welche Wahrscheinlichkeit für diese Curvenform empirisch besteht. Ja selbst dann, wenn ausnahmslos die Ellipse aus den immer und immer wieder angestellten Beobachtungen sich ergeben würde, könnte man keine empirische Gewissheit, sondern bloß einen der Gewissheit nahe kommenden, hohen Grad empirischer Wahrscheinlichkeit für das Auftreten der elliptischen Curve erhalten.

Handelt es sich anderseits um die Feststellung des physischen Vorgangs, der (nach dem Princip des psychophysischen Parallelismus) einer im Bewusstsein vorhandenen Empfindung als objectives Merkmal zur Seite steht, so versteht es sich von selbst, dass nur ein solcher Vorgang brauchbar ist, der sich der Empfindung eindeutig zuordnen lässt, so dass einer Veränderung der Empfindung eine bestimmte Veränderung des objectiven Thatbestandes entspricht. Dies setzt überdies voraus, dass nicht noch ein anderer objectiver Vorgang in ähnlicher eindeutiger Beziehung zu jener Empfindung steht. Wünscht man also beispielsweise das Kennzeichen für die objective oder physikalische Wärme zu finden, so muss sich an dem die

Wärmeempfindung erregenden Naturobjecte ein Merkmal angeben lassen, das sich mit der stärker oder schwächer werdenden Empfindung eindeutig ändert. Zeigt nun die Beobachtung, dass ein Körper, der eine Zunahme der Wärmeempfindung veranlasst, sich ausdehnt, sonst aber keine Veränderung in seiner objectiven Beschaffenheit erleidet, so ist es gewiss, dass die Ausdehnung des Körpers das objective Merkmal der Wärmezunahme ist; und man kann einen Normalkörper herstellen, der durch sein Volumen den physikalischen Wärmezustand kennzeichnet und so als Thermometer dient. — Ist aber nicht eine Empfindung sondern ein Gefühl, das den subjectiven Gemüthszustand charakterisirt, gegeben, so kann zwar wiederum ein bestimmter äußerer Vorgang als Erreger jenes Gefühls vorhanden sein. Es zeigen sich ja in der That die Gefühle der Lust und der Unlust vielfach an äußere Vorgänge geknüpft. Ein eindeutiges, mit empirischer Gewissheit zutreffendes Entsprechen eines bestimmten äußeren Vorgangs und eines bestimmten Gefühls wird man aber nicht annehmen, eben weil die Gefühle im Unterschied von den Empfindungen nicht als Zeichen für die Beschaffenheit der Objecte, sondern als Zeichen für den Zustand des Subjects dienen und der Zustand des Subjects nicht ausschließlich durch die jeweiligen äußeren Vorgänge bedingt wird. Man wäre daher selbst dann, wenn ausnahmslos, soweit die bisherige Erfahrung reicht, mit einem bestimmten äußeren Vorgange immer der nämliche Gefühlszustand sich verknüpft zeigen sollte, doch nicht zu der Ueberzeugung berechtigt, dass dies auch künftighin und bei jedermann so sein werde. Man müsste vielmehr in Rechnung ziehen, dass eine Ausnahme der vorläufig durchweg bestätigten Regel sehr wohl mit der sonstigen Erkenntniss des Seelenlebens vereinbar ist, und man dürfte aus der bisherigen Erfahrung nur dies mit Fug und Recht schließen, dass ein hoher Grad von Wahrscheinlichkeit für das Zusammentreffen jenes Gefühls und jenes äußeren Vorgangs besteht.

Die Theorie der Induction hat demzufolge in Rücksicht zu ziehen, dass Erfahrungsthatsachen je nach der in dem zugehörigen Wissensgebiete erreichten Stufe des Erkennens und der Anwendbarkeit allgemein gültiger, auf dem Erkenntnisstriebe beruhender Grundsätze wissenschaftlicher Forschung entweder zur Erkenntniss empirischer Gewissheit oder zur Einsicht in verschiedene Möglichkeiten, mit günstigen Falls bestimmbaren Wahrscheinlichkeiten, führen.

Dadurch wird zugleich die Frage Mill's[1]) beantwortet, warum in manchen Fällen eine einzige Instanz zu einer vollständigen Induction ausreiche, während in anderen Fällen Myriaden übereinstimmender Instanzen, ohne eine einzige bekannte oder vermuthete Ausnahme, so sehr wenig dazu beitragen, einen Satz von durchgängiger Allgemeinheit zu begründen. — Wenn nämlich Mill die beiden Fälle anführt, dass einerseits ein Chemiker keinem Zweifel begegne, falls er auch nur auf Grund eines einzigen, richtig angestellten Versuches Dasein und Eigenschaften einer neu entdeckten Substanz behauptet, während anderseits trotz ausnahmsloser Bestätigung des Satzes, dass Krähen schwarz sind, doch das Vorkommen einer grauen Krähe für möglich gehalten werden müsse, so ist im Einklange mit den vorstehenden Darlegungen klar, dass im ersten Falle eine wissenschaftliche Chemie nicht möglich wäre, wenn man die empirisch gefundenen Eigenschaften einer Substanz nicht als zum Begriff der Substanz gehörig anerkennen wollte, wogegen im zweiten Falle die zoologische Erkenntniss es ausschließt oder wenigstens nicht fordert, dass die Farbe der Krähe als zum Begriffe der Krähe gehörig aufgefasst werde. Darum ist es empirisch gewiss, dass eine bestimmte Substanz immer die nämlichen Eigenschaften hat, während günstigsten Falles bloß der Grad der Wahrscheinlichkeit ermittelt werden kann, der für die schwarze Farbe der Krähe besteht.

Auch ist nachdrücklich hervorzuheben, dass die Anzahl beobachteter Fälle nur dann von Bedeutung sein kann, wenn es sich um inductive Wahrscheinlichkeitserkenntniss handelt. Bei der üblichen Beschränkung der Induction auf Begriffsergänzung hingegen darf die Einführung der neuen Beziehung in den vorhandenen Begriff nicht durch den Hinweis auf die Häufigkeit zutreffender und den Mangel widersprechender Beobachtungen begründet werden. Selbst die bescheidene Forderung von mindestens zwei Beobachtungen ist unzulässig. Denn die Anzahl der Fälle als solche hat auf die Gewissheit der Erkenntniss keinen Einfluss, so wesentlich auch eine

---

1) John Stuart Mill, System der deductiven und inductiven Logik (übersetzt von Gomperz), I. S. 367.

Wiederholung der Beobachtung für die Sicherstellung des gefundenen Resultates sein mag[1]).

§ 6. Wahrscheinlichkeitsbestimmung.

Sieht man sich zu der Annahme genöthigt, dass A sowohl mit B vereinigt sein als auch ohne B vorkommen kann, so gilt es, die für beide Möglichkeiten bestehenden Wahrscheinlichkeiten zu bestimmen. Es muss darum klargestellt werden, worauf die Wahrscheinlichkeitsbestimmung abzielt und wie sie geleistet werden kann.

a. Die Definition der Wahrscheinlichkeit.

Als wahrscheinlich bezeichnet man ein Ereigniss, wenn es nicht mit Gewissheit erwartet wird, sondern — der vorhandenen Erkenntniss gemäß — ebensowohl eintreten als auch ausbleiben kann. Es ist ferner das Eintreten eines Ereignisses wahrscheinlicher als das Ausbleiben, wenn — auf Grund der vorhandenen Erkenntniss — das Eintreten häufiger zu erwarten ist als das Ausbleiben. Man kann schließlich die Wahrscheinlichkeit nicht bloß als ein unbestimmtes »mehr oder minder«, sondern zahlenmäßig bestimmen, wenn man zu der Erkenntniss gelangt, dass unter einer Gesammtzahl von t Fällen das in Frage stehende Ereigniss r-mal eintreten und s-mal ausbleiben wird. Es bezeichnen alsdann die Quotienten

$$p = \frac{r}{t} \text{ und } q = \frac{s}{t},$$

wo $r + s = t$ und $p + q = 1$, die Wahrscheinlichkeiten für das Eintreten und für das Ausbleiben des Ereignisses.

---

1) So hebt auch Sigwart (Logik, II., 2. Aufl. S. 430) hervor, »dass die Zahl der Fälle, aus denen ein allgemeiner Satz gewonnen wird, keinen fundamentalen Unterschied in dem logischen Processe begründet, der dabei stattfindet, und dass der Charakter des letzteren verhüllt wird, wo die Zusammenfassung einer Anzahl gleichartiger Fälle als sein wesentliches Moment gelten soll.« — Wenn B. Erdmann (Zur Theorie des Syllogismus und der Induction; Philos. Aufs. zu Zeller's Jubiläum, S. 211) sich damit nicht einverstanden erklärt und den inductiven Schluss auf die Zusammengehörigkeit von A und B darauf gründet, dass die bisherige Erfahrung einen constanten Zusammenhang zwischen A und B gezeigt habe, so ist darauf hinzuweisen, dass die obigen Beispiele das Unzulängliche eines solchen Schlusses zeigen.

Unter Wahrscheinlichkeit ist somit die relative Häufigkeit eines Ereignisses zu verstehen, mag nun dieselbe nur als ein unbestimmtes »mehr oder minder« oder als ein positiver echter Bruch bestimmbar sein.

Will man also die Wahrscheinlichkeit für die Verbindung von A mit B ermitteln, so ist die relative Häufigkeit für das Auftreten dieser Verbindung zu bestimmen.

Da diese Bestimmung einen Fortschritt in der Erkenntniss von A einschließen soll, so kann es sich dabei nicht um subjectives Meinen und Dafürhalten handeln. Auch wäre es werthlos, im Bewusstsein mangelnder Einsicht zu sagen, dass man keinen Grund habe, die eine Möglichkeit für wahrscheinlicher zu halten als die andere, und darum die Wahrscheinlichkeit beider Möglichkeiten als gleich groß, d. h. jede gleich $\frac{1}{2}$ annehmen wolle. Es ist vielmehr eine dem Denkgegenstande A wirklich zugehörige Bestimmung erforderlich, die ebenso wie die sonstigen, zum Begriffe oder zum Bereiche von A gehörenden Bestimmungen in der Beschaffenheit von A, nicht aber im subjectiven Ermessen begründet ist.

Nun hat man bereits erkannt, dass A je nach Umständen bald zusammen mit B bald ohne B vorkommen kann. Dies setzt voraus, dass A einer doppelten näheren Bestimmung fähig ist und entweder als A' oder als A'' auftreten kann, wo A' durch die Verbindung mit B, A'' durch das Fehlen von B ausgezeichnet ist. Es zerfällt somit der Begriffsumfang von A in die beiden Begriffsumfänge von A' und A''. Der Begriffsumfang von A bezeichnet aber alle Möglichkeiten für das Auftreten von A; denn er wird durch die Mannigfaltigkeit aller Bewusstseinsinhalte, die als Unterlage von A möglich sind, gebildet. Ebenso bezeichnet der Begriffsumfang von A' resp. A'' alle Möglichkeiten für das Auftreten von A' resp. A'', da er aus der Gesammtheit aller möglichen Unterlagen von A' resp. A'' besteht. Man hat demgemäß anzunehmen, dass A häufiger oder seltener mit B verbunden als ohne B vorkommen werde, je nachdem der Begriffsumfang von A' größer oder kleiner als der Begriffsumfang von A'' ist. Lassen sich überdies die Begriffsumfänge auszählen oder (wenn sie continuirliche Mannigfaltigkeiten darstellen) ausmessen und findet man auf diese Weise die Anzahl oder Maßzahl der Möglichkeiten für das

Auftreten von A' und A'' gleich r und gleich s, wonach für A selbst $r+s$ oder t Möglichkeiten bestehen, so wird die Wahrscheinlichkeit p für das Auftreten der Verbindung von A und B durch

$$p=\frac{r}{t}$$

und die Wahrscheinlichkeit q für das Ausbleiben dieser Verbindung durch

$$q=\frac{s}{t}$$

bezeichnet. Aber auch dann, wenn von den Zahlen r, s, t nicht jede für sich angebbar ist, kann es doch möglich sein festzustellen, den wie vielten Theil des Begriffsumfangs von A die Begriffsumfänge von A' und A'' ausmachen; und diese Bruchwerthe bestimmen alsdann die Wahrscheinlichkeiten p und q.

Das Urtheil über die Wahrscheinlichkeit oder relative Häufigkeit des Auftretens der Verbindung von A und B muss sich demnach auf den Vergleich von Begriffsumfängen stützen, wenn es eine objectiv gültige Erkenntniss aussagen soll. Dabei gilt folgende Bestimmung[1]):

Ist der gegebene Denkgegenstand A entweder als das mit B verbundene A' oder als das ohne B vorkommende A'' zu denken, so ist die Wahrscheinlichkeit für das Auftreten der Verbindung von A und B gleich dem durch den Begriffsumfang von A' dargestellten Bruchtheil des Begriffsumfangs von A.

1) Während hier die Wahrscheinlichkeitsbestimmung ganz allgemein auf den Vergleich von Begriffsumfängen zurückgeführt wird, mögen dieselben nun durch endliche oder unendlich große, discrete oder continuirliche Mannigfaltigkeiten dargestellt werden, findet J. von Kries (Die Principien der Wahrscheinlichkeitsrechnung, eine logische Untersuchung, 1886) die Wahrscheinlichkeitsbestimmung (S. 24) daran gebunden, dass »unsere Annahmen einen Gegenstand betreffen, für welchen, unserm Wissen gemäß, ein messbarer und in Theile zu zerlegender Spielraum des Verhaltens möglich erscheint«, und er gelangt (S. 36) zu dem Satz dass Annahmen in einem zahlenmäßig angebbaren Wahrscheinlichkeitsverhältniss stehen, wenn sie indifferente und ihrer Größe nach vergleichbare Spielräume umfassen, und dass bestimmte Wahrscheinlichkeitswerthe sich daher da ergeben, wo die Gesammtheit aller Möglichkeiten durch eine Anzahl solcher Annahmen erschöpft werden kann«.

b. Die deductive Bestimmung der Wahrscheinlichkeit.

Die Kenntniss dieses Bruchwerthes kann durch Deduction oder Induction gewonnen werden, wonach man deductive und inductive Wahrscheinlichkeitserkenntniss[1]) zu unterscheiden hat.

Gehört zum Begriff von A die Einsicht, dass A entweder als das mit B verbundene A' oder als das ohne B auftretende A'' sich darstellt, und sind außerdem die Begriffsumfänge von A sowohl wie von A' und A'' bekannt, so dass eine exacte Bestimmung oder wenigstens näherungsweise eine Abschätzung der Möglichkeiten für A' im Vergleich zu den Möglichkeiten für A ausführbar ist, so gelangt man auf deductivem Wege zur exacten oder angenäherten Kenntniss der Wahrscheinlichkeit für die Verbindung von A und B. Die Beziehung zwischen A und dem gefundenen Wahrscheinlichkeitswerthe gründet sich auf den Begriff von A und gehört zum logischen Bereich von A.

Als Beispiel denke man sich eine Urne, in die weiße und schwarze Kugeln gelegt sind. Dann wird der Denkgegenstand A durch jede Kugel der Urne und insbesondere der mit B (der weißen Farbe) verbundene Denkgegenstand A' durch jede weiße Kugel und demnach A'' durch jede schwarze Kugel dargestellt. Die Begriffsumfänge von A, A und A'' werden ferner durch die Anzahlen aller Kugeln, der weißen Kugeln und der schwarzen Kugeln bestimmt. Die Wahrscheinlichkeit oder relative Häufigkeit einer weißen Kugel in der Urne ist somit bekannt, wenn die Anzahl der weißen und der schwarzen

1) Es ist unmittelbar klar, dass hier die in der Wahrscheinlichkeitsrechnung sogenannten Wahrscheinlichkeiten a priori und a posteriori als deductiv erkannte und inductiv erkannte Wahrscheinlichkeiten bezeichnet werden. — Für andere Unterscheidungen, etwa für diejenige zwischen objectiver und subjectiver Wahrscheinlichkeit, die man bei A. Meyer (Vorlesungen über Wahrscheinlichkeitsrechnung; deutsch von Czuber, 1879, S. 6) findet, besteht kein genügender Grund. Denn jede Wahrscheinlichkeitsbestimmung ist, wenn sie einen Erkenntnisswerth besitzen soll, nothwendig objectiv, sofern sie in der Natur des Objectes, das der Wahrscheinlichkeitsbestimmung unterliegt, begründet sein muss; sie ist aber zugleich subjectiv, da sie stets »von dem Zustande unserer Kenntnisse« abhängt (was a. a. O. als charakteristisches Merkmal der subjectiven Wahrscheinlichkeit hervorgehoben wird). Eine rein subjective Wahrscheinlichkeitsbestimmung, die ein bloßes Meinen ohne objective Begründung zum Ausdruck brächte, wäre, wie schon weiter oben bemerkt wurde, ohne Erkenntnisswerth.

Kugeln oder wenigstens ihr Verhältniss bekannt ist. Beispielsweise ist die Wahrscheinlichkeit für eine weiße Kugel gleich $^3/_5$, wenn 500 Kugeln in der Urne und darunter 300 weiße sind, oder auch wenn unendlich viele Kugeln vorhanden sind, die sich in Gruppen von je fünf Kugeln mit je drei weißen Kugeln ordnen lassen. — Handelt es sich hingegen um die Wahrscheinlichkeitsbestimmung für das Ziehen einer weißen Kugel aus der Urne (wo alsdann bei einer endlichen Anzahl von Kugeln jede gezogene Kugel zurückgelegt und mit den anderen vermischt zu denken ist), so ist zwar wiederum klar, dass jeder Zug A entweder als das Ziehen einer weißen Kugel, als A', oder als das Ziehen einer schwarzen Kugel, als A'' zu denken ist; auch pflegt man ohne weiteres in der Wahrscheinlichkeitsrechnung die Wahrscheinlichkeit für das Ziehen einer Kugel von bestimmter Farbe der Anzahl vorhandener Kugeln mit der betreffenden Farbe proportional zu setzen. Dies ist aber an die Bedingung geknüpft, dass die Verschiedenartigkeit der Lage der Kugeln in der Urne und des Herausnehmens derselben aus der Urne, die in der That unendlich groß ist, zu keiner Bevorzugung einzelner Kugeln führt, dass vielmehr bei einer unbegrenzten Wiederholung des Ziehens jede einzelne Kugel verhältnissmäßig gleich häufig getroffen wird. Diese Voraussetzung ist unentbehrlich; denn an ein wirkliches Abzählen aller möglichen Fälle für das Ziehen einer weißen oder einer schwarzen Kugel oder einer Kugel überhaupt — an die Bestimmung der Begriffsumfänge von A', A'' oder A ist hier nicht zu denken.

c. Die Methoden zur inductiven Bestimmung der Wahrscheinlichkeit.

Ist die Wahrscheinlichkeitsbestimmung nicht auf deductivem Wege möglich, hat man also keine, mit dem vorhandenen Begriff von A mittelbar oder unmittelbar zusammenhängende Kenntniss der Begriffsumfänge von A, A', A'', so kann man bloß auf dem Wege der Erfahrung zu dieser Kenntniss gelangen. Die Erfahrung lehrt nämlich, ob an der jeweiligen Unterlage des gegebenen A sich zugleich B zeigt oder nicht, ob also das A in der gerade vorliegenden Gestalt ein A' oder A'' ist. Es ist nun aber der Inbegriff aller Möglichkeiten für das Auftreten von A oder der ganze Begriffsumfang von

A durchzugehen und festzustellen, wie oft A' und wie oft A'' auftritt. Der Quotient aus den so ermittelten Anzahlen der Möglichkeiten für A' und der Möglichkeiten für A, d. h. das empirisch bestimmte Verhältniss der Begriffsumfänge von A' und von A stellt die gesuchte Wahrscheinlichkeit oder relative Häufigkeit des Vorkommens von A' oder der Verbindung von A mit B dar. Auf diese Weise wird die Beziehung zwischen A und dem gefundenen Wahrscheinlichkeitswerthe in den Begriff oder in den logischen Bereich von A eingeführt und somit inductiv erkannt.

Diese Art der Wahrscheinlichkeitsbestimmung erfordert ein vollständiges Erschöpfen aller Einzelfälle — wie man es von der sogenannten vollständigen Induction zu fordern gewöhnt ist. Sie ist daher nur dann ohne weiteres ausführbar, wenn die Begriffsumfänge aus endlichen, zähl- oder messbaren Mannigfaltigkeiten bestehen: etwa in dem Falle, wo eine Urne mit einer unbekannten aber endlichen Anzahl schwarzer und weißer Kugeln vorhanden ist und die Kugeln einfach abzuzählen sind, um die Wahrscheinlichkeitswerthe zu ermitteln. Wird aber der Begriffsumfang durch eine unendlich große Mannigfaltigkeit dargestellt, so lässt er sich natürlich nicht erschöpfen; und auch bei einer endlichen Mannigfaltigkeit kann das Erproben aller Möglichkeiten wegen besonderer Umstände (z. B. wegen der Besonderheiten ihrer räumlichen oder zeitlichen Vertheilung) oder auch wegen der praktisch nicht zu bewältigenden großen Anzahl unmöglich werden. So z. B., wenn in der Urne unendlich viele oder überaus viele schwarze und weiße Kugeln in unbekannter Anzahl vorhanden sind, deren vollständiges Abzählen unausführbar ist.

Nun hat man aber nicht ausnahmsweise, sondern in der Regel solche Begriffsumfänge zu erwarten, deren Mannigfaltigkeiten unbegrenzt groß oder wenigstens größer sind, als dass man sie ohne weiteres abzählen könnte. Es kann daher die Frage nicht umgangen werden, ob nicht unter Umständen die Beobachtung eines zur Verfügung stehenden Theils des Begriffsumfanges von A genügen kann, um ein hinreichend sicheres Urtheil über das Verhältniss der Begriffsumfänge von A' und A'', die den Gesammtumfang des Begriffs von A ausmachen, zu gewinnen.

Dass dies nicht unter allen Umständen möglich ist, ergibt sich aus folgender Bemerkung.

Die Wahrscheinlichkeitsbestimmung ist nur von der Anzahl der auf A' und A'' treffenden Fälle, nicht von der Ordnung, in der diese Fälle auftreten, abhängig. Besteht nämlich die Mannigfaltigkeit des Begriffsumfangs von A aus t abzählbaren Elementen, von welchen r den Begriffsumfang von A' und s den Begriffsumfang von A'' bilden, so sind überhaupt

$$t(t-1)(t-2)\ldots 3.2.1 \text{ oder } t!$$

Anordnungen denkbar, indem jedes der t Elemente mit jedem der $t-1$ übrigen und jedes solche Paar wiederum mit jedem der $t-2$ nun noch verfügbaren Elemente combinirt werden kann u. s. w. Bei jeder Anordnung treten aber r Elemente A' und s Elemente A'' auf, deren relative Häufigkeiten durch

$$p = \frac{r}{t} \text{ und } q = \frac{s}{t}$$

angegeben werden. — Werden nun nicht alle t Elemente, sondern bloß m (wo $m < t$) abgezählt, so sind im Ganzen

$$t(t-1)(t-2)\ldots(t-m+1)$$

aus den t Elementen herstellbare Systeme von m Elementen möglich, und diese Systeme zerfallen, da die t Elemente aus r Elementen A' und s Elementen A'' sich zusammensetzen, in folgende $m+1$ Gruppen[1]:

1) $r(r-1)\ldots(r-m+1)$ Systeme aus m Elementen A;

2) $\frac{m}{1}\cdot r\ldots(r-m+2)\cdot s$ Systeme aus $m-1$ Elementen A' und einem Elemente A'';

3) $\frac{m(m-1)}{1\cdot 2}r\ldots(r-m+3)\cdot s(s-1)$ Systeme aus $m-2$ Elementen A' und 2 Elementen A'';

. . . .

m) $\frac{m}{1}r\cdot s(s-1)\ldots(s-m+2)$ Systeme aus einem Elemente A' und $m-1$ Elementen A'';

1 Die mitgetheilten Anzahlen der möglichen Systeme für jede Gruppe folgen unmittelbar aus den Regeln der Combinationslehre.

$m+1)\ s(s-1)\ldots(s-m+1)$ Systeme aus m Elementen A''.

Es sind somit, wenn

$$\frac{m(m-1)(m-2)\ldots(m-i+1)}{1\cdot 2\cdot 3\ldots i} \quad \text{oder} \quad \frac{m!}{i!(m-i)!}$$

durch

$$\binom{m}{i}=\binom{m}{m-i}$$

angedeutet wird, allgemein

$$\binom{m}{i} r\ldots(r-m+i+1)\cdot s\ldots(s-i+1) \text{ oder kurz } C_{m-i;\,i}$$

Systeme aus $m-i$ Elementen A' und i Elementen A'' möglich. Und es können folglich, falls weder r noch s kleiner als m ist, bei einer beliebigen Bestimmung von m Elementen entweder nur Elemente A' oder nur Elemente A'' oder sowohl Elemente A' als auch Elemente A'' in irgend welcher Mischung auftreten. Die Quotienten p und q, die hier das Verhältniss der Elemente A' und A'' zur Gesammtzahl m des abgezählten Systems angeben, werden demgemäß je nach Umständen das eine oder das andere der $m+1$ Werthenpaare

$$p=\frac{m-i}{m};\quad q=\frac{i}{m};\quad (i=0,\ 1,\ 2\ldots m)$$

darstellen, so dass eine sichere Wahrscheinlichkeitsbestimmung in der That unmöglich ist, wenn nicht besondere Bedingungen erfüllt sind.

Diese Bedingungen müssen das Hervortreten solcher Mischungsverhältnisse für die Elemente A' und A'' zur Folge haben, die mit dem Verhältniss der Begriffsumfänge von A' und A'' übereinstimmen, so dass diejenigen Werthe

$$p=\frac{m-i}{m};\quad q=\frac{i}{m}$$

bestimmt werden können, die den gesuchten Werthen

$$p=\frac{r}{t};\quad q=\frac{s}{t}$$

völlig oder nahezu gleich sind.

Ich finde nun zwei Arten solcher Bedingungen.

Es sind nämlich einerseits Bedingungen denkbar, durch welche für einen abzählbaren, aus m Elementen bestehenden Theil des Begriffsumfangs von A eine beliebige Mischung aus Elementen A' und A'' verhindert und gerade diejenige Mischung herbeigeführt wird, die genau oder mit möglichster Annäherung das wirkliche, zwischen den Begriffsumfängen von A' und A'' bestehende Verhältniss wiedergibt.

Es sind aber auch anderseits Bedingungen denkbar, durch welche bei der Bestimmung von einigen oder von vielen, kurz von n aus je m Elementen bestehenden Theilen des Begriffsumfangs von A die verschiedenen möglichen Mischungen aus Elementen A' und A'' zwar nicht zu Gunsten einer einzigen unterdrückt werden, wohl aber in einer solchen gesetzmäßigen Anordnung auftreten, dass diejenige Mischung, die das Verhältniss der Begriffsumfänge von A' und A'' genau oder am besten zum Ausdruck bringt, am häufigsten vorkommt und hierdurch kenntlich gemacht wird.

Die Bedingungen der ersten Art sind erfüllt, wenn der abgezählte Theil des Begriffsumfangs von A dem Gesammtumfange ähnlich ist, so dass er ein treues Abbild des letzteren ist. Dies ist möglich, wenn der Begriffsumfang von A in zwei oder noch mehr Theile zerfällt, die in ihrer Beschaffenheit und somit auch in den Verhältnissen der jedem Theile zukommenden Elemente A' und A'' übereinstimmen. Es ist dann nur dafür Sorge zu tragen, dass das abgezählte System von m Elementen einen solchen, dem Gesammtumfange ähnlichen Theil wirklich darstellt, um die Gewähr zu haben, dass die jenem System entnommenen Werthe

$$p = \frac{m-i}{m}; \quad q = \frac{i}{m}$$

mit den gesuchten Werthen

$$p = \frac{r}{t}; \quad q = \frac{s}{t}$$

ganz oder hinreichend nahe übereinstimmen. — Handelt es sich z. B. um eine Wahrscheinlichkeitsbestimmung für eine Pflanze wie Roggen oder Weizen, so darf man bei übereinstimmender Beschaffenheit des Bodens und des Klimas annehmen, dass das eine Roggen-oder Weizenfeld die nämlichen Verhältnisse darbietet wie das andere; und man kann sich demgemäß auf die Untersuchung eines gerade zur Verfügung stehenden Feldes beschränken. — Gilt ferner die Wahrscheinlichkeitsbestimmung einem gewissen Merkmal, das sich bei den Angehörigen eines Volksstammes mit einer constanten Häufigkeit vorfindet und mit dem Lebensalter sich nicht ändert (beispielsweise das Vorkommen blauer Augen), so genügt das Abzählen für eine bestimmte Altersclasse, etwa für die Schulkinder, um die wahren Werthe p und q mit hinreichender Sicherheit zu finden. Denn die Unterscheidung von Altersclassen führt hier zu Theilen der Gesammtmannigfaltigkeit, die untereinander und mit der letzteren hinsichtlich der gesuchten Verhältnisszahl mit hinreichender Genauigkeit übereinstimmen, weil die Häufigkeit jenes Merkmals constant ist und darum für die Jungen ebenso groß wie für die Alten angenommen werden muss. — Bei der Urne mit weißen und schwarzen Kugeln hingegen dürfte man nur dann mit Fug und Recht einen Theil dem Ganzen ähnlich voraussetzen, wenn man über ein Mischungsverfahren verfügen könnte, durch das in der ganzen Urne ein gleichartiges Gemenge schwarzer und weißer Kugeln erzeugt würde.

Fehlen solche Bedingungen, die ein Zerfallen des Begriffsumfangs in ähnliche Theile herbeiführen, so kann man nicht erwarten, dass ein einzelnes System von m Elementen das in Wirklichkeit zwischen den Begriffsumfängen von A' und A'' bestehende Verhältniss kennen lehrt. Es bleibt daher bloß übrig, außer dem einen Systeme noch andere abzuzählen. Aber auch bei der n-maligen, unter gleichen Umständen ausgeführten Bestimmung eines Systems von m Elementen kann man ohne das Hinzutreten besonderer Bedingungen nur ein regelloses Schwanken zwischen den $m + 1$ Mischungsverhältnissen $(m - i) : i$ erwarten. Denn man weiß zunächst nur, dass $t \ldots (t - m + 1)$ Systeme möglich sind, unter denen $C_{m-i;\,i}$ das Verhältniss $(m - i) : i$ zeigen. Dabei ist noch zu beachten, dass jedes abgezählte System der Gesammtmannigfaltigkeit wieder einverleibt zu denken ist, wenn die Bestimmung aller Systeme unter gleichen Umständen erfolgen soll: wie bei der Urne mit $t = r + s$ weißen und schwarzen Kugeln, aus der immer wieder m Kugeln gezogen werden sollen, nachdem die bereits gezogenen m Kugeln in die Urne zurückgelegt worden sind.

Es kann folglich ein bereits abgezähltes System nochmals zum Vorschein kommen und jedes in unbegrenzter Wiederholung zur Abzählung gelangen. Darum müssen Bedingungen von der an zweiter Stelle genannten Art erfüllt sein. Dann kann aus dem am häufigsten auftretenden Mischungsverhältniss dasjenige Paar von Wahrscheinlichkeitswerthen

$$p = \frac{m-i}{m}; \quad q = \frac{i}{m}$$

berechnet werden, das dem gesuchten Werthenpaare

$$p = \frac{r}{t}; \quad q = \frac{s}{t}$$

völlig oder nahezu gleich ist.

Diese Bedingungen sind insbesondere dann erfüllt, wenn bei der beliebig oft, unter gleichen Umständen wiederholten Bestimmung von je m Elementen aus der Gesammtzahl von $t = r + s$ Elementen keines dieser Elemente bevorzugt wird, sondern alle gleich häufig zur Verwendung kommen. Dann werden nämlich die verschiedenartigen Systeme von Elementen A' und A'' in solcher Vertheilung auftreten, dass durchschnittlich unter $t(t-1)\ldots(t-m+1)$ Systemen

$$C_{m-i;\,i} = \binom{m}{i} r \ldots (r-m+i+1) \cdot s \ldots (s-i+1)$$

Systeme mit $m - i$ Elementen A' und i Elementen A'' zu erwarten sind[1]). Es ist aber

1) Es werden also durchschnittlich

$C_{m;0} = r(r-1)\cdots(r-m+1)$ Systeme aus m-Elementen A' und

$C_{0;m} = s(s-1)\cdots(s-m+1)$ Systeme aus m-Elementen A''

zu erwarten sein. Solche Systeme nennt Marbe (Naturphilosophische Untersuchungen zur Wahrscheinlichkeitslehre, 1899) reine Gruppen und behauptet (S. 71), »dass in allen Fällen, wo man die Wahrscheinlichkeitsrechnung anzuwenden pflegt, für gewisse Gruppengrößen [d. h. wenn m eine gewisse Anzahl übersteigt] die reinen Gruppen niemals vorkommen«. Er stützt sich auf ähnliche Behauptungen d'Alembert's (Mélanges de littérature, d'histoire et de philosophie, Amsterdam 1767) und sucht durch eigene Versuche (Werfen eines Geldstücks, das Zahl oder Wappen zeigt) und aus den Resultaten des Roulettespiels eine Bestätigung für seine Ansicht zu gewinnen. Dabei begeht jedoch Marbe den Fehler, zu

$$C_{m,0} < C_{m-1;1} < \dots < C_{m-i+1;i-1} < C_{m-i;i} > C_{m-i-1;i+1} > \dots > C_{0;m},$$

falls die Proportion $(m - i) : i = r : s$ erfüllt ist. Denn es verhält sich:

$$C_{m-i-k;i+k} : C_{m-i-k-1;i+k+1} = \frac{r-m+i+k+1}{s-i-k} : \frac{m-i-k}{i+k+1};$$

$$C_{m-i+k;i-k} : C_{m-i+k+1;i-k-1} = \frac{s-i+k+1}{r-m+i-k} : \frac{i-k}{m-i+k+1};$$

und da, wenn $r = \lambda(m - i)$; $s = \lambda i$ gesetzt wird,

$$\frac{(\lambda-1)(m-i)+k+1}{(\lambda-1)i-k} > \frac{m-i}{i} > \frac{m-i-k}{i+k+1}$$

$$\frac{(\lambda-1)i+k+1}{(\lambda-1)(m-i)-k} > \frac{i}{m-i} > \frac{i-k}{m-i+k+1}$$

so ist auch für $k = 0, 1, 2 \dots.$

---

strenge Forderungen hinsichtlich der Uebereinstimmung zwischen theoretisch berechneten und empirisch gefundenen Wahrscheinlichkeitswerthen zu stellen. Ich finde darum, im Gegensatz zu Marbe, in seinen eigenen Versuchen eine hinreichende Bestätigung der üblichen Auffassungsweise, da man doch wohl nicht mehr verlangen kann, als dass die beobachtete Anzahl reiner Gruppen einmal kleiner (in der mitgetheilten Tabelle III) und ein andermal größer ist (in den bloß erwähnten Versuchen S. 14) als die theoretisch bestimmte Anzahl. Eine Uebereinstimmung kann man ja nur im Durchschnitt vieler, streng genommen unendlich vieler Versuche erwarten. Es sind ferner die Differenzen zwischen Theorie und Erfahrung beim Roulettespiel, einzeln betrachtet, gar nicht ungünstig. Nur der für die Tabellen V—X übereinstimmende Gang der Differenzen (die an Stelle der sachlich nicht gerechtfertigten Quotienten Marbe's zu treten haben) kann, meines Erachtens, zu dem Schlusse führen, dass speciell beim Roulettespiel die Tendenz zum Zurückbleiben der empirich beobachteten Anzahlen reiner Gruppen hinter den theoretisch zu fordernden Anzahlen bei wachsendem m vorhanden ist. Dass aber reine Gruppen von einer gewissen Größe des m ab überhaupt nicht mehr vorkommen, folgt hieraus gar nicht. Der angestrebte Beweis wird somit nicht erbracht. Seine Unmöglichkeit erhellt übrigens aus folgendem Beispiel: Es sind zwei Varietäten einer Pflanzenspecies denkbar, deren relative Häufigkeiten oder Wahrscheinlichkeiten für einen gewissen Landstrich bestimmt werden sollen. Die beiden Varietäten können nun gleichmäßig gemischt sein; dann wird man keine reinen Gruppen finden. Es kann aber auch die eine Varietät bloß auf den Bergen und die andere bloß in der Ebene vorkommen; dann wird man beliebig große reine Gruppen finden, falls man ausschließlich auf den Bergen oder in der Ebene botanisirt; man kann ferner eine beliebige Mischung der beiden Varietäten erhalten, wenn man die Pflanzen den verschiedenartigen Standorten entnimmt.

$$C_{m-i-k\,;\,i+k} > C_{m-i-k-1\,;\,i+k+1}$$

$$C_{m-i+k\,;\,i-k} > C_{m-i+k+1\,;\,i-k-1}\,.$$

Somit tritt dasjenige System am häufigsten auf, für welches das Verhältniss $(m - i) : i$ der zugehörigen Elemente A' und A'' dem Verhältnisse $r : s$ gleich ist oder am nächsten kommt, während die von jenem Verhältnisswerth mehr und mehr sich entfernenden Systeme durch ein immer selteneres Auftreten ausgezeichnet sind.

Ist z. B. $t = 15$; $r = 9$; $s = 6$; $m = 5$; sind also 15 Möglichkeiten für A, darunter 9 für A' und 6 für A'' vorhanden, während immer nur je 5 in übrigens unbegrenzter Wiederholung erprobt werden können, so hat man bei gleichmäßiger Inanspruchnahme aller 15 Möglichkeiten unter $15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 = 360\,360$ Bestimmungen folgende Fälle zu erwarten: 15 120 Fälle mit je 5 A'; 90 720 Fälle mit je 4 A' und 1 A''; 151 200 Fälle mit 3 A' und 2 A''; 86 400 Fälle mit 2 A' und 3 A''; 16 200 Fälle mit 1 A' und 4 A''; 720 Fälle mit 5 A''. Demnach findet man, aus dem am häufigsten auftretenden Systeme mit 3 A' und 2 A'' die Werthe

$$p = \frac{3}{5}; \quad q = \frac{2}{5},$$

die mit den wahren Werthen $r : t$ und $s : t$ übereinstimmen; und zwar werden sich diese Werthe mit der relativen Häufigkeit $151\,200 : 360\,360 = 0{,}42$ ergeben. Will man sich aber mit einer bloß angenäherten Bestimmung von p und q begnügen, etwa mit der Erkenntniss, dass

$$\frac{4}{5} \geqq p \geqq \frac{2}{5}; \quad \frac{1}{5} \leqq q \leqq \frac{3}{5},$$

so werden sich die Werthe von p und q innerhalb der angegebenen Grenzen mit der relativen Häufigkeit $(90\,720 + 151\,200 + 86\,400) : 360\,360 = 0{,}91$ finden lassen, da nunmehr zu den Fällen mit je drei A' und zwei A'' noch die Fälle mit je vier A' und einem A'' und je zwei A' und drei A'' hinzukommen.

Zu derselben Gesetzmäßigkeit in etwas einfacherer Form gelangt man, wenn nicht erst das ganze System, sondern schon jedes einzelne Element des zur Abzählung gelangenden Systems der Mannigfaltigkeit wieder einverleibt gedacht wird: wenn also beim Ziehen von m

Kugeln aus der Urne mit $t = r + s$ weißen und schwarzen Kugeln jede einzelne Kugel, nachdem sie gezogen worden, sogleich in die Urne zurückgelegt wird. Dann kann nämlich jedes der t Elemente mit sich selbst ebenso wie mit jedem anderen immer wieder combinirt werden, so dass im Ganzen t^m Systeme möglich sind, die wiederum in $m + 1$ Gruppen zerfallen, jedoch der Art, dass nun

$$\binom{m}{i} r^{m-i} s^i \text{ (für } i = 0, 1, 2 \ldots m)$$

Systeme mit $m - i$ Elementen A' und i Elementen A'' auftreten. Die Mischung der Elemente A' und A'' im Verhältniss $(m - i) : i$ wird somit, wenn keines der t Elemente vor den anderen bevorzugt wird, mit der relativen Häufigkeit

$$\binom{m}{i} \frac{r^{m-i} s^i}{t^m}$$

zu erwarten sein, wo nun bloß noch die Verhältnisse der Anzahlen r, s, t in Betracht kommen und folglich m beliebig gewählt werden kann. Und da

$$r^m < \ldots < \binom{m}{i-1} r^{m-i+1} s^{i-1} < \binom{m}{i} r^{m-i} s^i > \binom{m}{i+1} r^{m-i-1} s^{i+1}$$
$$> \ldots > s^m,$$

wenn $(m - i) : i = r : s$, so können wiederum aus dem am häufigsten auftretenden Mischungsverhältnisse die Werthe p und q bestimmt und die durchschnittlich zu erwartenden relativen Häufigkeiten oder Wahrscheinlichkeiten für das Vorkommen jedes Mischungsverhältnisses angegeben werden.

Wird z. B. $r : s : t = 3 : 2 : 5$ und $m = 5$ vorausgesetzt, so hat man das System von drei A' und zwei A'' am häufigsten und zwar mit der relativen Häufigkeit 0,35 zu erwarten, während die Systeme mit vier A' und einem A'', mit drei A' und zwei A'', mit zwei A' und drei A'' zusammengenommen mit der relativen Häufigkeit 0,84 auftreten werden. — Wird hingegen für die nämlichen Verhältnisswerthe von r, s und t die Anzahl $m = 10$ angenommen, so findet sich das am häufigsten vorkommende System von sechs A' und vier A'' mit der relativen Häufigkeit 0,25, während für die Systeme mit sieben, sechs, fünf A' und drei, vier, fünf A'' die relative Häufigkeit

0,67 und für die Systeme mit acht bis vier A' und zwei bis sechs A'' die relative Häufigkeit 0,89 besteht. — Hiernach hat man die gesuchten Wahrscheinlichkeitswerthe p und q zwischen den Grenzen

$$\frac{7}{10} \geqq p \geqq \frac{5}{10}\,; \quad \frac{3}{10} \leqq q \leqq \frac{5}{10}$$

für $m = 10$ mit der relativen Häufigkeit 0,67; und zwischen den Grenzen

$$\frac{8}{10} \geqq p \geqq \frac{4}{10}\,; \quad \frac{2}{10} \leqq q \leqq \frac{6}{10}$$

für $m = 5$ mit der relativen Häufigkeit 0,84; für $m = 10$ mit der relativen Häufigkeit 0,89 zu erwarten, so dass man die Wahrscheinlichkeitswerthe innerhalb gewisser, die wahren Werthe einschließender Grenzen um so häufiger finden wird, je größer m ist.

Auf diesem Wege hat Jakob Bernoulli im vierten Theile der Ars conjectandi[1]) das (im vierten Capitel) von ihm gestellte und in seinen principiell wichtigen Punkten geklärte Problem der Wahrscheinlichkeitsbestimmung aus einer großen Anzahl von Beobachtungen (im fünften Capitel) gelöst. Er hat gezeigt, dass die Sicherheit in der Bestimmung der Wahrscheinlichkeitswerthe mit der Anzahl der beobachteten Fälle wächst und zwar ohne Grenzen wächst, ohne dass — wie er sagt — das Problem »seine Asymptote« hat und ein bestimmter Grad der Gewissheit, das wahre Verhältniss der Fälle gefunden zu haben, auch bei beliebiger Vermehrung der Beobachtungen niemals überschritten werden kann. Er hat ferner hervorgehoben, dass die Wahrscheinlichkeitswerthe nicht absolut genau — »denn so würde ganz das Gegentheil herauskommen und desto unwahrscheinlicher werden, dass das richtige Verhältniss gefunden sei, je mehr Beobachtungen gemacht wären« — sondern nur angenähert »zwischen zwei Grenzen« bestimmt werden können. Und er gelangt zu dem Satze:

»Es möge sich die Zahl der günstigen Fälle zu der Zahl der ungünstigen Fälle genau oder näherungsweise wie $r : s$, also zu der

---

1) Unter dem Titel »Wahrscheinlichkeitsrechnung (Ars conjectandi) von Jakob Bernoulli (1713)« übersetzt und herausgegeben von Haussner (Ostwald's Klassiker der exacten Wissenschaften, No. 107 und 108).

Zahl aller Fälle wie $r : (r + s) = r : t$ — wenn $r + s = t$ gesetzt wird — verhalten, welches letzteres Verhältniss zwischen den Grenzen

$$\frac{r+1}{t} \quad \text{und} \quad \frac{r-1}{t}$$

enthalten ist. Nun können, wie zu beweisen ist, so viele Beobachtungen gemacht werden, dass es beliebig oft (z. B. c-mal) wahrscheinlicher wird, dass das Verhältniss der günstigen zu allen angestellten Beobachtungen innerhalb dieser Grenzen liegt als außerhalb derselben, also weder größer als $\frac{r+1}{t}$, noch kleiner als $\frac{r-1}{t}$ ist«.

Den Beweis dieses Satzes gründet Bernoulli auf die Eigenschaften der Glieder, die man durch die Entwicklung der Potenz

$$(r+s)^{nt} = r^{nt} + \binom{nt}{1} r^{nt-1} s + \ldots$$

(wo nt an Stelle des oben angenommenen m tritt) erhält, und er findet:

»Es können so viele Beobachtungen angestellt werden, dass die Anzahl der Fälle, in welchen das Verhältniss der günstigen zu allen überhaupt angestellten Beobachtungen die Grenzwerthe

$$\frac{nr+n}{nt} \quad \text{und} \quad \frac{nr-n}{nt} \quad \text{oder} \quad \frac{r+1}{t} \quad \text{und} \quad \frac{r-1}{t}$$

nicht überschreitet, mehr als c-mal größer ist als die Summe der übrigen Fälle, d. h. dass es mehr als c-mal wahrscheinlicher wird, dass das Verhältniss der Anzahl der günstigen zu der Anzahl aller Beobachtungen die Grenzen

$$\frac{r+1}{t} \quad \text{und} \quad \frac{r-1}{t}$$

nicht überschreitet, als dass sie es überschreitet«.

Hierdurch erhält die Methode der Wahrscheinlichkeitsbestimmung aus einer großen Anzahl von Beobachtungen ihre exacte Begründung. Es darf aber nicht außer Acht gelassen werden, dass diese Methode die Erfüllung der oben, an zweiter Stelle genannten Bedingungen voraussetzt, und dass insbesondere das Bernoulli'sche Theorem auf der Annahme beruht, nach welcher bei einer oftmaligen, streng genommen unendlich oftmaligen Bestimmung eines Systems von m Elementen aus der Mannigfaltigkeit des Begriffsumfangs von A kein Element dieser Mannigfaltigkeit vor den anderen bevorzugt wird. Auch ist daran festzuhalten, dass Wahrscheinlichkeit so viel wie relative Häufigkeit bedeutet, wonach die große Wahrscheinlichkeit, bei großem m die gesuchten Werthe von p und q innerhalb gewisser Grenzen zu finden, keineswegs mit empirischer Gewissheit verwechselt werden darf, sondern nur die Thatsache ausdrückt, dass bei einer oftmaligen Bestimmung eines Systems von m Elementen die Werthe von p und q mit überwiegender Häufigkeit zwischen jenen Grenzen sich halten werden.

Wird dies beachtet, so wird der principielle Unterschied zwischen der Wahrscheinlichkeitsbestimmung aus einem dem Gesammtumfange des Begriffs von A ähnlichen Theile und derjenigen aus einem sehr großen Theile dieses Begriffsumfangs nicht verkannt werden[1]. Denn man wird alsdann nicht etwa auf Grund des Bernoulli'schen Theorems einen sehr großen Theil ohne weiteres als einen dem Ganzen ähnlichen Theil voraussetzen. Lässt doch dieses Theorem nur erkennen, dass bei Erfüllung der nothwendigen Bedingungen sehr große Theile relativ häufiger als kleinere Theile dem Ganzen ähnlich sind und bei gleich großen Theilen die Aehnlichkeit innerhalb gewisser Grenzen häufiger ist als die Unähnlichkeit.

Es gibt sonach im ganzen drei Methoden für das induc-

1) Auf eine Verkennung dieses Unterschiedes weist bei Bernoulli die Bemerkung im vierten Capitel des vierten Theils der Ars conjectandi (Ostwald's Klassiker No. 108, S. 89): »Dabei [nämlich bei der Ermittelung der Wahrscheinlichkeiten a posteriori aus einer großen Anzahl von Beobachtungen, die durch das von ihm entwickelte Theorem ihre Begründung finden soll] muss angenommen werden, dass jedes einzelne Ereigniss in ebenso vielen Fällen eintreten oder nicht eintreten kann, als vorher bei einem gleichen Stande der Dinge beobachtet wurde, dass er eingetreten oder nicht eingetreten ist. Denn z. B. wenn man beobachtet hat, dass von 300 Menschen von dem Alter und der Constitution des Titius 200 vor Ablauf von 10 Jahren gestorben sind, die übrigen aber länger gelebt haben, so kann man mit hinreichender Sicherheit folgern, dass es doppelt so viele Fälle gibt, in welchen auch Titius innerhalb des nächsten Decenniums der Natur den schuldigen Tribut leisten muss, als Fälle, in welchen er diesen Zeitpunkt überleben kann«. Hier setzt demnach Bernoulli eine Aehnlichkeit der beiden Decennien voraus, während sein Theorem bei einer Vielheit solcher Zeiträume vielmehr eine bestimmte gesetzmäßige Gruppirung der möglichen Wahrscheinlichkeitswerthe um einen am häufigsten vorkommenden Werth verlangt.

tive Erkennen der Wahrscheinlichkeit, dass ein Denkgegenstand A als ein mit B verbundenes A' auftritt:

1. die unmittelbare Bestimmung der Begriffsumfänge von A' und A, deren Verhältniss den gesuchten Wahrscheinlichkeitswerth darstellt;

2. die Bestimmung eines dem Begriffsumfange von A ähnlichen Theils;

3. die Bestimmung von Elementen des Begriffsumfanges von A in hinreichend großer Zahl, wenn die Voraussetzungen für die Anwendung des Bernoulli'schen Theorems erfüllt sind, so dass aus n Systemen von je m Elementen der am häufigsten vorkommende Wahrscheinlichkeitswerth mit hinreichender Sicherheit ermittelt werden kann.

§ 7. Die Definition des Collectivgegenstandes.

Bisher war nur von den beiden Möglichkeiten die Rede, dass der gegebene Denkgegenstand A sowohl das mit B verbundene A' als auch das ohne B vorkommende A'' sein könne. Von besonderen Fällen abgesehen, wird man jedoch eine größere Anzahl verschiedener, einander ausschließender Möglichkeiten ins Auge zu fassen haben. Dann wird A entweder als das mit B_1 verbundene A_1, oder das mit B_2 verbundene A_2 u. s. w., oder schließlich das mit B_n verbundene A_n auftreten, so dass im ganzen n Modificationen von A möglich sind und der Umfang des Begriffs von A aus den n Begriffsumfängen von $A_1, A_2 \ldots A_n$ sich zusammensetzt. Die Anzahl n ist keiner Einschränkung unterworfen: sie kann jeden endlichen Werth 2, 3, 4 ... oder auch einen unbegrenzt großen Werth annehmen.

Einen solchen mehrerer oder beliebig vieler Modificationen fähigen Denkgegenstand nenne ich, indem ich eine von Fechner[1])

1) Unter einem Collectivgegenstand versteht Fechner (Collectivmaßlehre, 1897, S. 3) »einen Gegenstand, der aus unbestimmt vielen, nach Zufall variirenden Exemplaren besteht, die durch einen Art- oder Gattungsbegriff zusammengehalten werden«. Das Variiren der Exemplare bedeutet den Uebergang von einer Modification zu einer anderen. Dass dies zufällig sein soll, wurde in die obige Bestimmung nicht aufgenommen. Denn offenbar bedeutet das »zufällige« Variiren ein »regelloses« Schwanken zwischen den verschiedenen Modificationen. Die Anordnung der Exemplare innerhalb des Begriffsumfanges von A und die Art ihrer

eingeführte Bezeichnungsweise in erweitertem Sinne anwende, einen Collectivgegenstand (C.G.). Als Exemplare des C.G. sind dann alle denkbaren Unterlagen von A, die in ihrer Gesammtheit den Begriffsumfang von A bilden, aufzufassen. Die verschiedenen Modificationen $A_1, A_2 \ldots A_n$, in denen A auftreten kann, mögen die Varianten[1]) von A genannt werden. Jeder Variante kommt ein Wahrscheinlichkeitswerth zu, der das Verhältniss des Begriffsumfanges der Variante zum Gesammtumfange des Begriffs von A angibt; und da die Begriffsumfänge der Varianten zusammengenommen den Begriffsumfang von A darstellen, so ist die Summe der Wahrscheinlichkeitswerthe aller Varianten stets gleich 1. Es ist sonach, wenn die Wahrscheinlichkeiten für das Auftreten von $A_1, A_2 \ldots A_n$ der Reihe nach durch $p_1, p_2 \ldots p_n$ bezeichnet werden,

$$p_1 + p_2 + \cdots + p_n = 1 .$$

Die Kenntniss der Varianten und ihrer Wahrscheinlichkeiten kann, wie die bisherigen Erörterungen zeigen, auf deductivem und auf inductivem Wege gewonnen werden. Sie ist durch Deduction erreichbar, wenn zu dem vorhandenen Begriff von A die Erkenntniss der Bedingungen gehört, unter denen A auftritt, so dass ein Ueberblick über den Begriffsumfang von A und dessen Zusammensetzung aus den Begriffsumfängen der Varianten gewonnen werden kann. Ist dies nicht der Fall, so muss die Kenntniss auf inductivem Wege, durch empirisches Erforschen des Begriffsumfanges von A angestrebt werden. Wird zu diesem Zwecke der ganze Begriffsumfang, oder, da dies im allgemeinen nicht ausführbar ist, ein dem Gesammtumfange ähnlicher Theil (mag er direct oder auf Grund des Bernoulli'schen Theorems bestimmbar sein) abgezählt und ergeben

---

Vertheilung auf die Begriffsumfänge von $A_1, A_2 \ldots A_n$ kommt aber bei der Wahrscheinlichkeitsbestimmung an und für sich nicht in Betracht. Dass die Beschränkung auf zufällige Variation unwesentlich sei, bemerkt auch H. Bruns (zur Collectivmaßlehre, Philosophische Studien XIV. S. 346). Es ist indessen die Berücksichtigung des Zufalls für die Wahrscheinlichkeitsrechnung überhaupt entbehrlich. Die verschiedenen Arten, den Zufall zu begreifen, erörtert Windelband (Die Lehren vom Zufall, 1870).

1) Von G. Duncker (Die Methode der Variationsstatistik, Archiv für Entwicklungsmechanik der Organismen, VIII. S. 115) werden speciell die Einzelformen, die innerhalb einer Thier- oder Pflanzenspecies zur Beobachtung gelangen, Varianten genannt.

sich auf diese Weise unter m Exemplaren z_1 Exemplare A_1, z_2 Exemplare A_2, ... z_n Exemplare A_n (wo $m = z_1 + z_2 + \cdots + z_n$), so findet man zugleich mit der Kenntniss der Varianten $A_1, A_2 \ldots A_n$ die Wahrscheinlichkeiten derselben, nämlich:

$$p_1 = \frac{z_1}{m}, \quad p_2 = \frac{z_2}{m}, \quad \ldots p_n = \frac{z_n}{m},$$

die der Bedingung

$$p_1 + p_2 + \ldots + p_n = 1$$

genügen. Alsdann bietet sich die Aufgabe dar, die inductiv gewonnene Kenntniss der Varianten und ihrer Wahrscheinlichkeiten für den Begriff von A selbst fruchtbar zu machen.

Kann diese Aufgabe durch besondere, den C.G. eigenthümliche Methoden gelöst werden, so veranlasst sie eine »Theorie der Collectivgegenstände«, die einen Bestandtheil der empirischen Wahrscheinlichkeitslehre bildet. Denn das Untersuchungsmaterial besteht aus empirisch bestimmten Wahrscheinlichkeiten.

Die auf inductive Wahrscheinlichkeitserkenntniss gegründete Theorie der C.G. hat alsdann die Aufgabe, aus der von der Erfahrung dargebotenen Kenntniss der Varianten und ihrer Wahrscheinlichkeiten eine für den C.G. selbst gültige Erkenntniss zu gewinnen.

Bei der Entwicklung der Methoden, die zur Lösung dieser Aufgabe dienlich sind, ist im Auge zu behalten, dass im allgemeinen (wenn nicht der ganze Begriffsumfang zur Verfügung steht) nur eine angenäherte Bestimmung der Wahrscheinlichkeitswerthe möglich ist. Demzufolge ist einestheils jeder von den n gefundenen Werthen:

$$p_1 = \frac{z_1}{m}, \quad p_2 = \frac{z_2}{m}, \quad \ldots p_n = \frac{z_n}{m}$$

nur innerhalb gewisser Grenzen als zuverlässig anzusehen, und es ist anderntheils in Rechnung zu ziehen, dass außer den thatsächlich beobachteten Varianten $A_1, A_2 \ldots A_n$ noch andere vorhanden sein können, die wegen ihrer geringen relativen Häufigkeit oder Wahrscheinlichkeit in dem beobachteten Theil des Begriffsumfangs von A nicht vertreten sind.

Die Theorie der C.G. muss daher einestheils die im allgemeinen unvermeidlichen Ungenauigkeiten in der Bestimmung der Wahrscheinlichkeitswerthe und anderntheils die Möglichkeit des Vorkommens von Varianten mit geringen, der Beobachtung sich entziehenden Wahrscheinlichkeiten berücksichtigen.

Durch die Aufstellung dieses Grundsatzes erhalten die zur Bestimmung der C.G. tauglichen Methoden von vorn herein ein besonderes Gepräge. Denn es ist zwar möglich, dass die Varianten aus dem Begriff von A deducirt oder durch eine willkürliche Eintheilung des Begriffsumfangs unzweideutig festgestellt werden können und nur die Wahrscheinlichkeitswerthe empirisch zu bestimmen sind; es ist ferner denkbar, dass auch die Wahrscheinlichkeiten völlig genau gefunden werden können. Eine allgemein gültige Methode lässt sich aber auf die Annahme solcher besonderer Fälle nicht gründen.

II. Die Bestimmung der Collectivgegenstände durch Mittelwerthe.

§ 1. Die Collectivgegenstände als Objecte mathematischer Untersuchung.

Die Aufgabe, aus den inductiv erkannten Wahrscheinlichkeiten der Varianten eines durch A bezeichneten C.G. eine für den C.G. gültige Erkenntniss zu gewinnen, bedarf einer genaueren Bestimmung, da es sich ebensowohl um eine den Inhalt wie um eine die Form von A betreffende Erkenntniss handeln kann.

Der Inhalt von A beruht wesentlich auf den Beziehungen, welche den Begriff von A bilden und die Zugehörigkeit des A zu einem bestimmten Wissensgebiete bedingen. Jede den Inhalt von A betreffende neue Erkenntniss wird daher gleichfalls jenem Wissensgebiete angehören. Ist z. B. A eine Pflanzen- oder Thierspecies, und werden die Wahrscheinlichkeiten ihrer Varianten unter verschiedenen klimatischen oder sonstigen Verhältnissen beobachtet, so kann man hierdurch einen etwaigen Einfluss des Klimas oder der sonstigen Verhältnisse auf A feststellen und eine Mehrung botanischer oder zoologischer Erkenntniss erzielen. Um solche Erkenntnisse handelt es sich in der allgemeinen Lehre von den C.G. nicht.

Die Form von A ist hingegen durch das bloße Vorhandensein von Varianten und ihre Wahrscheinlichkeitswerthe bedingt. Auf Grund derselben zerfällt der Begriffsumfang von A in eine bestimmte oder unbestimmt große Anzahl von Theilumfängen, die den einzelnen Varianten zugehören. Und man kann nun, ohne auf den Inhalt von A Rücksicht zu nehmen, fragen, welche Merkmale für A aus der Art und Weise der Zusammensetzung seines Begriffsumfangs oder (mit anderen Worten) aus den Wahrscheinlichkeitswerthen seiner Varianten sich ergeben. Solche auf der Form von A beruhenden Erkenntnisse zu gewinnen, ist die zunächst liegende Aufgabe der Lehre von den C.G. Sie ist mit mathematischen Hülfsmitteln zu lösen, da zahlenmäßig bestimmte Wahrscheinlichkeiten das Untersuchungsmaterial bilden.

Hierbei sind von vorn herein zwei Fälle zu unterscheiden.

Die Varianten des C.G. A können nämlich einerseits beziehungslos neben einander stehen, so dass sie keiner objectiv begründeten Ordnung fähig sind und nur die Möglichkeit besteht, nach subjectiver Willkür die eine als A_1, eine andere als A_2, eine dritte als A_3 u. s. w. zu bezeichnen und so aufeinander folgen zu lassen, weil eben das Erfassen einer Vielheit nicht anders als in der Form einer Reihe oder eines Systemes von Reihen möglich ist. Dann gehören, wenn unter m Exemplaren von A z_1 Exemplare A_1, z_2 Exemplare A_2 u. s. f. gefunden wurden, die Häufigkeiten $z_1, z_2 \ldots z_n$ oder die Wahrscheinlichkeiten

$$p_1 = \frac{z_1}{m}, \quad p_2 = \frac{z_2}{m}, \quad \ldots \quad p_n = \frac{z_n}{m}$$

zwar zusammen, weil sie den Varianten $A_1, A_2 \ldots A_n$ zugehören und insgesammt auf A sich beziehen; es kann jedoch nicht von einem objectiv begründeten Verlauf dieser Werthe und von einer Gesetzmäßigkeit des Verlaufs die Rede sein. — Werden z. B. unter den Säugethieren Raubthiere auf Grund der Beschaffenheit ihres Gebisses, Wiederkäuer mit Rücksicht auf die Einrichtung des Magens, Hufthiere wegen des Vorkommens von Zehen, die mit Hufen bekleidet sind u. s. w., unterschieden, und denkt man sich die Wahrscheinlichkeiten oder relativen Häufigkeiten der verschiedenen Ordnungen, in welche demgemäß die Classe der Säugethiere eingetheilt wird, für einen Erdtheil etwa und für einen gegebenen Zeitraum bestimmt, so ist ein Gesetz, nach welchem sich die Wahrscheinlichkeiten auf die Ordnungen vertheilen möchten, nicht angebbar, so lange keine objectiv begründete Reihenfolge der Ordnungen (beispielsweise nach einem entwicklungsgeschichtlich motivirten Stammbaum) hergestellt werden kann.

Trifft dieser Fall zu, so ist das System der gefundenen Wahrscheinlichkeitswerthe unmittelbar der Charakterisirung von A zu Grunde zu legen. Es besteht aus einer Mannigfaltigkeit reeller, positiver Zahlenwerthe, deren Summe stets gleich 1 ist.

Die Varianten von A können anderseits in solcher Beziehung zu einander stehen, dass sie als Elemente eines wohlgeordneten, aus einer Reihe oder aus einem Gewebe zusammenhängender Reihen bestehenden Systems aufzufassen sind. Dieses System wird von der Gesammtheit aller denkbaren Varianten gebildet, sodass jede thatsächlich beobachtete Variante einer bestimmten Stelle des Systemes angehört und durch diese Stelle bezeichnet werden kann. Als Stellenzeichen dienen die Ordnungszahlen. Denn jede Reihe des Systems findet in der Reihe der Ordnungszahlen ihr zutreffendes Bild, sofern es nicht auf die Beschaffenheit der Varianten, sondern nur auf die Möglichkeit, sie zu ordnen, ankommt. Die Ordnung einer Reihe besteht aber darin, dass auf jedes Glied ein anderes folgt, wenn es nicht das letzte ist, und jedem ein anderes vorangeht, wenn es nicht das erste ist. Dabei hängt es ganz von der Beschaffenheit der ordnenden Beziehungen ab, ob für eine bestimmte Variantenreihe ein Anfangsglied und Endglied vorhanden ist oder nicht und ob ein solches zwar denkbar, aber nicht angebbar oder nicht nur denkbar, sondern auch angebbar ist. Es kann hingegen nicht durch die Beobachtung der Varianten festgestellt werden, ob eine Reihe begrenzt oder unbegrenzt ist; denn die Beobachtung führt stets zu einer bestimmten endlichen Anzahl von Varianten, ohne dass auf diese Weise das Nichtvorhandensein der Varianten, die der Beobachtung aus dem einen oder dem anderen Grunde sich vielleicht entziehen, erwiesen werden könnte.

Man kann demnach, wenn die Gesammtheit aller denkbaren

Varianten in eine einzige Reihe sich ordnen lässt, ein beliebiges Glied dieser Reihe durch irgend eine positive oder negative Ordnungszahl (oder auch durch die Null) markiren und den auf der einen Seite folgenden Varianten die aufsteigenden, den auf der anderen Seite folgenden Varianten die absteigenden Glieder der Zahlenreihe zuweisen. Es wird alsdann, wenn die als Ausgangsglied gewählte Variante durch die Ordnungszahl α bezeichnet wird, die Reihe aller denkbaren Varianten durch

$$\ldots \alpha - 2, \quad \alpha - 1, \quad \alpha, \quad \alpha + 2, \quad \alpha + 4, \ldots$$

dargestellt.

Besteht jedoch das System der denkbaren Varianten nicht aus einer einzigen Reihe, sondern aus einem Gewebe von Reihen, so ist jede Variante, die mehreren Reihen angehört, durch ebenso viele Ordnungszahlen, als die Anzahl der Reihen beträgt, zu markiren. Eine Doppelreihe von Varianten wird z. B., wenn eine beliebige Variante durch das Paar von Ordnungszahlen (α, β) angegeben wird, durch

$$\begin{array}{lccc}
 & \cdot & \cdot & \cdot \\
 & \cdot & \cdot & \cdot \\
.. & (\alpha - 1, \beta - 1) & (\alpha, \beta - 1) & (\alpha + 1, \beta - 1) \;.. \\
.. & (\alpha - 1, \beta) & (\alpha, \beta) & (\alpha + 1, \beta) \;.. \\
.. & (\alpha - 1, \beta + 1) & (\alpha, \beta + 1) & (\alpha + 1, \beta + 1) \;.. \\
 & \cdot & \cdot & \cdot \\
 & \cdot & \cdot & \cdot
\end{array}$$

dargestellt.

Eine solche Abbildung durch Reihen von Ordnungszahlen ist immer dann ausführbar, wenn die Varianten Abstufungen eines und desselben Merkmals, oder mehrerer zusammengehöriger Merkmale des C.G. sind. Beispielsweise, wenn es sich um die in Millimetern etwa angegebenen Längen der Früchte einer Pflanze, oder um die Anzahlen der Samenkörner in den Früchten handelt. Als abstufbare Merkmale können aber nicht nur stetig veränderliche Größen oder zählbare Mengen, sondern auch Qualitäten, für die es eine Scala der möglichen Unterschiede gibt, auftreten. Es können also z. B. Farben- und Tonempfindungen, nicht aber Geruchsempfindungen in Betracht kommen, weil eine Farbenscala (für die unterscheidbaren Farbentöne, Sättigungsgrade und Helligkeitsstufen) ebenso wie eine Tonscala (für die unterscheidbaren Tonhöhen) vorhanden ist, während eine vollständige und unanfechtbare Scala der Gerüche fehlt. Dabei ist es gleichgültig, ob die aufeinanderfolgenden Stufen letzte, unüberbrückbare Unterschiede (ähnlich den Anzahlen völlig gleichartiger Dinge) oder nur einen gewissen Grad der Unterscheidbarkeit (ähnlich den auf Millimeter abgerundeten Längenangaben) zum Ausdruck bringen, so dass bei schärferer Unterscheidung jede Stufe selbst wieder Abstufungen enthält und schließlich bei unbegrenzt feiner Abtönung der Unterschiede in unendlich viele unendlich wenig verschiedene Stufen zerfallend zu denken ist.

Die einer bestimmten, objectiv begründeten Ordnung fähigen Varianten der C.G. können daher ebensowohl Vielfache einer Maßeinheit oder Anzahlen gleichartiger Dinge, wie auch Stufen einer stetig oder gradweise veränderlichen Qualität sein.

Eben derselben Ordnung werden nun auch die beobachteten Häufigkeiten oder Wahrscheinlichkeiten durch ihre Zugehörigkeit zu den Varianten theilhaftig. Es ist daher ein objectiv, durch die Bezugnahme auf die Varianten begründeter Verlauf vorhanden, dessen Besonderheiten festgestellt und zur Charakterisirung des C.G. benutzt werden können. Man wird darum nicht, wie bei den C.G., deren Varianten keiner objectiv begründeten Ordnung fähig sind, das System der Wahrscheinlichkeitswerthe für sich allein (was allerdings auch jetzt wieder — eben unter Absehen von der für die Varianten bestehenden Ordnung — möglich wäre), sondern vielmehr die absoluten oder relativen Häufigkeiten in ihrer Zugehörigkeit zu den Varianten ins Auge fassen müssen. Demgemäß sind die Häufigkeiten der Ordnungszahlen, welche die zur Beobachtung gelangten Varianten repräsentiren, der Charakterisirung des C.G. zu Grunde zu legen.

Lassen sich also die denkbaren Varianten des gegebenen C.G. in eine einzige Reihe ordnen und werden unter m Exemplaren z_α Exemplare der durch α markirten Variante, $z_{\alpha-1}$ Exemplare der Variante $\alpha - 1$, $z_{\alpha+1}$ Exemplare der Variante $\alpha + 1$ u. s. w. gefunden, wo jede zwar denkbare, aber thatsächlich nicht beobachtete Variante mit einem $z = 0$ behaftet aufzufassen ist, so wird das von z_α Werthen α, $z_{\alpha-1}$ Werthen $\alpha - 1$, $z_{\alpha+1}$ Werthen $\alpha + 1$ u. s. w. gebildete

System den Ausgangspunkt der Untersuchung bilden. Es kann in der Form

$$\frac{\ldots\ \alpha-1,\quad \alpha,\quad \alpha+1,\ \ldots}{\ldots\ z_{\alpha-1},\quad z_{\alpha},\quad z_{\alpha+1},\ \ldots}$$

oder hiermit gleichbedeutend, wenn $p_\alpha = \frac{z_\alpha}{m}$, $p_{\alpha-1} = \frac{z_{\alpha-1}}{m}$, $p_{\alpha+1} = \frac{z_{\alpha+1}}{m}$ u. s. w. gesetzt wird, in der Form

$$\frac{\ldots\ \alpha-1,\quad \alpha,\quad \alpha+1,\ \ldots}{\ldots\ p_{\alpha-1},\quad p_{\alpha},\quad p_{\alpha+1},\ \ldots}$$

vorausgesetzt werden. Diese tabellarische Zuordnung der absoluten oder relativen Häufigkeitswerthe zu den durch Ordnungszahlen repräsentirten Varianten soll nach **Fechner** die **Vertheilungstafel** des C.G. genannt werden.

Bilden hingegen die denkbaren Varianten des C.G. eine Doppelreihe und befinden sich unter m Exemplaren $z_{\alpha,\beta}$ Exemplare der Variante (α, β), $z_{\alpha+1,\beta}$ Exemplare der Variante $(\alpha+1, \beta)$, $z_{\alpha,\beta+1}$ Exemplare der Variante $(\alpha, \beta+1)$ u. s. w., so wird das aus $z_{\alpha,\beta}$ Werthenpaaren (α, β), $z_{\alpha+1,\beta}$ Werthenpaaren $(\alpha+1, \beta)$, $z_{\alpha,\beta+1}$ Werthenpaaren $(\alpha, \beta+1)$ u. s. w. bestehende System die Grundlage der Untersuchung bilden. Die Vertheilungstafel kann in der Form

	. .	$\alpha-1$	α	$\alpha+1$	. .
.	. .	.	.	.	. .
.	. .	.	.	.	. .
$\beta-1$	. .	$z_{\alpha-1,\beta-1}$	$z_{\alpha,\beta-1}$	$z_{\alpha+1,\beta-1}$	. .
β	. .	$z_{\alpha-1,\beta}$	$z_{\alpha,\beta}$	$z_{\alpha+1,\beta}$	. .
$\beta+1$	. .	$z_{\alpha-1,\beta+1}$	$z_{\alpha,\beta+1}$	$z_{\alpha+1,\beta+1}$	. .
.	. .	.	.	.	. .
.	. .	.	.	.	. .

vor Augen gestellt werden, wo jedes z die Häufigkeit desjenigen Werthenpaares angibt, dessen Glieder in der zu z gehörenden Horizontal- und Verticalreihe stehen, und wo an Stelle der absoluten Häufigkeiten $z_{\alpha,\beta}$ wiederum die relativen Häufigkeiten

$$p_{\alpha,\beta} = \frac{z_{\alpha,\beta}}{m}$$

treten können.

Hieraus erhellt, dass in jedem Falle, mögen die Varianten des C.G. einer objectiv begründeten Ordnung fähig sein oder nicht, eine Mannigfaltigkeit reeller Zahlenwerthe die Grundlage für die Charakterisirung des C.G. bildet. Diese Mannigfaltigkeit besteht aus Einzelwerthen, wenn es sich um das System der Wahrscheinlichkeitswerthe selbst oder um eine einzige Variantenreihe handelt; sie besteht aus Werthenpaaren bei einer Doppelreihe von Varianten und allgemein aus Combinationen von je n Werthen bei n-fach zusammenhängenden Variantenreihen.

Dieser Gemeinsamkeit des mathematischen Ausgangspunktes der Untersuchung entsprechend könnte auch die weitere Behandlung für alle Fälle gleichartig gestaltet werden: denn man kann stets entweder von den gefundenen Werthen selbst ausgehen und die Bestimmung dieses Systems sich zur Aufgabe machen, oder aber ein Schema möglicher Werthe voraussetzen, die Fächer desselben mit den gefundenen Werthen besetzen (wo dann im allgemeinen neben einfach und mehrfach besetzten Fächern leer bleibende Fächer auftreten werden) und nun das Gesetz, nach dem sich die Werthe auf das vorausgesetzte Schema vertheilen, zu bestimmen suchen.

In der That ist es, wenn die n Wahrscheinlichkeitswerthe p_1, p_2 $\ldots$ p_n ohne Bezugnahme auf die zugehörigen Varianten zur Charakterisirung eines C.G. dienen sollen, zulässig, statt zur Bestimmung des Systems dieser Werthe zu schreiten, eine Scala von Wahrscheinlichkeitswerthen, die sich gleichmäßig über das Intervall der Zahlenwerthe von 0—1 erstreckt, vorauszusetzen und die Häufigkeit des Vorkommens jeder Stufe (ob sie 0mal, 1mal, 2mal u. s. w. auftritt) anzugeben, um sodann die so construirte Vertheilungstafel der Wahrscheinlichkeitswerthe weiter zu behandeln. Man erhält so, wenn die Stufen der Scala allgemein durch

$$p^{(o)} = \frac{o}{l}, \quad p' = \frac{1}{l}, \quad p'' = \frac{2}{l}, \ldots$$

(wo l eine ganze Zahl vorstellt) bezeichnet werden und die Anzahlen n_0, n_1, n_2 $\ldots$ der Reihe nach für jede einzelne Stufe angeben, ob dieselbe und wie oft sie unter den gefundenen Wahrscheinlichkeitswerthen p_1, p_2 $\ldots$ p_n vorkommt, als Darstellung des ursprünglich vorhandenen Systems die tabellarische Zuordnung

$$\frac{p^{(o)},\quad p',\quad p'',\ldots}{n_0,\quad n_1,\quad n_2,\ldots},$$

wo $n_0 p^{(o)} + n_1 p' + n_2 p'' + \cdots = 1$. Und diese Vertheilungstafel ist im Grunde nichts anderes als eine Anordnung der gefundenen Werthe $p_1, p_2 \ldots p_n$ nach der Größe unter gleichzeitiger Beschränkung der Werthangabe auf die durch $1:l$ bestimmten Bruchtheile. — Sollen anderseits die n absoluten oder relativen Häufigkeiten $z_1, z_2 \ldots z_n$ oder $p_1, p_2 \ldots p_n$ in ihrer Zugehörigkeit zu den einer Reihe angehörigen und durch die Ordnungszahlen $\alpha_1, \alpha_2 \ldots \alpha_n$ bezeichneten Varianten zur Charakterisirung des C.G. dienen, so kann an Stelle der Vertheilungstafel auch das aus z_1 Werthen α_1, z_2 Werthen $\alpha_2, \ldots z_n$ Werthen α_n bestehende System unmittelbar zum Gegenstande der Untersuchung gemacht werden. Denn die Zugehörigkeit der Häufigkeitswerthe zu den Ordnungszahlen der Varianten ergibt sich zugleich mit der Bestimmung jenes Systems, dessen Besonderheit gerade darin liegt, dass n Gruppen von Werthen, nämlich z_1 Werthe α_1, z_2 Werthe α_2 u. s. w. vorhanden sind.

Man wird indessen doch im ersteren Falle bei dem System der Wahrscheinlichkeitswerthe stehen bleiben und im letzteren Falle auf die directe Ermittelung des Gesetzes, nach dem die beobachteten Häufigkeiten auf die Reihe der denkbaren Varianten sich vertheilen, nicht verzichten. Denn die leer bleibenden Stufen der Scala von Wahrscheinlichkeiten $p^{(o)}, p', p'' \ldots$ sind mit Ausnahme der Stufe $p^{(o)} = 0$ (die im allgemeinen mit einer gewissen, aber nicht angebbaren Anzahl n_0 verschwindender, der Beobachtung sich entziehender und darum gleich Null zu setzender Wahrscheinlichkeiten belegt zu denken ist) in Wirklichkeit als nicht vorhanden anzusehen, so dass kein Bedürfniss besteht, dieselben neben den thatsächlich vertretenen Stufen im Auge zu behalten und ein Vertheilungsgesetz der Anzahlen $n_0, n_1 \ldots n_l$ unter Berücksichtigung all dieser Stufen zu entwickeln. Hingegen sind alle mit der empirischen Häufigkeit Null auftretenden Stufen der Variantenreihe stets als mögliche Stufen festzuhalten und mit geringen, der Beobachtung sich entziehenden Häufigkeitswerthen behaftet zu denken, so dass sie neben den thatsächlich beobachteten Stufen beachtet werden müssen. Darum beansprucht hier das Vertheilungsgesetz unter Zugrundelegen aller denkbaren Stufen der Variantenreihe das hauptsächliche Interesse, das nicht bloß auf dem Umwege durch Bestimmung des unmittelbar vorliegenden Systems von z_1 Werthen $\alpha_1 \ldots z_n$ Werthen α_n befriedigt werden kann.

Es ist darum trotz der gemeinsamen mathematischen Grundlage die Verschiedenheit zwischen den C.G., deren Varianten keiner objectiv begründeten Ordnung fähig sind, und den C.G., deren Varianten in Reihen geordnet werden können, anzuerkennen und sowohl die Methode für die Bestimmung eines unmittelbar vorliegenden Systems von Wahrscheinlichkeitswerthen als auch die Methode für die Feststellung von Gesetzmäßigkeiten gegebener Vertheilungstafeln oder von Vertheilungsgesetzen zu entwickeln. Die erstere Methode soll jedoch nur wegen des theoretischen Interesses, das sie an sich und durch den Vergleich mit der letzteren Methode besitzt, hier in Kürze behandelt werden. Den Hauptgegenstand der Untersuchung werden hingegen die Vertheilungstafeln der C.G. bilden.

§ 2. Die Charakterisirung eines Collectivgegenstandes auf Grund des Systems der Wahrscheinlichkeitswerthe.

Ein C.G. mit den empirisch ermittelten Varianten $A_1, A_2, \ldots A_n$ und den zugehörigen Wahrscheinlichkeitswerthen $p_1, p_2, \ldots p_n$ sei gegeben. Sind nun in Folge des Mangels einer objectiv begründeten Ordnung alle subjectiv möglichen Anordnungen der Wahrscheinlichkeitswerthe gleichberechtigt, so sind nur solche Combinationen von $p_1, p_2, \ldots p_n$ zur Charakterisirung des C.G. tauglich, die bei einer beliebigen Vertauschung ihrer Elemente unverändert bleiben. Es können somit nur die sogenannten symmetrischen Functionen von $p_1, p_2, \ldots p_n$ in Betracht kommen, und man wird der Einfacheit wegen algebraische Functionen wählen.

Solcher Functionen gibt es unbegrenzt viele. Indessen lässt sich bekanntlich[1]) jede derselben durch die Coefficienten der Gleichung n-ten Grades, welche die n Werthe $p_1, p_2 \ldots p_n$ zu Wurzeln hat, ausdrücken. Schreibt man diese Gleichung in der Form

1) In den Lehrbüchern der Algebra wird bewiesen, dass eine beliebige rationale, symmetrische Function der Wurzeln einer Gleichung rational durch die Coefficienten der Gleichung ausgedrückt werden kann.

$$p^n - b_1 p^{n-1} + b_2 p^{n-2} - \cdots \pm b_n = 0\,, \tag{1}$$

so ist

$$\left.\begin{array}{l} b_1 = p_1 + p_2 + \cdots + p_n \\ b_2 = p_1 \cdot p_2 + p_1 \cdot p_3 + \cdots + p_{n-1} \cdot p_n \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ b_n = p_1 \cdot p_2 \ldots p_n\,. \end{array}\right\} \tag{2}$$

Man kann sich daher auf die n durch b_1, $b_2 \ldots b_n$ bezeichneten symmetrischen Grundfunctionen von p_1, $p_2 \ldots p_n$ beschränken, um eine vollständige Bestimmung des Systems dieser Werthe und eine möglichst weitgehende Charakterisirung des C.G. zu erhalten. Dabei ist stets

$$b_1 = p_1 + p_2 + \cdots + p_n = 1\,. \tag{3}$$

An Stelle von b_1, $b_2 \ldots b_n$ kann man auch die Potenzsummen

$$\left.\begin{array}{l} s_1 = p_1 + p_2 + \cdots + p_n \\ s_2 = p_1^2 + p_2^2 + \cdots + p_n^2 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ s_n = p_1^n + p_2^n + \cdots + p_n^n \end{array}\right\} \tag{4}$$

wählen, wo $s_1 = b_1 = 1$. Denn die Werthe s_1, $s_2 \ldots s_n$ hängen mit den Werthen b_1, $b_2 \ldots b_n$ in einfacher Weise zusammen, da nach den Newton'schen Formeln:

$$\begin{array}{l} s_1 - b_1 = 0 \\ s_2 - b_1 s_1 + 2b_2 = 0 \\ s_3 - b_1 s_2 + b_2 s_1 - 3b_3 = 0 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{array}$$

und somit das eine Werthensystem aus dem andern gefunden werden kann.

Die Werthe b_1, $b_2 \ldots b_n$ und s_1, $s_2 \ldots s_n$ sind aber untereinander und mit den p-Werthen nicht unmittelbar vergleichbar, da die Summen, durch welche jene Werthe dargestellt werden, verschiedene Anzahlen von Gliedern und Glieder von verschiedenen Dimensionen besitzen. Es besteht nämlich b_i aus

$$\binom{n}{i} = \frac{n(n-1)\ \ldots\ (n-i+1)}{1 \cdot 2\ \ \ \ldots\ \ \ i}$$

Summanden, von welchen jeder ein aus i Factoren gebildetes Product ist, und s_i ist aus n Potenzen vom i-ten Grade zusammengesetzt (für $i = 1, 2 \ldots n$).

Es empfiehlt sich darum zu Mittelwerthen überzugehen, die ich in folgender Weise definire.

Es sei p_0 ein beliebiger reeller Zahlenwerth. Man bilde nun die Differenzen $p_1 - p_0$, $p_2 - p_0$, $\ldots$ $p_n - p_0$ und erhebe jede in die ν-te Potenz. Dann erhält man aus der Summe

$$n \cdot \pi_\nu^\nu = (p_1 - p_0)^\nu + (p_2 - p_0)^\nu + \cdots + (p_n - p_0)^\nu \tag{5}$$

nach Division mit n und Ausziehen der ν-ten Wurzel den reellen Werth π_ν, der für ein geradzahliges ν dem absoluten Betrage nach zu nehmen ist. Den Werth π_ν bezeichne ich als den auf den Ausgangswerth p_0 bezogenen Mittelwerth ν-ter Ordnung des Systems p_1, $p_2 \ldots p_n$.

Entwickelt man die rechte Seite von (5) nach Potenzen von p_0, so erhält man

$$n \cdot \pi_\nu^\nu = s_\nu - \binom{\nu}{1} s_{\nu-1} p_0 + \binom{\nu}{2} s_{\nu-2} p_0^2 - \cdots \pm n p_0^\nu \tag{6}$$

Es ist somit:

$$\begin{aligned} n\pi_1 &= s_1 - np_0 = 1 - np_0; \\ n\pi_2^2 &= s_2 - 2s_1 p_0 + np_0^2; \\ n\pi_3^3 &= s_3 - 3s_2 p_0 + 3s_1 p_0^2 - np_0^3; \\ &\ldots\ldots\ldots\ldots \end{aligned}$$

Hieraus wird ersichtlich, dass nach Wahl eines bestimmten Ausgangswerthes p_0 einerseits π_1, $\pi_2 \ldots \pi_n$ aus s_1, $s_2 \ldots s_n$ und anderseits s_1, $s_2 \ldots s_n$ aus π_1, $\pi_2 \ldots \pi_n$ gefunden werden können. Das System der Wahrscheinlichkeitswerthe p_1, $p_2 \ldots p_n$ wird somit durch die auf einen beliebig gewählten Ausgangswerth p_0 bezogenen Mittelwerthe π_1, $\pi_2 \ldots \pi_n$ vollständig bestimmt.

Eine angenäherte Bestimmung wird durch eine kleinere Anzahl von Mittelwerthen, etwa durch π_1, $\pi_2 \ldots \pi_\nu$ (wo $\nu < n$) geleistet. Denn aus π_1, $\pi_2 \ldots \pi_\nu$ kann man s_1, $s_2 \ldots s_\nu$ und somit b_1, $b_2 \ldots b_\nu$ finden, so dass in der Gleichung (1) die ν ersten Coefficienten bekannt sind, die $n - \nu$ folgenden aber unbekannt bleiben. Dabei sind aber die unbekannt bleibenden Coefficienten keineswegs willkürlich bestimmbar, sondern an gewisse Bedingungen gebunden, die — allgemein gesagt — darin ihre Quelle haben, dass die n Wurzeln der Gleichung (1) sämmtlich reell und positiv sind.

Demzufolge sind, wenn die Gleichung (1) in der Form

$$p^n - \binom{n}{1} \beta_1 p^{n-1} + \binom{n}{2} \beta_2^2 p^{n-2} - \ldots \pm \beta_n^n = 0 \tag{7}$$

geschrieben, wenn also

$$b_i = \binom{n}{i} \beta_i^i \quad \text{für } i = 1,\ 2 \ldots n$$

gesetzt wird, die Werthe $\beta_1, \beta_2 \ldots \beta_n$ gleichfalls reelle Mittelwerthe der Größen $p_1, p_2 \ldots p_n$. Insbesondere ist β_1 das arithmetische und β_n das geometrische Mittel jener Größen. Und es bestehen zwischen diesen Mittelwerthen [wie im nächsten Capitel (§ 1) gezeigt wird] Beziehungen, in denen die Thatsache, dass die n Wurzeln von (7) sämmtlich reell und positiv sind, einen Ausdruck findet. Beispielsweise ist

$$\beta_1 > \beta_2 > \beta_3 > \cdots > \beta_{n-1} > \beta_n . \tag{8}$$

Es ist aber nun zu beachten, dass die Wahrscheinlichkeitswerthe $p_1, p_2 \ldots p_n$ im allgemeinen nur innerhalb gewisser Grenzen als zuverlässig gelten können. Man wird darum nicht erwarten dürfen, bei einer erneuten, selbst unter gleichen Umständen ausgeführten Bestimmung genau dieselben Werthe wieder zu finden. Mit den Wahrscheinlichkeitswerthen ändern sich aber im allgemeinen auch ihre Mittelwerthe nach Maßgabe der Definitionsgleichung (5), so dass $\pi_1 + \varDelta_1, \pi_2 + \varDelta_2 \ldots \pi_n + \varDelta_n$ an Stelle von $\pi_1, \pi_2 \ldots \pi_n$ tritt, wenn $p_1, p_2 \ldots p_n$ durch $p_1 + \delta_1, p_2 + \delta_2 \ldots p_n + \delta_n$ ersetzt wird, und

$$n(\pi_\nu + \varDelta_\nu)^\nu = (p_1 - p_0 + \delta_1)^\nu + (p_2 - p_0 + \delta_2)^\nu + \cdots + (p_n - p_0 + \delta_n)^\nu$$

ist. Dabei muss

$$\delta_1 + \delta_2 + \ldots + \delta_n = 0$$

sein, indem sowohl $p_1 + p_2 + \ldots + p_n$ als auch $p_1 + \delta_1 + p_2 + \delta_2 + \ldots + p_n + \delta_n$ gleich 1 ist. Da indessen $\delta_1, \delta_2 \ldots \delta_n$ als kleine

Größen vorauszusetzen sind, so darf man die zweiten und höheren Potenzen derselben vernachlässigen und in erster Annäherung

$$n\pi_\nu^\nu + n\nu\pi_\nu^{\nu-1}\Delta_\nu = (p_1 - p_0)^\nu + (p_2 - p_0)^\nu + \ldots + (p_n - p_0)^\nu$$
$$+ \nu\delta_1(p_1 - p_0)^{\nu-1} + \nu\delta_2(p_2 - p_0)^{\nu-1} + \ldots + \nu\delta_n(p_n - p_0)^{\nu-1}$$

oder

$$\Delta_\nu = \frac{1}{n\pi_\nu^{\nu-1}}(\delta_1(p_1 - p_0)^{\nu-1} + \delta_2(p_2 - p_0)^{\nu-1} + \ldots + \delta_n(p_n - p_0)^{\nu-1})$$

setzen. Hieraus erhält man schließlich, wenn

$$\delta_n = -\delta_1 - \delta_2 - \ldots - \delta_{n-1}$$

und zur Abkürzung

$$c_1 = (p_1 - p_0)^{\nu-1} - (p_n - p_0)^{\nu-1};\ c_2 = (p_2 - p_0)^{\nu-1} - (p_n - p_0)^{\nu-1};$$

u. s. w. gesetzt wird,

$$\Delta_\nu = \frac{1}{n\pi_\nu^{\nu-1}}(\delta_1 c_1 + \delta_2 c_2 + \ldots + \delta_{n-1} c_{n-1}) \,. \tag{9}$$

Denkt man sich jetzt die Bestimmung der Wahrscheinlichkeitswerthe $p_1, p_2 \ldots p_n$ beliebig oft wiederholt, so werden andere und andere, positive und negative Abweichungswerthe $\delta_1, \delta_2 \ldots \delta_{n-1}$ resultiren und zugehörige Werthe Δ_ν erzeugen. Es lässt sich dann aus den Quadraten der Abweichungswerthe δ_1 der mittlere Werth m_1 berechnen und nach Gauß als Maßstab für die bei Bestimmung von p_1 zu befürchtende Unsicherheit ansehen. Desgleichen existirt je ein mittlerer Werth $m_2, m_3 \ldots m_{n-1}$ für die Abweichungswerthe $\delta_2, \delta_3 \ldots \delta_{n-1}$. Aus denselben wird der mittlere Werth der Abweichungen Δ_ν, der durch M_ν bezeichnet werden soll, nach der Formel[1])

$$M_\nu = \frac{1}{n\pi_\nu^{\nu-1}}\sqrt{m_1^2 c_1^2 + m_2^2 c_2^2 + \ldots + m_{n-1}^2 c_{n-1}^2} \tag{10}$$

gefunden. Er dient unabhängig von dem Gesetze, das die Häufigkeit des Auftretens der verschiedenen, möglichen Abweichungswerthe $\delta_1, \delta_2 \ldots \delta_{n-1}$ regelt, als Maßstab für die bei Bestimmung von π_ν zu erwartenden Schwankungen.

---

1) Gauß, Theoria combinationis observationum erroribus minimis obnoxiae, Art. 18.

Von den Schwankungen der thatsächlich beobachteten Wahrscheinlichkeitswerthe $p_1, p_2 \ldots p_n$ abgesehen ist noch in Rücksicht zu ziehen, dass stets Varianten mit geringen Wahrscheinlichkeitswerthen vorhanden sein können, die unter den zur Beobachtung gelangenden Exemplaren des C.G. nicht vertreten sind. Es mögen r solcher Varianten mit den Wahrscheinlichkeitswerthen $p_{n+1}, p_{n+2} \ldots p_{n+r}$ vorhanden sein. Alsdann müssen an Stelle von (4) die Summenwerthe

$$s_\nu = p_1^\nu + \ldots p_n^\nu + p_{n+1}^\nu + \ldots p_{n+r}^\nu$$

treten, die indessen von den ursprünglichen Werthen nicht merklich verschieden sein werden, da mit Rücksicht auf die vorauszusetzenden geringen Werthe von $p_{n+1}, p_{n+2} \ldots p_{n+r}$ die Summe

$$p_{n+1}^\nu + p_{n+2}^\nu + \ldots + p_{n+r}^\nu \quad \text{neben} \quad p_1^\nu + p_2^\nu + \ldots + p_n^\nu$$

nicht in Betracht kommt. Insbesondere ist nach wie vor $s_1 = 1$.

Anders verhält es sich mit den Mittelwerthen π_ν, die nach Vornahme der Correctur durch (π_ν) bezeichnet werden sollen. Denn es tritt nun an Stelle von (5) vielmehr

$$(n+r)\cdot(\pi_\nu)^\nu = (p_1 - p_0)^\nu + \ldots + (p_n - p_0)^\nu + (p_{n+1} - p_0)^\nu \\ + \ldots + (p_{n+r} - p_0)^\nu$$

oder in Anbetracht der Kleinheit von $p_{n+1}, \ldots p_{n+r}$

$$(n+r)\cdot(\pi_\nu)^\nu = (p_1 - p_0)^\nu + \ldots + (p_n - p_0)^\nu + (-p_0)^\nu \cdot r. \qquad (11)$$

Man erhält daher an Stelle von (6) jetzt:

$$(n+r)\cdot(\pi_\nu)^\nu = s_\nu - \binom{\nu}{1} s_{\nu-1} p_0 + \binom{\nu}{2} s_{\nu-2} p_0^2 - \ldots \pm (n+r) p_0^\nu, \qquad (12)$$

woraus zu ersehen ist, dass (π_ν) aus p_0 und $s_1, s_2 \ldots s_\nu$ unter der Annahme, als gehörten diese Werthe nicht zu n sondern zu $n+r$ Werthen, zu berechnen ist; und der Unterschied zwischen (π_ν) und π_ν tritt in der Gleichung

$$(n+r)\cdot(\pi_\nu)^\nu = n\cdot\pi_\nu^\nu + (-p_0)^\nu \cdot r$$

hervor, wonach

$$(\pi_\nu) = \sqrt[\nu]{\frac{n}{n+r}\pi_\nu^\nu + \frac{r}{n+r}(-p_0)^\nu}. \qquad (13)$$

Die Anzahl r kann nicht bestimmt, sondern nur mit Rücksicht auf die Beschaffenheit des Collectivgegenstandes und die Umstände, unter denen die Bestimmung der Wahrscheinlichkeitswerthe vorgenommen wurde, abgeschätzt werden. Sie ist im allgemeinen als klein anzunehmen. Insbesondere kann sie gleich 1 vorausgesetzt werden, was darauf hinaus kommt, alle der Beobachtung sich entziehenden Varianten zu einer einzigen vereinigt zu denken. Bei dieser offenbar erlaubten Annahme wird

$$(\pi_\nu) = \sqrt[\nu]{\frac{n}{n+1}\pi_\nu^\nu + \frac{1}{n+1}(-p_0)^\nu} \tag{13a}$$

also

$$(\pi_1) = \frac{1}{n+1} - p_0$$

$$(\pi_2) = \sqrt{\frac{n}{n+1}\pi_2^2 + \frac{1}{n+1}p_0^2}$$

u. s. w.

Das arithmetische Mittel ist nun gleich $\frac{1}{n+1}$.

Wählt man dasselbe als Ausgangswerth p_0, so wird

$$(\pi_1) = 0$$

$$(\pi_2) = \sqrt{\frac{n}{n+1}\pi_2^2 + \frac{1}{(n+1)^3}}$$

$$= \sqrt{\frac{1}{n+1}(p_1^2 + \cdots + p_n^2) - \frac{1}{(n+1)^2}},$$

und dieser Werth kann als Maß für die durchschnittliche Schwankung der Wahrscheinlichkeitswerthe um ihr arithmetisches Mittel dienen.

§ 3. Die Charakterisirung eines Collectivgegenstandes auf Grund der Vertheilungstafel.

Lassen sich alle unterscheidbaren Varianten des gegebenen C.G. in eine einzige Reihe ordnen, die in der Reihe der Ordnungszahlen

$$\ldots \alpha - 1,\ \alpha,\ \alpha + 1, \ldots$$

ihr Bild findet, so gibt es stets ein endliches, durch eine kleinste und eine größte Ordnungszahl $\alpha - \beta$ und $\alpha + \gamma$ begrenztes, mehr oder minder ausgedehntes Intervall der Zahlenreihe, das von den empirisch gefundenen Häufigkeiten in Anspruch genommen wird. Denn man kann nur eine endliche Anzahl von Varianten und demgemäß nur ein in bestimmter Weise abgegrenztes Gebiet der Variantenreihe (die im allgemeinen als unbegrenzt oder in unbestimmter Weise begrenzt vorauszusetzen ist) empirisch bestimmen.

Die Varianten außerhalb jenes Gebietes besitzen alsdann durchweg die empirische Häufigkeit Null und sind in Wahrheit mit sehr kleinen oder auch unendlich kleinen Wahrscheinlichkeiten behaftet zu denken, die sich der Beobachtung entziehen. Sie können zunächst unberücksichtigt bleiben. Die Varianten des abgegrenzten Gebietes sind hingegen insgesammt in Rücksicht zu ziehen, mögen ihre empirischen Häufigkeiten gleich Null oder von Null verschieden sein. Denn es handelt sich um die Feststellung des Verlaufs der Wahrscheinlichkeitswerthe, wobei die Häufigkeiten aller Varianten des abgegrenzten Gebietes in Betracht kommen.

Ist nun die Anzahl der letzteren gleich n, so können die aufeinanderfolgenden Ordnungszahlen der Varianten Einfachheitshalber durch $\alpha_1, \alpha_2 \ldots \alpha_n$ bezeichnet werden, so dass die Vertheilungstafel des C.G. in der Form

$$\frac{\alpha_1, \alpha_2 \ldots \alpha_n}{z_1, z_2 \ldots z_n} \quad \text{oder} \quad \frac{\alpha_1, \alpha_2 \ldots \alpha_n}{p_1, p_2 \ldots p_n} \tag{14}$$

darstellbar ist, wo z_λ (für $\lambda = 1, 2 \ldots n$) angibt, wie oft die Variante α_λ unter $m = z_1 + z_2 + \cdots + z_n$ Exemplaren des C.G. gefunden wurde, und wo $p_\lambda = z_\lambda : m$ die relative Häufigkeit von α_λ bedeutet.

Diese Darstellungsweise der Vertheilungstafel, bei der die aufeinanderfolgenden Varianten durch aufeinanderfolgende Ordnungszahlen bezeichnet werden, kann für jeden C.G. gewählt werden. Es ist aber denkbar, dass die bei einem gewissen Grade der Unterscheidung resultirenden Varianten sich bei schärferer Unterscheidung als Zusammenfassungen von je zwei oder noch mehr Varianten ergeben. Soll nun die hierdurch bedingte neue Form der Vertheilungstafel mit der ursprünglichen vergleichbar sein, so muss man gebrochene

Zahlenwerthe zu Hülfe nehmen und z. B. α_1, $\alpha_1 + \frac{1}{2}$ oder α_1, $\alpha_1 + \frac{1}{3}$, $\alpha_1 + \frac{2}{3}$ als Zeichen für die zwei oder drei, ursprünglich gemeinsam durch α_1 bezeichneten Varianten benutzen. Es tritt alsdann an Stelle der ganzen Zahlen von α_1 bis α_n die Reihe der $2n$ oder $3n$ Zahlen

$$\alpha_1, \alpha_1 + \frac{1}{2}, \alpha_2, \alpha_2 + \frac{1}{2}, \ldots\ldots \alpha_n, \alpha_n + \frac{1}{2};$$

oder

$$\alpha_1, \alpha_1 + \frac{1}{3}, \alpha_1 + \frac{2}{3}, \alpha_2, \alpha_2 + \frac{1}{3}, \alpha_2 + \frac{2}{3}, \ldots\ldots \alpha_n, \alpha_n + \frac{1}{3}, \alpha_n + \frac{2}{3}.$$

Es kann ferner, wenn die Varianten stetig veränderliche Größen sind, aus Gründen der Bequemlichkeit geboten sein, die Maßzahlen der Varianten in einer Einheit auszudrücken, die von derjenigen, welche den Unterschied zweier aufeinanderfolgender Varianten angibt, verschieden ist. Dann werden wiederum im allgemeinen gebrochene Zahlen als Zeichen der Varianten sich einstellen.

Darum scheint es angebracht, die Vertheilungstafel in einer solchen Form vorauszusetzen, dass statt der Ordnungszahlen reelle Zahlenwerthe, die ganz oder gebrochen sein können, auftreten. Dies lässt sich ganz allgemein in folgender Weise erreichen.

Man theile das Continuum der reellen, von $-\infty$ bis $+\infty$ sich erstreckenden Zahlenwerthe irgendwie in aneinandergrenzende Intervalle. Dann folgen diese Intervalle ganz ebenso aufeinander wie die Glieder der Zahlenreihe, so dass jeder Ordnungszahl ein Intervall entspricht, wenn nur eine Ordnungszahl einem beliebigen Intervall zugeordnet wird. Besitzt nun die Mitte des zur Ordnungszahl α_λ gehörenden Intervalls den Zahlenwerth a_λ und ist die Intervalllänge gleich i_λ, so dass die obere Grenze durch $a_\lambda + \frac{1}{2} i_\lambda$, die untere durch $a_\lambda - \frac{1}{2} i_\lambda$ und das ganze Intervall durch $a_\lambda \pm \frac{1}{2} i_\lambda$ bezeichnet werden kann, so kann man den absoluten oder relativen Häufigkeitswerth z_λ oder p_λ statt der Ordnungszahl α_λ vielmehr der reellen Zahl

a_λ oder auch dem Intervall $a_\lambda \pm \frac{1}{2} i_\lambda$ zuweisen. Man erhält alsdann die Vertheilungstafel in der Form:

$$\frac{a_1, a_2 \ldots a_n}{z_1, z_2 \ldots z_n} \text{ oder } \frac{a_1, a_2 \ldots a_n}{p_1, p_2 \ldots p_n}. \tag{14a}$$

Wählt man die Intervalle so, dass sie alle die Länge 1 besitzen und ihre Mitten durch die aufeinanderfolgenden ganzen Zahlen angegeben werden, so gelangt man wieder zu der früheren Form (14) mit den Intervallen $a_1 \pm \frac{1}{2}, a_2 \pm \frac{1}{2} \ldots a_n \pm \frac{1}{2}$.

Um für eine solche Vertheilungstafel den Verlauf der Wahrscheinlichkeitswerthe festzustellen, gilt es Größen zu bestimmen, aus denen der zu einem Gliede a_λ der Reihe $a_1, a_2 \ldots a_n$ gehörige Werth p_λ gefunden werden kann.

Solche Größen sind die Summenwerthe $s_1, s_2 \ldots s_{n-1}$, die durch

$$s_\nu = p_1 a_1^\nu + p_2 a_2^\nu + \cdots + p_n a_n^\nu; \; \nu = 1, 2 \ldots n-1 \tag{15}$$

definirt werden. Denn die n Werthe $p_1, p_2 \ldots p_n$ können aus jenen Summenwerthen und den bekannten Zahlenwerthen $a_1, a_2 \ldots a_n$ und zwar in ihrer Zugehörigkeit zu den letzteren auf Grund des Systems der Gleichungen

$$\left.\begin{array}{llll} 1 & = p_1 & + p_2 & + \ldots p_n \\ s_1 & = p_1 a_1 & + p_2 a_2 & + \ldots p_n a_n \\ s_2 & = p_1 a_1^2 & + p^2 a_2^2 & + \ldots p_n a_n^2 \\ \ldots & \ldots & \ldots & \ldots \\ s_{n-1} & = p_1 a_1^{n-1} & + p_2 a_2^{n-1} & + \ldots p_n a_n^{n-1} \end{array}\right\} \tag{16}$$

berechnet werden, da die Determinante

$$\begin{vmatrix} 1 & 1 & \ldots 1 \\ a_1 & a_2 & \ldots a_n \\ \cdot & \cdot & \\ \cdot & \cdot & \\ a_1^{n-1} & a_2^{n-1} & \ldots a_n^{n-1} \end{vmatrix}$$

gleich dem Producte $(a_2 - a_1)(a_3 - a_1) \ldots (a_n - a_{n-1})$ und somit von Null verschieden ist. Es genügt demgemäß die Bestimmung der s-Werthe, die als symmetrische Functionen der z_1 Varianten a_1, z_2

Varianten $a_2 \ldots z_n$ Varianten a_n für den C.G. charakteristisch sind, um den Verlauf der p-Werthe vollständig zu bestimmen.

Die s-Werthe sind indessen, weil der Grad der Potenzen, aus denen sie zu berechnen sind, ständig wächst, nicht gleichartig. Es empfiehlt sich darum, aus ihnen andere, unmittelbar mit einander vergleichbare Werthe abzuleiten, die als Mittelwerthe der Vertheilungstafel aufzufassen sind.

Um diese Mittelwerthe ganz allgemein zu definiren, möge b eine beliebige reelle Zahl bezeichnen und an Stelle von a_λ die Differenz $a_\lambda - b$ für $\lambda = 1, 2 \ldots n$ gesetzt werden, so dass die z_λ Differenzen $a_\lambda - b$ die Rolle der z_λ Werthe a_λ in der Verteilungstafel des C.G. übernehmen. Bildet man nun die Summe

$$m \cdot \varepsilon_\nu^\nu = z_1 (a_1 - b)^\nu + z_2 (a_2 - b)^\nu + \cdots + z_n (a_n - b)^\nu \tag{17}$$

so erhält man nach Division mit m

$$\varepsilon_\nu^\nu = p_1 (a_1 - b)^\nu + p_2 (a_2 - b)^\nu + \cdots + p_n (a_n - b)^\nu \tag{17a}$$

und hieraus nach Ausziehen der ν-ten Wurzel den reellen Werth ε_ν, der für ein geradzahliges ν dem absoluten Betrage nach zu nehmen ist. Den Werth ε_ν nenne ich den auf den Ausgangswerth b bezogenen Mittelwerth ν-ter Ordnung der Vertheilungstafel des C.G.

Entwickelt man (17a) nach Potenzen von b, so resultirt mit Rücksicht auf (15)

$$\varepsilon_\nu^\nu = s_\nu - \binom{\nu}{1} s_{\nu-1} b + \binom{\nu}{2} s_{\nu-2} b^2 - \cdots \pm b^\nu, \tag{18}$$

woraus ersichtlich wird, dass für einen beliebig, aber fest gegebenen Ausgangswerth b_0 die Mittelwerthe $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_{n-1}$ aus den Summenwerthen $s_1, s_2 \ldots s_{n-1}$ und ebenso die letzteren aus den ersteren berechnet werden können. Die Mittelwerthe sind daher in gleichem Umfange wie die Summenwerthe zur Charakterisirung des Verlaufs der Wahrscheinlichkeitswerthe dienlich. Dies erhellt übrigens unmittelbar auf Grund des Systems der Gleichungen:

$$\left.\begin{array}{lllll} 1 & = p_1 & + p_2 & + \cdots + p_n \\ \varepsilon_1 & = p_1 (a_1 - b) & + p_2 (a_2 - b) & + \cdots + p_n (a_n - b) \\ \varepsilon_2^2 & = p_1 (a_1 - b)^2 & + p_2 (a_2 - b)^2 & + \cdots + p_n (a_n - b)^2 \\ \cdot\ \cdot\ \cdot \\ \varepsilon_{n-1}^{n-1} & = p_1 (a_1 - b)^{n-1} & + p_2 (a_2 - b)^{n-1} & + \cdots + p_n (a_n - b)^{n-1} \end{array}\right\} (19)$$

Demgemäß gilt folgender Satz:

Hat man für eine n-gliedrige Variantenreihe $a_1, a_2 \ldots a_n$ die $n-1$ Mittelwerthe $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_{n-1}$ auf Grund der Vertheilungstafel berechnet, so wird durch dieselben der zu a_λ gehörige Werth p_λ für $\lambda = 1, 2 \ldots n$ und somit der Verlauf der Wahrscheinlichkeitswerthe innerhalb der Vertheilungstafel vollständig bestimmt.

Es fragt sich jetzt, welche Bedeutung eine geringere Anzahl von Mittelwerthen hat: ob und in wie weit sie eine angenäherte Kenntniss des Verlaufs der Wahrscheinlichkeitswerthe gewährt.

Um dies klar zu legen, mögen zunächst die $n-2$ Mittelwerthe $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_{n-2}$ als bekannt vorausgesetzt werden, so dass die Gleichungen

$$\left.\begin{array}{lllll} 1 & = p_1 & + p_2 & + \cdots + p_n \\ \varepsilon_1 & = p_1 (a_1 - b) & + p_2 (a_2 - b) & + \cdots + p_n (a_n - b) \\ \cdot\ \cdot\ \cdot \\ \varepsilon_{n-2}^{n-2} & = p_1 (a_1 - b)^{n-2} & + p_2 (a_2 - b)^{n-2} & + \cdots + p_n (a_n - b)^{n-2} \end{array}\right\} (20)$$

zur Verfügung stehen.

Fasst man nun in jeder Gleichung zwei aufeinanderfolgende, mit denselben Indiceswerthen behaftete Glieder der rechten Seite, beispielsweise das erste und zweite zusammen, indem man

$$\begin{array}{l} p_1 (a_1 - b) + p_2 (a_2 - b) = (p_1 + p_2)(b_1 - b) \\ p_1 (a_1 - b)^2 + p_2 (a_2 - b)^2 = (p_1 + p_2)(b_2 - b)^2 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{array}$$

setzt, so sind $b_1 - b, b_2 - b \ldots$ Mittelwerthe der z_1 Werthe $a_1 - b_0$ und z_2 Werthe $a_2 - b$. Denn es ist ja $p_1 = z_1 : m$; $p_2 = z_2 : m$, so dass nach Multiplication mit m an Stelle von p_1 und p_2 die Anzahlen z_1 und z_2 treten. Es gehören daher — wie man sich leicht überzeugt — die Werthe $b_1, b_2 \ldots$ insgesammt dem von a_1 und a_2 begrenzten Intervall der reellen Zahlen an.

Dürfte man diese Werthe als bekannt voraussetzen, so fände man aus den $n-1$ Gleichungen

$$\begin{array}{lllll} 1 & = (p_1 + p_2) & + p_3 & + \cdots + p_n \\ \varepsilon_1 & = (p_1 + p_2)(b_1 - b) & + p_3(a_3 - b) & + \cdots + p_n(a_n - b) \\ \cdots & \cdots & \cdots & \cdots \\ \varepsilon_{n-2}^{n-2} & = (p_1 + p_2)(b_{n-2} - b)^{n-2} & + p_3(a_3 - b)^{n-2} & + \cdots + p_n(a_n - b)^{n-2} \end{array}$$

die $n-1$ Werthe $p_1 + p_2$, $p_3 \ldots p_n$ und zwar die Summe $p_1 + p_2$ als zugehörig zu den beiden Werthen a_1 und a_2, zwischen denen b_1, $b_2 \ldots b_{n-2}$ liegen, während $p_3 \ldots p_n$ in ihrer Zugehörigkeit zu $a_3 \ldots a_n$ bestimmt werden. Denn die Determinante dieses Systems von Gleichungen ist im allgemeinen von Null verschieden, obschon sie in besonderen Fällen gleich Null sein kann.

In Wirklichkeit sind aber, da lediglich die Mittelwerthe ε_1, $\varepsilon_2 \ldots \varepsilon_{n-2}$ und die Zahlenreihe a_1, $a_2 \ldots a_n$ nebst dem Ausgangswerthe b als gegeben anzunehmen sind, die Werthe b_1, $b_2 \ldots b_{n-2}$ unbekannt, und man weiß nur, dass sie gewisse, für die Mittelwerthe allgemein gültige (im nächsten Capitel entwickelte) Bedingungen erfüllen müssen. Man ist daher darauf angewiesen, irgend welche mit jenen Bedingungen verträgliche Werthe für b_1, $b_2 \ldots b_{n-2}$ vorauszusetzen oder auch willkürlich ohne Rücksicht auf jene Bedingungen zwischen a_1 und a_2 anzunehmen. Beispielsweise kann man der Einfachheit wegen

$$b_1 = b_2 = \cdots = \frac{1}{2}(a_1 + a_2)$$

setzen. Bezeichnet man diesen Werth durch a_1^1, so lässt sich aus

$$\left.\begin{array}{lllll} 1 & = (p_1 + p_2) & + p_3 & + \cdots + p_n \\ \varepsilon_1 & = (p_1 + p_2)(a_1^1 - b) & + p_3(a_3 - b) & + \cdots + p_n(a_n - b) \\ \cdots & \cdots & \cdots & \cdots \\ \varepsilon_{n-2}^{n-2} & = (p_1 + p_2)(a_1^1 - b)^{n-2} & + p_3(a_3 - b)^{n-2} & + \cdots + p_n(a_n - b)^{n-2} \end{array}\right\} \quad (21)$$

eine angenäherte Bestimmung der $n-1$ Werthe $p_1 + p_2$, $p_3 \ldots p_n$ in ihrer Zugehörigkeit zu dem Zahlenpaare a_1, a_2 und zu den Zahlen a_3, a_4, $\ldots a_n$ gewinnen, die in jedem Falle ausführbar ist, da die Determinante dieses Gleichungensystems stets von Null verschieden ist. Und eine solche angenäherte Bestimmung der Wahrscheinlichkeitswerthe ist schon aus dem Grunde hinreichend, weil die empirisch gefundenen Wahrscheinlichkeiten im allgemeinen nur innerhalb gewisser Grenzen als zuverlässig gelten können.

Setzt man weiterhin $n-3$ Mittelwerthe $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_{n-3}$ als bekannt voraus, so kann man entweder drei aufeinanderfolgende Varianten oder zweimal je zwei aufeinanderfolgende Varianten zusammennehmen. Fasst man z. B. die drei ersten Varianten a_1, a_2, a_3 oder die beiden ersten a_1 und a_2 und dann die beiden folgenden a_3 und a_4 zusammen, so findet man aus den nun zur Verfügung stehenden $n-2$ Gleichungen entweder die angenäherten Werthe von $p_1+p_2+p_3, p_4 \ldots p_n$ in ihrer Zugehörigkeit zu dem Werthentripel a_1, a_2, a_3 und den Werthen $a_4 \ldots a_n$, oder die angenäherten Werthe von $p_1+p_2, p_3+p_4, p_5 \ldots p_n$ in ihrer Zugehörigkeit zu den Werthenpaaren a_1, a_2 und a_3, a_4 und den Werthen $a_5 \ldots a_n$.

Zugleich erhellt, dass in ganz entsprechender Weise eine angenäherte Bestimmung des noch mehr reducirten Verlaufs der Wahrscheinlichkeitswerthe durch eine noch kleinere Anzahl von Mittelwerthen geleistet wird. Man gelangt daher zu folgender Erkenntniss.

Hat man für eine n-gliedrige Variantenreihe $a_1, a_2 \ldots a_n$ die ν Mittelwerthe $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_\nu$, wo $\nu < n-1$, auf Grund der Vertheilungstafel berechnet, so wird durch dieselben der Verlauf der Wahrscheinlichkeitswerthe angenähert und zwar in der Weise bestimmt, als wenn in der Variantenreihe benachbarte Glieder zusammengefasst und nur noch $\nu+1$ Varianten mit ihren zugehörigen Wahrscheinlichkeitswerthen unterschieden würden.

Die Wahl des Ausgangswerthes ist, soweit nicht besondere Gründe denselben bestimmen, gleichgültig, da man aus den für b berechneten Mittelwerthen $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_\nu$ zu dem für $b+l$ geltenden Mittelwerthe ν-ter Ordnung mittelst der Formel

$$\varepsilon_\nu^\nu + \binom{\nu}{1} l \varepsilon_{\nu-1}^{\nu-1} + \binom{\nu}{2} l^2 \varepsilon_{\nu-2}^{\nu-2} + \cdots + l^\nu \tag{22}$$

gelangt, die aus (17a) durch Entwicklung nach dem binomischen Lehrsatze folgt, wenn $b+l$ an Stelle von b gesetzt wird. Im allgemeinen empfiehlt es sich, das arithmetische Mittel der z_1 Größen a_1, z_2 Größen $a_2 \ldots z_n$ Größen a_n oder

$$b = p_1 a_1 + p_2 a_2 + \cdots + p_n a_n$$

als Ausgangswerth zu wählen. In diesem Falle ist $\varepsilon_1 = 0$ und der Werth ε_2 kann als Maßstab für die Ausdehnung der Vertheilungstafel dienen. Man ersieht hieraus, dass die beiden Mittelwerthe ε_1 und ε_2 den arithmetischen Mittelwerth nebst der Ausdehnung der Vertheilungstafel charakterisiren und somit das Minimum an Bestimmungsstücken für einen C.G. darstellen.

Es erübrigt noch der Unsicherheit, mit der die empirische Bestimmung der Wahrscheinlichkeitswerthe im allgemeinen verknüpft ist, Rechnung zu tragen.

Wenn die Beobachtung von m Exemplaren eines C.G. zu z_1 Exemplaren a_1, z_2 Exemplaren $a_2 \ldots z_n$ Exemplaren a_n geführt hat, die in der Reihenfolge, in der sie sich der Beobachtung dargeboten haben, als m Exemplare a', a'', $\ldots$ $a^{(m)}$ notirt werden mögen, so kann man nicht erwarten, bei einer erneuten Beobachtung von m Exemplaren die nämlichen Werthe wieder zu finden. An Stelle von a', $a'' \ldots a^{(m)}$ wird vielmehr die Reihe x_1, $x_2 \ldots x_m$ auftreten, so dass auch die Summe

$$\varepsilon_\nu^\nu = \frac{1}{m}\{(a'-b)^\nu + (a^\nu - b)^\nu + \cdots + (a^{(m)} - b)^\nu\} \ ,$$

welche den nämlichen Werth wie (17a) nur in anderer Form darstellt, durch die Summe

$$\frac{1}{m}\{(x_1 - b)^\nu + (x_2 - b)^\nu + \cdots (x_m - b)^\nu\}$$

ersetzt wird, die um den Betrag

$$\varDelta = \frac{1}{m}\{(x_1 - b)^\nu + (x_2 - b)^\nu + \cdots + (x_m - b)^\nu\} - \varepsilon_\nu^\nu$$

von der anfänglichen Summe abweicht. Denkt man sich nun die Beobachtung von m Exemplaren des C.G. beliebig oft, streng genommen unendlich oft wiederholt, so ergeben sich andere und andere Summenwerthe an Stelle von ε_ν^ν und dementsprechend andere und andere Abweichungswerthe $\varDelta$, da jeder Werth x_1, $x_2 \ldots x_m$ die Reihe der Varianten durchläuft und mit jeder Variante in der ihr zukommenden Häufigkeit zusammenfällt. Der Mittelwerth M_ν aus den Quadraten der $\varDelta$ dient alsdann nach Gauß als Maß der Unsicherheit, mit der die Bestimmung von ε_ν^ν aus m Exemplaren des

C.G. behaftet ist. Die Berechnung dieses Mittelwerthes führt bei Anwendung des von Gauß[1]) benutzten Verfahrens zu

$$M_\nu = \sqrt{\frac{\varepsilon_{2\nu}^{2\nu} - \varepsilon_\nu^{2\nu}}{m}}, \tag{23}$$

wobei vorauszusetzen ist, dass die aus der Vertheilungstafel berechneten Werthe $\varepsilon_{2\nu}^{2\nu}$ und ε_ν^ν mit hinreichender Annäherung die wahren Werthe darstellen.

Um die Annäherung der beobachteten Mittelwerthe ε_ν an die wahren Werthe zu verschärfen, gilt es ferner zu berücksichtigen, dass wegen der Beschränktheit der Anzahl m von Exemplaren, die aus dem Begriffsumfang des C.G. herausgegriffen wurden, Varianten mit kleinen Wahrscheinlichkeitswerthen verborgen bleiben können, die erst bei der Durchforschung des ganzen Begriffsumfangs nothwendig zu Tag treten müssen. Da der ganze Begriffsumfang im allgemeinen aus unendlich vielen Exemplaren besteht, so kann man auch sagen, dass der aus m Exemplaren bestimmte Werth von ε_ν wegen der Endlichkeit des m zu corrigiren ist, damit er als übereinstimmend mit dem aus unendlich großen m bestimmten ε_ν angesehen werden dürfe.

Zu dieser Correction gelange ich auf Grund folgender Erwägungen.

Es ist vorauszusetzen, dass außer den bei endlichem m thatsächlich beobachteten Varianten auch die nicht beobachteten, aber denkbaren Varianten bei der Erforschung des ganzen Begriffsumfangs des C.G. (oder bei unendlich großem m) mit bestimmten, wenn auch kleinen Wahrscheinlichkeiten auftreten werden. Da aber die Summe der Wahrscheinlichkeitswerthe nach wie vor gleich 1 ist, so kann dies nur in der Weise geschehen, dass die Werthe $p_1, p_2 \ldots p_n$ nicht mehr vollständig auf $a_1, a_2 \ldots a_n$ fallen, sondern sich irgendwie auf die Gesammtheit der Varianten vertheilen. Demzufolge wird von dem zu a_λ gehörenden Werthe p_λ nur noch ein Bruchtheil, der allerdings im allgemeinen nicht erheblich von p_λ verschieden sein wird, bei a_λ verbleiben, während der Rest mit kleinen Beträgen auf die übrigen Varianten $a_{\lambda+\varkappa}$ (wo $\varkappa = \pm 1, \pm 2, \pm 3$ u. s. w.) fällt. Es gehört alsdann, wenn allgemein durch $p \cdot \gamma$, wo $\gamma < 1$, ein Bruch-

1) Theoria combinationis observ. error. min. obn. Art. 13, 15.

theil von p bezeichnet wird, zu a_λ der Werth $p_\lambda \cdot \gamma_0$, zu $a_{\lambda+1}$ der Werth $p_\lambda \cdot \gamma_1$, zu $a_{\lambda-1}$ der Werth $p_\lambda \cdot \gamma_{-1}$ und allgemein zu $a_{\lambda+\varkappa}$ der Werth $p_\lambda \cdot \gamma_\varkappa$ für $\varkappa = 0, \pm 1, \pm 2, \pm 3$ u. s. w.; so dass

$$p_\lambda = \sum_\varkappa p_\lambda \cdot \gamma_\varkappa \quad \text{oder} \quad 1 = \sum_\varkappa \gamma_\varkappa \,.$$

Das Gesetz, welches diese Vertheilung von p_λ auf die Varianten $a_{\lambda+\varkappa}$ regelt, ist unbekannt. Es soll aber die nahe liegende Annahme gemacht werden, dass jeder von den beobachteten Wahrscheinlichkeitswerthen $p_1, p_2 \ldots p_n$ nach dem nämlichen Gesetze sich vertheile[1]). Alsdann wird man für jeden Indexwerth $\lambda = 1, 2 \ldots n$

$$p_\lambda \cdot \eta_\nu^\nu = \sum_\varkappa p_\lambda \cdot \gamma_\varkappa \,(a_{\lambda+\varkappa} - a_\lambda)^\nu \tag{24}$$

$$\nu = 0, 1, 2, 3 \ldots$$

setzen dürfen, wo die Summation über alle Werthe $\varkappa = 0, \pm 1, \pm 2 \ldots$ zu erstrecken ist.

Nun tritt bei der Berechnung des corrigirten Mittelwerthes ε_ν, der durch (ε_ν) bezeichnet werden soll, die Summe

$$\sum_\varkappa p_\lambda \cdot \gamma_\varkappa \,(a_{\lambda+\varkappa} - b)^\nu$$

an Stelle von

$$p_\lambda \,(a_\lambda - b)^\nu \,.$$

Da jedoch $a_{\lambda+\varkappa} - b = (a_{\lambda+\varkappa} - a_\lambda) + (a_\lambda - b)$ und

$$(a_{\lambda+\varkappa} - b)^\nu = (a_{\lambda+\varkappa} - a_\lambda)^\nu + \binom{\nu}{1}(a_{\lambda+\varkappa} - a_\lambda)^{\nu-1}(a_\lambda - b) + \cdots + (a_\lambda - b)^\nu,$$

so erhält man mit Rücksicht auf (24)

$$\sum_\varkappa p_\lambda \cdot \gamma_\varkappa (a_{\lambda+\varkappa} - b)^\nu = p_\lambda \left\{ \eta_\nu^\nu + \binom{\nu}{1} \eta_{\nu-1}^{\nu-1} (a_\lambda - b) + \cdots + (a_\lambda - b)^\nu \right\}.$$

Es ist somit der Werth

1) Man wird also, wenn die a-Werthe — was als Regel anzunehmen — äquidistant sind, vorauszusetzen haben, dass der Bruchtheil von p_1, der auf $a_{1+\varkappa}$ fällt, ebenso groß ist, wie der Bruchtheil von p_2, der auf $a_{2+\varkappa}$ fällt, oder wie der Bruchtheil von p_3, der auf $a_{3+\varkappa}$ fällt u. s. w., oder dass für jeden Werth p_λ ($\lambda = 1, 2 \ldots n$) der Bruchtheil $p_\lambda \cdot \gamma_\varkappa$ zu $a_{\lambda+\varkappa}$ gehört.

$$\varepsilon_\nu^\nu = \sum_\lambda p_\lambda (a_\lambda - b)^\nu$$

durch den Werth

$$\left.\begin{aligned}(\varepsilon_\nu)^\nu &= \sum_\lambda \sum_\varkappa p_\lambda \cdot \gamma_\varkappa (a_{\lambda+\varkappa} - b)^\nu \\ &= \varepsilon_\nu^\nu + \binom{\nu}{1} \varepsilon_{\nu-1}^{\nu-1} \eta_1 + \binom{\nu}{2} \varepsilon_{\nu-2}^{\nu-2} \eta_2^2 + \cdots + \binom{\nu}{1} \varepsilon_1 \eta_{\nu-1}^{\nu-1} + \eta_\nu^\nu\end{aligned}\right\} (25)$$

zu ersetzen. Man erhält so den corrigirten Werth von ε_ν, nämlich (ε_ν), für jedes Gesetz, welches die Vertheilung der bei endlichem m gefundenen Wahrscheinlichkeitswerthe $p_1, p_2 \ldots p_n$ auf die Reihe aller denkbaren Varianten bei unendlich großem m, d. h. bei Berücksichtigung des ganzen Begriffsumfangs des C.G. in Uebereinstimmung mit (24) regelt.

Es liegt indessen nahe, bestimmte Voraussetzungen über jenes Vertheilungsgesetz oder — was dasselbe ist — über die Werthe von $\eta_1, \eta_2, \eta_3 \ldots$ zu machen.

Zunächst wird man η_ν^ν mit $1 : m$ proportional setzen dürfen, da für $m = \infty$ die η-Werthe verschwinden müssen, indem nur für endliches m eine Correctur in Betracht kommt. Sodann wird wohl die Voraussetzung gestattet sein, dass die Vertheilung von p_λ auf die oberhalb und unterhalb a_λ liegenden Werthe $a_{\lambda+1}, a_{\lambda+2} \ldots$ und $a_{\lambda-1}, a_{\lambda-2} \ldots$ der Gruppirung der Wahrscheinlichkeitswerthe, welche die Vertheilungstafel darbietet, ähnlich sei. Unter der Voraussetzung dass der Ausgangswerth b der Mittelwerthe ε_ν das arithmetische Mittel $b = p_1 a_1 + p_2 a_2 + \cdots + p_n a_n$ ist, wird es daher als annehmbar gelten können, wenn man

$$\eta_\nu^\nu = \frac{\varepsilon_\nu^\nu}{m}; \quad \eta_1 = \varepsilon_1 = 0$$

setzt. Dann erhält man:

$$(\varepsilon_\nu)^\nu = \varepsilon_\nu^\nu + \frac{1}{m}\left\{\binom{\nu}{2} \varepsilon_{\nu-2}^{\nu-2} \varepsilon_2^2 + \binom{\nu}{3} \varepsilon_{\nu-3}^{\nu-3} \varepsilon_3^3 + \cdots + \varepsilon_\nu^\nu\right\}, \quad (25a)$$

so dass insbesondere

$$(\varepsilon_1) = \varepsilon_1 = 0$$

$$(\varepsilon_2)^2 = \varepsilon_2^2\left(1 + \frac{1}{m}\right)$$

$$(\varepsilon_3)^3 = \varepsilon_3^3\left(1+\frac{1}{m}\right)$$

$$(\varepsilon_4)^4 = \varepsilon_4^4\left(1+\frac{1}{m}\right)+\frac{6}{m}\,\varepsilon_2^4$$

$$\cdots\cdots\cdots\cdots$$

Es ist bemerkenswerth, dass die Correctur von ε_2 mit der in der Fehlertheorie üblichen, auf das gewöhnliche Fehlergesetz gegründeten insoweit übereinstimmt, als man bei hinreichend großem m

$$1+\frac{1}{m}=\frac{m}{m-1} \qquad \text{oder} \qquad m^2-1=m^2$$

setzen darf. Denn alsdann ist

$$(\varepsilon_2)^2 = \varepsilon_2^2\,\frac{m}{m-1}\,.$$

Indessen mögen die Annahmen, die zu (25a) führen, nur als mögliche, nicht aber als nothwendig gültige angesehen werden.

§ 4. Die Darstellung der Vertheilungstafel durch eine mathematische Function.

Die soeben entwickelte Methode, den Verlauf der Wahrscheinlichkeitswerthe durch eine Reihe aufeinanderfolgender Mittelwerthe zu bestimmen, ist in jedem Falle anwendbar und zur Charakterisirung eines gegebenen C.G. ausreichend.

Man kann aber wünschen, die Wahrscheinlichkeitswerthe der Vertheilungstafel (14) in mathematischer Abhängigkeit von den zugehörigen Ordnungszahlen, d. h. als Functionen der Ordnungszahlen, darzustellen. Dann ist eine Function F zu suchen, so dass

$$p_1 = F(\alpha_1)\,; \qquad p_2 = F(\alpha_2)\,; \ldots \tag{26}$$

und mithin ein bestimmtes mathematisches Gesetz jeden p-Werth aus der zugehörigen Ordnungszahl abzuleiten gestattet.

Als Argumente der Function F treten die Ordnungszahlen auf. Da es indessen im allgemeinen bequemer ist, Functionen von stetig veränderlichen, reellen Argumenten zu Grunde zu legen, so kann man sich auch die Aufgabe stellen, eine Function $f(a)$ der reellen

Variablen a zu suchen, so dass für jede in Betracht kommende Ordnungszahl α

$$F(\alpha) = \int\limits_{\alpha-\frac{1}{2}}^{\alpha+\frac{1}{2}} f(a) \cdot da \; . \tag{27}$$

Es wird so der Reihe der Ordnungszahlen das Gebiet der reellen Zahlenwerthe zugeordnet und zwar jeder Ordnungszahl ein Intervall von der Länge 1, dessen Mitte durch den Zahlenwerth der betreffenden Ordnungszahl bestimmt wird. Die Function F der Ordnungszahlen aber wird durch die über jene Intervalle erstreckten Integrale einer Function mit stetig veränderlichem, reellem Argumente dargestellt.

Nun ist es offenbar unwesentlich, dass die den Ordnungszahlen zugeordneten Intervalle gerade die Länge 1 und den Zahlenwerth der Ordnungszahlen zu Intervallmitten haben. Man kann daher von der Darstellungsform (14) der Vertheilungstafel zu der allgemeineren Form (14a) übergehen, indem man das Gebiet der reellen Zahlenmannigfaltigkeit irgendwie in aneinandergrenzende Intervalle theilt und die Reihe der Ordnungszahlen der Intervallreihe zuordnet. Gehört alsdann zu der Ordnungszahl α das Intervall $a_\alpha \pm \frac{1}{2}\, i_\alpha$ mit der Intervallmitte a_α und der Intervalllänge i_α, so ist die Function $F(\alpha)$ als das über das Intervall $a_\alpha \pm \frac{1}{2}\, i_\alpha$ erstreckte Integral

$$F(\alpha) = \int\limits_{a_\alpha-\frac{1}{2}i_\alpha}^{a_\alpha+\frac{1}{2}i_\alpha} f(a) \cdot da \tag{27a}$$

zu bestimmen.

Die Lösung dieser Aufgabe ist auf zwei verschiedenen Wegen möglich. Es ist nämlich einerseits denkbar, dass auf Grund der vorhandenen Erkenntniss bestimmte Angaben oder wenigstens einigermaßen begründete Vermuthungen bezüglich der Varianten des C.G. möglich sind, die einen Anhalt zur Ableitung des Vertheilungsgesetzes darbieten. Man kann anderseits — was für eine auf inductive Wahrscheinlichkeitserkenntniss gegründete Theorie der C.G. von vorn herein geboten ist — rein empirisch das Vertheilungsgesetz festzustellen suchen.

a. Die Deduction des Vertheilungsgesetzes aus Hypothesen.

Der an erster Stelle genannte Weg wurde in einem Specialgebiete der Lehre von den C.G., in der Fehlertheorie zuerst beschritten. Dabei kommen zwei Hypothesen in Betracht. Die eine legt beispielsweise Encke[1]) der Ableitung des Fehlergesetzes in folgender Fassung zu Grunde: »Wenn eine beliebige Anzahl gleich guter directer Beobachtungen einer unbekannten Größe gegeben ist, so bestimmt das arithmetische Mittel aus allen beobachteten Werthen den wahrscheinlichsten Werth der unbekannten Größe, so weit er aus diesen Beobachtungen folgt, ganz allein, ohne dass außer ihm noch eine andere Bedingung erforderlich und im allgemeinen zulässig ist.« Aus dieser Hypothese vom arithmetischen Mittel ergibt sich nach den Regeln der Wahrscheinlichkeitsrechnung der Wahrscheinlichkeitswerth p_x einer Abweichung x vom arithmetischen Mittel als dem wahrscheinlichsten Werthe in der Form[2])

$$p_x = \frac{h}{\sqrt{\pi}}\, xep(-h^2x^2)\cdot dx\,. \tag{28}$$

Die andere, zuerst von Hagen[3]) aufgestellte Hypothese über die Entstehung der Fehler aus Combinationen positiver und negativer Elementarfehler lautet: »Der Fehler im Resultate einer Messung ist die algebraische Summe aus einer unendlich großen Anzahl elementarer Fehler, die alle gleich groß sind, und von denen jeder einzelne ebenso leicht positiv wie negativ sein kann«. Diese Hypothese führt dazu, die Glieder der nach dem binomischen Lehrsatz aus der Potenz

$$\left(\frac{1}{2}+\frac{1}{2}\right)^n.$$

entwickelten Reihe als Wahrscheinlichkeitswerthe für die verschiedenen, positiven und negativen Fehlergrößen in Anspruch zu nehmen. Und diese Werthe lassen sich für unendlich großes n wiederum in der Form (28) darstellen.

---

1) Ueber die Methode der kleinsten Quadrate. Berliner astronomisches Jahrbuch für 1834, S. 264.

2) Es wird hier $exp\ (x)$ statt e^x gesetzt.

3) Grundzüge der Wahrscheinlichkeitsrechnung, 1837, S. 34.

Die eine und die andere Hypothese bildete den Ausgangspunkt für allgemeinere Voraussetzungen, als man nach Quetelet's[1]) Vorgang das Fehlergesetz auf C.G. anderer Gebiete (z. B. der Anthropologie, Botanik, Zoologie, Meteorologie) anwandte und sich dabei von der Unzulänglichkeit dieses Gesetzes überzeugen musste.

Der hierdurch bedingte Fortschritt in der Theorie der C.G. fand einerseits in Fechner's[2]) »Collectivmaßlehre«, anderseits in Pearson's[3]) »Mathematical theory of evolution« seine Verwirklichung.

Fechner erkannte, dass bei den C.G. an Stelle der von dem Fehlergesetz geforderten symmetrischen Gruppirung der Häufigkeitswerthe um das arithmetische Mittel vielmehr Asymmetrie als Regel zu gelten habe; und dass darum das arithmetische Mittel im allgemeinen nicht als der wahrscheinlichste Werth angesehen werden dürfe. Er suchte demzufolge nach einem Vertheilungsgesetze, welches das gewöhnliche Fehlergesetz für den Fall symmetrischer Vertheilung in sich schließt, im allgemeinen aber eine asymmetrische Gruppirung um den (vom arithmetischen Mittel verschiedenen) wahrscheinlichsten Werth bedingt. Er fand die Lösung dieser Aufgabe in der Annahme, dass ein bestimmter wahrscheinlichster Werth existirt, der die Vertheilungstafel in einen oberen und unteren Theil trennt, und dass für jeden Theil unabhängig vom anderen die Vertheilung nach dem Fehlergesetze geregelt werde. Hierdurch wird gefordert, dass die Anzahlen m' und $m_{,}$ der Häufigkeitswerthe des oberen und des unteren Theiles sich verhalten wie die einfachen Mittelwerthe e' und $e_{,}$ (d. h. die auf den wahrscheinlichsten Werth als Ausgangswerth bezogenen Mittelwerthe erster Ordnung) des oberen und des unteren Theils, so dass der wahrscheinlichste Werth nicht mehr durch die Gleichheit der beiderseitigen einfachen Mittelwerthe, sondern durch die Proportion

$$e' : e_{,} = m' : m_{,} \tag{29}$$

1) Lettres sur la théorie des probabilités, 1846. — Physique sociale, 1869.

2) Im Auftrage der kgl. Sächs. Gesellsch. d. Wissensch. herausgegeben, 1897. Die Untersuchungen Fechner's über die Probleme der Collectivmaßlehre reichen bis in die 50er Jahre des zu Ende gegangenen Jahrhunderts zurück (vergl. Collectivmaßlehre, S. 229).

3) Philosophical Transactions of the Royal Society of London, Vol. 185 A (1894), 186 A (1895); 191 A (1898).

definirt wird, und nur für den Fall, dass $m' = m$, zu $e' = e$, und somit zum arithmetischen Mittel zurückführt.

In Fechner's Collectivmaßlehre liegt demgemäß unverkennbar eine durch die Erfahrung gebotene und an Beispielen bewährte Verallgemeinerung der Hypothese vom arithmetischen Mittel vor, deren Geltungsbereich noch dadurch erweitert wird, dass bei großer Ausdehnung der Vertheilungstafel (14a) die Einzelmaße a des C.G. durch ihre Logarithmen log a ersetzt werden.

Diese erweiterte Hypothese hat den Vorzug, in jedem Falle zu einer genaueren Uebereinstimmung zwischen Theorie und Erfahrung zu führen als die Hypothese vom arithmetischen Mittel. Sie findet jedoch die — von Fechner selbst erstrebte — Grenze ihrer Leistungsfähigkeit in der Uebereinstimmung mit der letzteren Hypothese bei vorwaltender Symmetrie. Wenn also Fechner aus den Maßwerthen a eines C.G. oder aus den Logarithmen dieser Werthe auf Grund des Proportionalgesetzes den wahrscheinlichsten Werth berechnet und mit dem interpolatorisch bestimmten Maximalwerthe der Vertheilungstafel vergleicht, außerdem das arithmetische Mittel und den Centralwerth nebst den oberen und unteren Abweichungszahlen und dem oberen und unteren Abweichungsmittel angibt, — so wird zwar durch das System dieser Werthe eine Charakterisirung des C.G. erreicht; dieselbe ist aber nur dann ausreichend, wenn die Asymmetrie eine bloße Störung des normalen, durch das gewöhnliche Fehlergesetz bedingten Verlaufs der Tafelwerthe bildet, so dass nach Wegfall der Asymmetrie dieses Gesetz zu ungetrübter Geltung käme. Da dies nicht allgemein zutrifft, so kann durch die Methode Fechner's nur eine beschränkte Classe von C.G. ausreichend charakterisirt werden.

Eine Erweiterung der Hypothese über die Entstehung der Fehler aus Combinationen positiver und negativer Elementarfehler bildet den Ausgangspunkt der Theorie Pearson's, die in der Abhandlung[1]) »Contributions to the mathematical theory of evolution. Skew variation in homogeneous material« dargelegt wird. Pearson setzt zunächst an Stelle der Potenz $\left(\frac{1}{2}+\frac{1}{2}\right)^n$ die allgemeinere $(p+q)^n$, wo p und q ein beliebiges Verhältniss haben. Er sagt: »Just as the

1) Philosoph. Transactions Roy. Soc. London, 186 A, 1895, S. 343—414.

normal frequency curve may be obtained by running a continuous curve through the point-binomial $\left(\frac{1}{2}+\frac{1}{2}\right)^n$ when n is fairly large, so a more general form of the probability curve may be obtained by running a continuous curve through the general binomial $(p+q)^n$. As the great and only true test of the normal curve is: Does it really fit observations and measurments of a symmetrical kind? so the best argument for the generalised probability curve deduced in this paper is that it does fit, and fit surprisingly accurately observations of an asymmetrical character. Indeed, there are very few results which have been represented by the normal curve which do not better fit the generalised probability curve, — a slight degree of asymmetry being probably characteristic of nearly all groups of measurments«. In diesen Worten zeigt sich die klare Erkenntniss, dass Asymmetrie als Regel vorauszusetzen und dementsprechend das gewöhnliche Fehlergesetz durch ein asymmetrisches zu ersetzen sei. Zugleich wird die Erfahrung als einziger Prüfstein anerkannt. Pearson stellt sich somit das nämliche Problem wie Fechner; er erstrebt jedoch die Lösung mit anderen Hülfsmitteln.

Die stetige Curve, welche der discreten, durch Entwicklung von $(p+q)^n$ nach dem binomischen Satze resultirenden Werthenreihe zur Seite steht, findet Pearson mittelst einer ihm eigenthümlichen Methode bestimmt durch

$$p_x = p_0\left(1+\frac{x}{a}\right)^{\gamma\, a} \cdot exp(-\gamma x)\,. \tag{30}$$

Dass hierbei in der That die erwähnte Hypothese, die hier in der Annahme von Elementarursachen (»contributory causes«) an Stelle der positiven und negativen Elementarfehler besteht, den Ausgangspunkt bildet, erhellt aus den Worten: »So long as we remain in ignorance of the nature and number of contributory causes in physics and biology, so long as we do find markedly skew distributions, it seems to me that we must seek more general results than flow from the assumption that $p=q$ and $n=\infty$. The form of curve given above is suggested as a possible form of skew frequency curves. Its justification lies essentially, like that of the normal curve, in its capacity to express statistical observations«.

Nun tritt aber ein weiteres Moment in den Gedankengang Pearson's. Die verallgemeinerte Wahrscheinlichkeitscurve ist nach einer Richtung begrenzt, nach der anderen unbegrenzt. »This limitation at one end only, corresponds theoretically to many cases in economics, physics, and biology. But there are a great variety of cases in which there is theoretical limitation at both ends; that is to say, there is a limited range of possible deviations«. Und nach Angabe von Beispielen, in denen eine beiderseits begrenzte (z. B. zwischen den Argumentwerthen 0 und 1 verlaufende) Vertheilungstafel theoretisch gefordert wird, sagt Pearson: »Thus the problem of range seems a very important one, it theoretically excludes the use of the normal curve in many classes of statistics; it is quite true that, for many practical purposes, frequency curves of limited range may be sensibly identical either with unlimited curves, or even with normal curves, but, in other cases, this is not so, and under any circumstances the limited curve may actually give information as to the possible range — the limits of stability — which is itself of great value. — We have, thus, reached this point: that to deal effectively with statistics we require generalised probability curves which include the factors of skewness and range.«

Dementsprechend stellt sich Pearson die Aufgabe, fünf Curventypen zu finden: 1) beiderseitig begrenzte asymmetrische; 2) beiderseitig begrenzte symmetrische; 3) einseitig begrenzte asymmetrische; 4) beiderseits unbegrenzte asymmetrische; 5) beiderseits unbegrenzte symmetrische.

Es ist somit die zunächst vorausgesetzte Potenz $(p+q)^n$ nicht ausreichend zur Entwicklung des allgemeinen Vertheilungsgesetzes, da sich eine nur einseitig begrenzte Wahrscheinlichkeitscurve ergab. Pearson geht darum von der allgemeineren Reihe[1])

$$\frac{pn(pn-1)\ldots(pn-r+1)}{n(n-1)\ldots(n-r+1)}\left\{1+r\frac{qn}{pn-r+1}+\ldots\ldots\right\}$$

aus, wo $pn+qn=n$ und $r<n$; und entwickelt aus ihr nach der schon benutzten Methode die beiden stetigen Vertheilungsgesetze:

1) Dieselbe reducirt sich für $n=\infty$ auf die Reihe

$$(p+q)^r = p^r + r\cdot p^{r-1}q + \frac{r(r-1)}{1\cdot 2}p^{r-2}q^2 + \cdots$$

$$p_x = p_o\left(1+\frac{x}{a_1}\right)^{\nu a_1}\left(1-\frac{x}{a_2}\right)^{\nu a_2} \tag{31}$$

und

$$p_x = p_o \cdot \cos^{2m}\vartheta \cdot exp(-\nu\vartheta)\,;\; x = a \operatorname{tang}\vartheta\,, \tag{32}$$

wo je nach dem Werthe gewisser Constanten die eine oder die andere Function in Kraft tritt. Zur Rechtfertigung sagt er: »Until we know very much more definitely, than we do at present, how the size of an organ in any individual, say, depends on the sizes of the same organ in its ancestors, ore what are the nature of the causes which lead to the determination of prices, or of income, or of mortality at a given age, I do not see that we have any right to select as our sole frequency curve the normal type

$$y = y_o \cdot c^{-px^2}$$

in preference to the far more general

$$y = y_o\left(1+\frac{x}{a_1}\right)^{\nu a_1}\left(1-\frac{x}{a_2}\right)^{\nu a_2},$$

which not only includes the former, but supplies the element of skewness which is undoubtedly present in many statistical frequency distributions. As we may look upon the former as a limit to a coin-tossing series, so the latter represents a limit to teetotum-spinning and card-drawing experiments. It is not easy to realise why nature or economics should, from the standpoint of chance, be more akin to tossing than to teetotum-spinning or card-dealing. At any rate, from purely utilitarian and prudent motives, we are justified so long as the analysis is manageable, in using the more general form. It will always give us a measure of the divergence of particular statistics from the normal type, and in many cases of skew frequency, it can be used when it would be the height of absurdity to apply the normal curve at all.«

Der kritische Werth, von dem die Anwendbarkeit des einen oder anderen Curventypus abhängt, ist in der Schreibweise Pearson's:

$$2\mu_2(3\mu_2^2 - \mu_4) + 3\mu_3^2\,, \tag{33}$$

wo μ_2, μ_3, μ_4 die sogenannten Momente der Vertheilungstafel oder Häufigkeitscurve bezüglich des arithmetischen Mittels sind. Diese

Momente stehen in einfacher Beziehung zu den von mir zur Charakterisirung der C.G. eingeführten Mittelwerthen. Es ist nämlich

$$\mu_2 = \varepsilon_2^2\,;\ \mu_3 = \varepsilon_3^3\,;\ \mu_4 = \varepsilon_4^4\,,$$

wenn als Ausgangswerth der Mittelwerthe ε_2, ε_3, ε_4 das arithmetische Mittel gewählt wird.

So kommt denn Pearson schließlich zu dem Resultate: »We may say that a skew frequency curve will have limited range, range limited in one direction only, or unlimited range according as

$$2\mu_2(3\mu_2^2 - \mu_4) + 3\mu_3^2$$

is greater than, equal to or less than zero. Thus the calculation of this expression is the first step towards the classification of a frequency curve given by observation.«

Demnach wäre die Variantenreihe eines C.G. bei asymmetrischer Gruppirung der Häufigkeitswerthe um das arithmetische Mittel entweder beiderseits begrenzt oder nur einerseits begrenzt oder beiderseits unbegrenzt, je nachdem der Werth von

$$2\varepsilon_2^2(3\varepsilon_2^4 - \varepsilon_4^4) + 3\varepsilon_3^6 \tag{33a}$$

größer als Null, gleich Null oder kleiner als Null ist, wo ε_2, ε_3, ε_4 die auf das arithmetische Mittel als Ausgangswerth bezogenen Mittelwerthe zweiter, dritter oder vierter Ordnung sind [1]).

Dies hätte zur Folge, dass die oberen und unteren extremen Werthe der Variantenreihe, die natürlich für eine endliche Anzahl m von beobachteten Exemplaren des C.G. unter allen Umständen sich im Endlichen halten, für unbegrenzt wachsendes m entweder einem bestimmten, angebbaren Werthe sich asymptotisch nähern oder

1) In Uebereinstimmung damit sagt Duncker in seiner Darlegung der Methode Pearson's (»Die Methode der Variationsstatistik«. Archiv für Entwicklungsmechanik der Organismen, VIII. S. 140): »Die asymmetrischen Curventypen besagen, dass die beiden Gruppen positiv und negativ wirksamer Elementarursachen der individuellen Variation ungleich, die asymmetrischen, dass dieselben gleich groß sind. Ferner bedeutet Begrenztheit in der Abscissenaxe des betr. Typus, dass die gleichzeitig auf einen Organismus einwirkende Quantität von Elementarursachen eine endliche, die Unbegrenztheit, dass diese Quantität unendlich groß ist.« — Mit der Anzahl der Elementarursachen wächst aber die Reihe der möglichen Varianten. Ist nämlich die Anzahl der positiv wirkenden Ursachen gleich m, die Anzahl der negativ wirkenden gleich n, so sind $m + n + 1$ Varianten möglich.

keinen solchen Grenzwerth besitzen, je nachdem auf Grund der Momente μ_2, μ_3, μ_4 oder der Mittelwerthe ε_2, ε_3, ε_4 der obige kritische Werth positiv oder negativ wird. (Der Nullwerth kann von vornherein keine besondere Bedeutung haben, da die Momente oder Mittelwerthe im allgemeinen nicht exact bestimmbar sind).

Dass hier ein Fehlschluss vorliegt, der zu Illusionen über die Beschaffenheit des C.G. führt, erhellt aus folgender Bemerkung[1]).

Hat man empirisch die Variantenreihe A_1, $A_2 \ldots A_n$ mit den Wahrscheinlichkeitswerthen p_1, $p_2 \ldots p_n$ festgestellt, so bleibt man völlig in Unkenntniss darüber, ob keine oder nur einzelne oder unbegrenzt viele Varianten, die mit sehr kleinen oder unendlich kleinen Wahrscheinlichkeitswerthen behaftet zu denken sind, eben wegen der geringen Häufigkeit ihres Auftretens sich der Beobachtung entzogen haben. Erst die im allgemeinen nicht durchführbare Erforschung aller Exemplare des C.G. würde auf inductivem Wege zu einer Entscheidung führen. Es lassen sich daher etwaige Grenzen der Variantenreihe nur deductiv, auf Grund anderweitig gewonnener und bereits vorhandener Erkenntniss angeben. — Demzufolge steht es mit den Principien der inductiven Wahrscheinlichkeitserkenntniss nicht in Einklang, wenn Pearson sagt[2]): »We may not know the range, a priori, but we are quite certain that one exists, and it is a quantity to be determined — just as the mean or the standard deviation — from our measurments themselves«. In der That liegen auch in den von Pearson angeführten Beispielen begrenzter Variation durchweg deductive Erkenntnisse vor, denen zufolge ein Ueberschreiten bestimmter (angebbarer oder nicht angebbarer) Grenzen undenkbar ist.

Mag man aber von vorn herein über die möglichen extremen Werthe einer Variantenreihe orientirt sein oder nicht, so können doch in jedem Falle die empirisch bestimmten Wahrscheinlichkeitswerthe p_1, $p_2 \ldots p_n$ jedes mit der Bedingung $p_1 + p_2 + \cdots + p_n = 1$ verträgliche System reeller positiver Werthe darstellen. Es kann

---

1) Dieselbe ergibt sich unmittelbar als Folgerung aus dem am Schluss des I. Cap. (§ 7) aufgestellten Grundsatze.

2) a. a. O. S. 359.

daher jener kritische Werth positiv oder negativ werden, sowohl wenn ein endlicher Grenzwerth für die Variantenreihe angebbar, als auch wenn ein solcher Grenzwerth nicht angebbar ist. Man ist somit auch nicht befugt, aus dem positiven oder negativen Vorzeichen jenes kritischen Werthes auf Begrenztheit oder Unbegrenztheit der Variantenreihe zu schließen[1]).

Vielmehr zeigt sich hierdurch, dass Pearson's Theorie nicht für alle Fälle ausreichen kann. Sie stellt nur für solche Fälle ein möglicher Weise gültiges Vertheilungsgesetz zur Verfügung, in denen eine asymmetrische Vertheilungstafel bei negativem kritischen Werthe theoretisch unbegrenzt oder bei positivem kritischen Werthe theoretisch begrenzt ist. Sie versagt hingegen von vorn herein in den Fällen, wo ein negativer kritischer Werth mit theoretischer Begrenztheit und ein positiver kritischer Werth mit theoretischer Unbegrenztheit der Variantenreihe verknüpft ist.

b. Die Unzulänglichkeit der Hypothesen.

Wenn, wie aus Vorstehendem erhellt, weder Fechner's noch Pearson's Theorie eine allgemeine Lösung des Problems, C.G. jeder möglichen Art ausreichend zu bestimmen, bietet, so liegt der wesentliche Grund nicht in der zu engen Fassung der Hypothesen, sondern in der Natur des Problems.

Die Function, durch welche eine Vertheilungstafel dargestellt werden soll, ist nämlich von vorn herein nur an die Bedingung gebunden, für jeden in Betracht kommenden Argumentwerth einen endlichen, positiven Werth darzubieten oder gleich Null zu werden. Im übrigen kann sie stetig oder unstetig sein; sie kann nur ein

1) Es verdient erwähnt zu werden, dass auch Fechner's Collectivmaßlehre (im XX. Cap.) Extremgesetze gibt. Fechner glaubt aber nicht aus beobachteten Werthen die absolut möglichen Extreme, wenn auch nur angenähert, bestimmen zu können. Er sagt vielmehr (S. 322): »Aber diese Annahme einer approximativ erreichbaren Grenze der Extreme bei wachsendem m hat weder empirisch noch theoretisch etwas für sich; sondern wahr ist nur nach beiden Gesichtspunkten, dass die Größe der Extreme in sehr viel kleinerem Verhältniss als die Größe des m wächst, aber, wenn m bis ins Unendliche steigend gedacht wird, immer in angebbarer Weise mit fortwächst«. — Dementsprechend werden die im Werke Fechner's entwickelten Extremgesetze aus dem gewöhnlichen Fehlergesetze deducirt.

Maximum oder mehrere relative Maxima besitzen und überhaupt jede denkbare Gesetzmäßigkeit in ihrem Verlaufe zeigen, da keine einschränkende Bestimmungen von allgemeiner Geltung angebbar sind. Es kann daher sowohl die Annahme eines, durch bestimmte Eigenschaften ausgezeichneten wahrscheinlichsten Werthes, aus dessen Existenz ein Vertheilungsgesetz folgt, als auch die Voraussetzung von Elementarursachen, deren Combinationen ein System von Wahrscheinlichkeitswerthen liefern und zur Aufstellung eines Vertheilungsgesetzes verhelfen, im günstigsten Falle nur zur Bestimmung einer gewissen Gruppe von C.G. führen.

Dabei ist jedoch wohl zu beachten, dass ein fundamentaler Unterschied zwischen den beiden Arten von Hypothesen besteht. Denn die Hypothese vom arithmetischen Mittel und das gewöhnliche Fehlergesetz, nicht minder die Erweiterung dieser Hypothese in Fechner's Collectivmaßlehre und das erweiterte (zweiseitige) Fehlergesetz stehen in solidarischem Zusammenhang, so dass die Hypothese und die Form des Gesetzes sich wechselweise bedingen. Hingegen führt zwar die Voraussetzung von Elementarursachen, die abhängig oder unabhängig von einander mit bestimmten Wahrscheinlichkeiten wirken, nothwendig zu einem bestimmten Vertheilungsgesetze; es kann aber jedes Vertheilungsgesetz durch unbegrenzt viele, verschiedenartige Systeme von Elementarursachen erzeugt gedacht werden. Die vielfach hervortretende Neigung[1]), aus der Form der Vertheilungstafel oder aus den Werthen gewisser, auf Grund der Vertheilungstafel berechneter Constanten auf die Beschaffenheit der Elementarursachen zu schließen, hat daher keine Berechtigung und führt zu haltlosen Vermuthungen ohne wissenschaftlichen Werth.

Um dies in der einfachsten Weise zu erhärten, möge eine aus vier Varianten A_0, A_1, A_2, A_3 bestehende Reihe mit den zugehörigen Wahrscheinlichkeitswerthen p_0, p_1, p_2, p_3, für die

$$p_0 + p_1 + p_2 + p_3 = 1,$$

gegeben sein; und es sollen die Varianten aus den Combinationen von drei Elementarursachen resultiren, in der Weise, dass dem Wir-

1) Bei Pearson an verschiedenen Stellen seiner Abhandlung (a. a. O. S. 379, 389, 410).

ken keiner Ursache A_0 entspricht, das Wirken einer Ursache zu A_1, das Wirken von zwei und drei Ursachen zu A_2 und A_3 führt.

Nimmt man zunächst an, dass die Elementarursachen unabhängig von einander mit den Wahrscheinlichkeiten p, q, r wirken und mit den Wahrscheinlichkeiten $p' = 1 - p$, $q' = 1 - q$, $r' = 1 - r$ nicht wirken, so werden die Varianten A_3, A_2, A_1, A_0 der Reihe noch in folgenden relativen Häufigkeiten auftreten:

$$\begin{aligned} p_3 &= p \cdot q \cdot r \\ p_2 &= p \cdot q \cdot r' + p \cdot q' \cdot r + p' \cdot q \cdot r \\ p_1 &= p \cdot q' \cdot r' + p' \cdot q \cdot r' + p' \cdot q' \cdot r \\ p_0 &= p' \cdot q' \cdot r' . \end{aligned}$$

Die Werthe p, q, r, die sämmtlich positiv und kleiner als 1 sein müssen, sind daher die Wurzeln der Gleichung

$$x^3 - (p_1 + 2p_2 + 3p_3)x^2 + (p_2 + 3p_3)x - p_3 = 0. \qquad (34)$$

Die Wurzeln dieser Gleichung sind aber jedenfalls dann nicht alle reell und positiv, wenn die Ungleichungen[1])

$$\frac{1}{3}(p_1 + 2p_2 + 3p_3) > \sqrt{\frac{1}{3}(p_2 + 3p_3)} > \sqrt[3]{p_3}$$

nicht erfüllt sind. Setzt man nun z. B.

$$p_0 = p_1 = p_2 = p_3 = \frac{1}{4},$$

so wird

$$\frac{1}{2} < \sqrt{\frac{1}{3}} < \sqrt[3]{\frac{1}{4}}.$$

Es gibt daher in diesem Falle kein System unabhängig wirkender Elementarursachen, durch welche die Varianten mit den angegebenen Wahrscheinlichkeitswerthen erzeugt werden können.

---

1) Im nächsten Capitel (III, § 1) beweise ich, dass, falls eine Gleichung n-ten Grades

$$x^n - n\beta_1 x^{n-1} + \frac{n(n-1)}{1 \cdot 2}\beta_2^2 x^{n-2} - \cdots \pm \beta_n^n = 0$$

lauter positive, reelle Wurzeln hat, die Ungleichungen

$$\beta_1 > \beta_2 > \cdots > \beta_n$$

bestehen.

Man muss somit im allgemeinen Abhängigkeit zwischen den Elementarursachen voraussetzen. Dann wird zunächst die erste Ursache mit der Wahrscheinlichkeit p wirken und mit der Wahrscheinlichkeit $p' = 1 - p$ außer Wirkung bleiben. Es wird ferner die zweite Ursache, je nachdem die erste gewirkt oder nicht gewirkt hat, mit der Wahrscheinlichkeit q resp. r wirksam werden und mit der Wahrscheinlichkeit $q' = 1 - q$ resp. $r' = 1 - r$ nicht wirksam werden. Zuletzt wird die dritte Ursache mit der Wahrscheinlichkeit s oder t oder u oder v eingreifen und mit der Wahrscheinlichkeit $s' = 1 - s$, oder $t' = 1 - t$ oder $u' = 1 - u$ oder $v = 1 - v$ nicht eingreifen, je nachdem die erste Ursache im Verein mit der zweiten oder die erste ohne die zweite oder die zweite ohne die erste oder weder die erste noch die zweite ihre Wirkung ausgeübt hat. Dies wird durch folgendes Schema erläutert:

$$\begin{array}{lcccccccccc}
\text{1. Ursache:} & & p & & ; & & p' & & \\
\text{2. Ursache:} & q & & q' & ; & r & & r' & \\
\text{3. Ursache:} & s \quad s' ; & & t \quad t' ; & & u \quad u' ; & & v \quad v' .
\end{array}$$

Ist hier $q = r$ und $s = t = u = v$, so wirken die Ursachen unabhängig von einander.

Man erhält so auf Grund der gegebenen Wahrscheinlichkeitswerthe p_3, p_2, p_1, p_0 der vier Varianten zur Bestimmung der Elementarursachen die Gleichungen:

$$\begin{aligned}
p_3 &= p \cdot q \cdot s \\
p_2 &= p \cdot q \cdot s' + p \cdot q' \cdot t + p' \cdot r \cdot u \\
p_1 &= p \cdot q' \cdot t' + p' \cdot r \cdot u' + p' \cdot r' \cdot v \\
p_0 &= p' \cdot r' \cdot v' .
\end{aligned}$$

Setzt man nun

$$\begin{aligned}
p_{21} &= pqs' ; \; p_{22} = pq't ; \; p_{23} = p'ru , \\
p_{11} &= pq't' ; \; p_{12} = p'ru' ; \; p_{13} = p'r'v ,
\end{aligned}$$

jedoch so, dass

$$p_{21} + p_{22} + p_{23} = p_2 ; \; p_{11} + p_{12} + p_{13} = p_1 ,$$

so erhält man:

$$\left.\begin{aligned}
&s=\frac{p_3}{p_3+p_{21}};\quad t=\frac{p_{22}}{p_{22}+p_{11}};\quad u=\frac{p_{23}}{p_{23}+p_{12}};\quad v=\frac{p_{13}}{p_{13}+p_0};\\
&q=\frac{p_3+p_{21}}{p_3+p_{21}+p_{11}+p_{22}};\quad r=\frac{p_{12}+p_{23}}{p_{12}+p_{23}+p_0+p_{13}};\\
&p=p_3+p_{21}+p_{22}+p_{11}=1-p_0-p_{13}-p_{12}-p_{23};
\end{aligned}\right\}\quad(35)$$

Da aber im allgemeinen $p_{21}, p_{22}, p_{23}; p_{11}, p_{12}, p_{13}$ auf unbegrenzt viele Arten als positive echte Brüche bestimmt werden können, so dass

$$p_{21}+p_{22}+p_{23}=p_2;\; p_{11}+p_{12}+p_{13}=p_1$$

$$p_3+p_{21}+p_{22}+p_{11}<1;\; p_0+p_{13}+p_{12}+p_{23}<1$$

so gibt es auch dementsprechend unbegrenzt viele verschiedene Systeme von Wahrscheinlichkeitswerthen für das Wirken der vorauszusetzenden drei Elementarursachen.

Man kann daher in der That aus dem Verlaufe der Werthe einer Vertheilungstafel keinen Schluss auf die Art und Weise des Wirkens von Elementarursachen ziehen. Insbesondere schließt die Gültigkeit des gewöhnlichen Fehlergesetzes nicht ein, dass — wie die Hypothese Hagen's annimmt — die Messungsfehler aus Elementarfehlern sich zusammensetzen, die ebenso leicht positiv wie negativ sein können.

c. Eine Methode zur Darstellung willkürlich gegebener Functionen.

Verzichtet man demzufolge auf die Hypothesenbildung, so kann man in der Vertheilungstafel lediglich die Werthe einer willkürlich gegebenen Function erblicken und die Methoden zur Darstellung willkürlich gegebener Functionen anwenden.

Eine solche Methode, welche den Besonderheiten der C.G. Rechnung trägt, will ich hier entwickeln.

Da nur ausnahmsweise, auf Grund eines Actes deductiven Erkennens, eine Orientirung über die extremen Werthe einer Variantenreihe vorausgesetzt werden darf, so ist im allgemeinen eine vollständige Zahlenmannigfaltigkeit ohne willkürlichen Ausschluss bestimmter, mehr oder minder ausgedehnter Gebiete als Träger der Reihe denkmöglicher Varianten anzunehmen. Diese Mannigfaltigkeit sei zunächst, im Einklang mit der früher (§ 3) an erster Stelle angegebenen Form der Vertheilungstafel (14) die Reihe der reellen ganzen Zahlen.

Nun soll jedem Gliede α der von $-\infty$ bis $+\infty$ sich erstreckenden Reihe ganzer Zahlen ein bestimmter reeller Werth $F(\alpha)$ zugeordnet gedacht werden, der positiv oder gleich Null sein soll, da die Vertheilungstafel eines C.G. keine negativen Werthe darbieten kann. Dann ist $F(\alpha)$ eine willkürlich gegebene Function des ganzzahligen Argumentes α. Diese Function soll aber, da sie ins Unbegrenzte sich erstreckt, noch folgenden Bedingungen genügen, die ohne weiteres erfüllt sind, wenn $F(\alpha)$ durchaus im Endlichen verläuft, d. h. unterhalb und oberhalb einer gewissen Grenze durchweg gleich Null ist.

Es sei erstens die über alle ganzen Zahlen α von $-\infty$ bis $+\infty$ erstreckte Summe

$$\sum (\alpha - b)^{\nu} \cdot F(\alpha) = s_{\nu}, \tag{36}$$

wo b einen beliebigen reellen Werth bezeichnet, endlich für jede endliche ganze Zahl $\nu = 0, 1, 2 \ldots.$

Bezeichnet man die Summe, welche über ein aus ϱ Gliedern bestehendes Intervall J der Zahlenreihe erstreckt wird, durch Σ_{ϱ} und setzt man

$$\sum_{\varrho} (\alpha - b)^{\nu} \cdot F(\alpha) + d_{\nu} = s_{\nu}, \tag{37}$$

so soll zweitens das Intervall J stets der Art bestimmbar sein, dass die absoluten Beträge der Größen d_{ν} den Bedingungen

$$|d_{\nu}| \leqq k_{\nu} \text{ für } \nu = 0, 1, 2 \ldots n-1 \tag{38}$$

genügen, wo $k_0, k_1, k_2 \ldots k_{n-1}$ Größen sind, die bei dem für die Berechnung von $s_0, s_1, s_2 \ldots s_{n-1}$ angestrebten Grade von Genauigkeit nicht in Betracht kommen.

Die Erstreckung des durch die Bedingung (38) geforderten Intervalls J ist sowohl von der Größe der k als auch von der Anzahl n der Summen $s_0, s_1 \ldots s_{n-1}$, zu deren Berechnung J genügen soll, abhängig. Insbesondere wird für kleine Anzahlen n die Gliederanzahl ϱ von J stets größer als n sein. Es soll aber drittens von einem bestimmten Werthe n ab, den man sich beispielsweise gleich 10 oder gleich 100 oder gleich 1000 denken kann, das zu $k_0, k_1 \ldots$

k_{n-1} gehörige Intervall J eine Gliederanzahl $\varrho \leqq n$ besitzen. Man kann dann offenbar stets $\varrho = n$ voraussetzen.

Es gibt somit bei Erfüllung dieser Bedingungen ein bestimmtes, kleinstes n-gliedriges Intervall J, so dass unter Vernachlässigung der den Bedingungen (38) genügenden Größen d_ν auf Grund der n Gleichungen

$$\sum_n (\alpha - b)^\nu \cdot F(\alpha) = s_\nu \tag{39}$$

$$\nu = 0, 1, 2 \ldots n-1$$

einerseits die Summenwerthe $s_0, s_1, \ldots s_{n-1}$ aus den n Functionswerthen $F(\alpha)$ des Intervalls J, andererntheils die n Functionswerthe $F(\alpha)$ des Intervalls aus den gegebenen Werthen $s_0, s_1 \ldots s_{n-1}$ berechnet werden können. Und da die Werthe $F(\alpha)$ außerhalb des Intervalls J bei dem angestrebten Genauigkeitsgrade außer Acht gelassen und gleich Null gesetzt werden dürfen, so wird demnach der Verlauf der willkürlich gegebenen Function durch die n Summenwerthe mit einer durch die angegebenen Vernachlässigungen bedingten Annäherung bestimmt. Je größer die erstrebte Genauigkeit ist, um so größer muss im allgemeinen das Intervall J und zugleich die Reihe der Summenwerthe sein. Man darf aber erwarten, dass die Function $F(\alpha)$ durch die unbegrenzte Reihe der Summenwerthe $s_0, s_1, s_2 \ldots$ vollständig bestimmt wird.

Diese Eigenschaft der Function $F(\alpha)$ kann man auch ohne weiteres voraussetzen und sagen: Es soll angenommen werden, dass die willkürlich gegebene Function $F(\alpha)$ durch die Summenwerthe s_ν der nullten bis $(n-1)$ten Ordnung mit einer gewissen Annäherung und durch die unbegrenzte Reihe der Summenwerthe vollständig bestimmt werde.

Um nun zu einer Darstellung von $F(\alpha)$ zu gelangen, soll eine reelle, eindeutige, im Gebiete der positiven Ordinaten verlaufende Function $\varphi(\alpha)$ des ganzzahligen Argumentes α als bekannt vorausgesetzt werden, so dass die von $\alpha = -\infty$ bis $\alpha = +\infty$ erstreckte Summe

$$\sum (\alpha - b)^\nu \cdot \varphi(\alpha) = \sigma_\nu \tag{40}$$

für jede endliche ganze Zahl ν endlich ist und mit einem beliebig vorgegebenen Genauigkeitsgrade durch die auf ein bestimmtes, kleines

Intervall beschränkte Summe repräsentirt wird. Dies ist möglich, wenn $\varphi(\alpha)$ einen Parameter enthält, über den in geeigneter Weise verfügt werden kann. Ueberdies möge der Einfachheit wegen $\sigma_0 = 1$ sein.

Denn ist auch für eine beliebig gewählte ganze Zahl β

$$\sum (\alpha + \beta - b)^\nu \cdot \varphi(\alpha + \beta) = \sigma_\nu ,$$

und da

$$(\alpha - b)^\nu = (\alpha + \beta - b)^\nu - \binom{\nu}{1}(\alpha + \beta - b)^{\nu-1} \cdot \beta$$
$$+ \binom{\nu}{2}(\alpha + \beta - b)^{\nu-2} \cdot \beta^2 \cdots\cdots \pm \beta^\nu ,$$

so ist zugleich

$$\sum (\alpha - b)^\nu \cdot \varphi(\alpha + \beta) = \sigma_\nu - \binom{\nu}{1}\sigma_{\nu-1}\beta + \binom{\nu}{2}\sigma_{\nu-2}\beta^2 - \cdots \pm \beta^\nu .$$

Es ist daher, wenn $\beta_1, \beta_2 \ldots \beta_n$ irgend welche von einander verschiedene ganze Zahlen bedeuten und als darstellende Function

$$\Phi(\alpha) = c_1\varphi(\alpha + \beta_1) + c_2\varphi(\alpha + \beta_2) + \cdots + c_n\varphi(\alpha + \beta_n) \tag{41}$$

mit den noch unbestimmten Coefficienten $c_1, c_2 \ldots c_n$ vorausgesetzt wird:

$$\sum \Phi(\alpha) = c_1 + c_2 + \cdots + c_n$$
$$\sum (\alpha - b) \cdot \Phi(\alpha) = \sigma_1(c_1 + c_2 + \cdots + c_n) - (c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n)$$

. . . .

und allgemein

$$\sum (\alpha - b)^\nu \cdot \Phi(\alpha) = \sigma_\nu[c] - \binom{\nu}{1}\sigma_{\nu-1}[c\beta]$$
$$+ \binom{\nu}{2}\sigma_{\nu-2}[c\beta^2] - \cdots \pm [c\beta^\nu], \tag{42}$$

wo $[c]$, $[c\beta]$, $[c\beta^2] \ldots$ zur Abkürzung für die Summen $c_1 + c_2 + \cdots + c_n$, $c_1\beta_1 + c_2\beta_2 + \cdots + c_n\beta_n$, $c_1\beta_1^2 + c_2\beta_2^2 + \cdots + c_n\beta_n^2 \ldots$ dient.

Setzt man jetzt in (42) für $\nu = 0, 1, 2 \ldots n - 1$

$$\sum (\alpha - b)^\nu \cdot \Phi(\alpha) = s_\nu ,$$

so kann man aus den n Gleichungen die Coefficienten $c_1, c_2 \ldots c_n$

berechnen, da die Determinante des Gleichungensystems wegen der vorausgesetzten Verschiedenheit der β nicht verschwindet; und die Function $\Phi(\alpha)$ erhält somit die nämlichen Summenwerthe $s_0, s_1 \ldots s_{n-1}$ wie die willkürlich gegebene Function $F(\alpha)$. Da aber σ_ν mit einem beliebig vorgegebenen Genauigkeitsgrade durch die auf ein bestimmtes, kleines Intervall beschränkte Summe (40) repräsentirt wird, so existirt auch für die Function $\Phi(\alpha)$ ein endliches ϱ-gliedriges Intervall J', so dass für $\nu = 0, 1, 2 \ldots n-1$

$$\left.\begin{array}{c} \sum_{\varrho}(\alpha - b)^\nu \cdot \Phi(\alpha) + \delta_\nu = \sum(\alpha - b)^\nu \cdot \Phi(\alpha) \\ |\delta_\nu| \leqq k_\nu . \end{array}\right\} \qquad (43)$$

Die Lage und Ausdehnung von J' wird durch die Werthe $\beta_1, \beta_2 \ldots \beta_n$ und den Parameter von $\varphi(\alpha)$ bedingt.

Lässt es sich nun durch passende Wahl der Werthe β_1, $\beta_2 \ldots \beta_n$ und des Parameters von $\varphi(\alpha)$ erreichen, dass das zur Bestimmung von $\Phi(\alpha)$ genügende Intervall J' mit dem zur Bestimmung von $F(\alpha)$ hinreichenden Intervalle J sich deckt, so wird auf Grund der Uebereinstimmung in den Summenwerthen nullter bis $(n-1)$ter Ordnung $F(\alpha)$ durch $\Phi(\alpha)$ angenähert dargestellt.

Die einzelnen Glieder der Function $\Phi(\alpha)$ lassen sich umformen, indem man für einen positiven Werth β

$$\varphi(\alpha + \beta) = \varphi(\alpha) + \binom{\beta}{1}\varphi'(\alpha) + \binom{\beta}{2}\varphi''(\alpha) + \cdots + \varphi^{(\beta)}(\alpha)$$

und für einen negativen Werth $-\beta$

$$\varphi(\alpha - \beta) = \varphi(\alpha) - \binom{\beta}{1}\varphi_1(\alpha) + \binom{\beta}{2}\varphi_2(\alpha) - \cdots \pm \varphi_\beta(\alpha)$$

setzt, wo

$$\varphi'(\alpha) = \varphi(\alpha + 1) - \varphi(\alpha)\,;\; \varphi''(\alpha) = \varphi'(\alpha + 1) - \varphi'(\alpha)\,;\; \ldots$$
$$\varphi_1(\alpha) = \varphi(\alpha) - \varphi(\alpha - 1)\,;\; \varphi_2(\alpha) = \varphi_1(\alpha) - \varphi_1(\alpha - 1)\,;\; \ldots$$

Man erhält so die darstellende Function (41) in der Form

$$\left.\begin{array}{l} \Phi(\alpha) = \gamma_0\varphi(\alpha) + \gamma'\varphi'(\alpha) + \gamma''\varphi''(\alpha) + \ldots \\ \qquad + \gamma_1\varphi_1(\alpha) + \gamma_2\varphi_2(\alpha) + \ldots \end{array}\right\} \qquad (44)$$

Ist insbesondere $\varphi(\alpha)$ eine Function, die für negative ganze Zahlen α durchweg gleich Null ist, und wählt man zur Herstellung von (41) nur negative ganze Zahlen $\beta_1, \beta_2 \ldots \beta_n$, so wird auch $\Phi(\alpha)$

für negative Argumentwerthe gleich Null und in (44) treten bloß die abgeleiteten Functionen $\varphi_1, \varphi_2 \ldots$ auf. Auf diese Weise kann man zur Darstellung einer den positiven ganzen Zahlen $0, 1, 2 \ldots$ zugeordneten, willkürlich gegebenen Werthenreihe gelangen.

Eine hiezu geeignete Function ist

$$\varphi(\alpha) = \frac{\lambda^\alpha}{\Pi(\alpha)}, \tag{45}$$

wo λ einen Parameter und $\Pi(\alpha)$ die bekannte, der Relation $\Pi(\alpha) = \alpha \cdot \Pi(\alpha - 1)$ genügende, Gauß'sche Function für ganzzahlige Argumente bezeichnet, so dass, wenn α positiv ist,

$$\frac{1}{\Pi(\alpha)} = \frac{1}{1 \cdot 2 \cdots \alpha}; \quad \frac{1}{\Pi(0)} = 1; \quad \frac{1}{\Pi(-\alpha)} = 0.$$

Werden die Varianten des C.G. nicht — wie zunächst angenommen wurde — der Reihe der ganzen Zahlen, sondern der gesammten reellen Zahlenmannigfaltigkeit zugeordnet, so ist auch für die willkürlich gegebene Function ein reelles, stetig variables Argument a vorauszusetzen. In diesem Falle hat man in Uebereinstimmung mit der früher (§ 3) an zweiter Stelle (14a) angegebenen Form der Vertheilungstafel von einer Eintheilung der reellen Zahlenmannigfaltigkeit in aneinander grenzende Intervalle auszugehen und jedem Intervall einen reellen endlichen Zahlenwerth (der positiv oder gleich Null sein kann) zugewiesen zu denken. Diese Werthe sind sodann auf die zugehörigen Intervalle vertheilt zu denken, so dass schließlich jedem von a und $a + da$ begrenzten, unendlich kleinen Bereiche ein bestimmter unendlich kleiner Werth $f(a) \cdot da$ zugehört. Nunmehr ist $f(a)$ eine willkürlich gegebene Function des reellen Arguments a, für welche die anfängliche Eintheilung der Zahlenmannigfaltigkeit in Intervalle nicht weiter in Betracht kommen soll.

Da diese Function sich ins Unbegrenzte erstreckt, so soll sie im Interesse ihrer Darstellbarkeit ähnlich wie $F(\alpha)$ folgenden, von jeder im Endlichen verlaufenden Function ohne weiteres erfüllten Bedingungen genügen. Es besitze das von $a = -\infty$ bis $a = +\infty$ erstreckte Integral

$$\int (a - b)^\nu \cdot f(a) \cdot da = s_\nu, \tag{46}$$

wo b eine beliebige reelle Größe angibt, für endliche Zahlen $\nu = 0$, 1, 2 . . . einen endlichen Werth. — Setzt man ferner

$$\int_J (a-b)^\nu \cdot f(a) \cdot da + d_\nu = s_\nu , \tag{47}$$

wo die Beifügung von J an das Integralzeichen die Beschränkung der Integration auf das Intervall J andeutet, so soll J so bestimmbar sein, dass für beliebig vorgegebene, bei der Berechnung von s_ν nicht in Betracht gezogene Größen k_ν der absolute Betrag von d_ν die Bedingung

$$|d_\nu| \leqq k_\nu ; \quad \text{für } \nu = 0,\ 1,\ 2 \ldots n-1 \tag{47a}$$

erfüllt. — Es soll schließlich das zur Berechnung von $s_0, s_1 \ldots s_{n-1}$ ausreichende Intervall J von einem gewissen Werthe n ab (der etwa gleich 10 oder 100 oder 1000 sein möge) klein genug und zugleich der Verlauf von $f(a)$ ruhig genug sein, so dass nach Zerlegung von J in n aneinandergrenzende Theilintervalle $a_1 \pm \frac{1}{2} i_1$, $a_2 \pm \frac{1}{2} i_2 \ldots a_n \pm \frac{1}{2} i_n$ (mit den Längen $i_1, i_2 \ldots i_n$ und den Mitten $a_1, a_2 \ldots a_n$) die Angabe der auf die Theilintervalle fallenden und durch die von $a_k - \frac{1}{2} i_k$ bis $a_k + \frac{1}{2} i_k$ erstreckten Integrale

$$\int_{(k)} f(a) \cdot da = s_0^{(k)} \tag{48}$$

dargestellten Beträge $s_0^{(1)}, s_0^{(2)} \ldots s_0^{(n)}$ genügt, um den Functionsverlauf angenähert zu bestimmen.

Sind diese Bedingungen erfüllt, so kann für $\nu = 0, 1, 2 \ldots n-1$

$$\left.\begin{aligned} s_\nu &= s_\nu^{(1)} + s_\nu^{(2)} + \cdots + s_\nu^{(n)} \\ s_\nu^{(k)} &= \int_{(k)} (a-b)^\nu \cdot f(a) \cdot da \end{aligned}\right\} \tag{49}$$

gesetzt werden. Da es aber für jedes ν (mag b oberhalb oder unterhalb oder innerhalb des Intervalls $a_k + \frac{1}{2} i_k$ liegen) zwei den Intervallgrenzen oder dem Intervalle selbst angehörige Werthe a'_k und a''_k gibt, so dass

$$(a_k' - b)^\nu \cdot \int_{(k)} f(a) \cdot da < s_\nu^{(k)} < (a_k'' - b)^\nu \cdot \int_{(k)} f(a) \cdot da\,,$$

so existirt stets ein dem Intervalle $a_k \pm \frac{1}{2} i_k$, nicht aber seinen Grenzen angehöriger Werth $a_k^{(\nu)}$, so dass

$$s_\nu^{(k)} = (a_k^{(\nu)} - b)^\nu \cdot s_0^{(k)}. \tag{50}$$

Demnach bestehen die n Gleichungen

$$\left.\begin{aligned} s_0 &= s_0^{(1)} + s_0^{(2)} + \cdots + s_0^{(n)} \\ s_1 &= (a_1^{(1)} - b) s_0^{(1)} + (a_2^{(1)} - b) s_0^{(2)} + \cdots + (a_n^{(1)} - b) s_0^{(n)} \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ s_{n-1} &= (a_1^{(n-1)} - b)^{n-1} \cdot s_0^{(1)} + (a_2^{(n-1)} - b)^{n-1} \cdot s_0^{(2)} + \cdots \\ &\qquad + (a_n^{(n-1)} - b)^{n-1} \cdot s_0^{(n)}, \end{aligned}\right\} \tag{51}$$

aus denen, wenn die Werthe $a_k^{(\nu)}$ bekannt sind, oder willkürlich angenommen und etwa durch a_k ersetzt werden dürfen, die Beträge $s_0^{(1)}$, $s_0^{(2)} \ldots s_0^{(n)}$ auf Grund der Integralwerthe s_0, s_1, $s_2 \ldots s_{n-1}$ bestimmt werden können.

Diese Integralwerthe gestatten somit eine angenäherte Bestimmung der willkürlich gegebenen Function $f(a)$. Ihre Anzahl n muss offenbar um so größer sein, je schärfer die Bestimmung sein soll. Es ist aber anzunehmen, dass durch die unbegrenzte Reihe von Integralwerthen die Function $f(a)$ vollständig bestimmt wird. Die Festsetzung der obigen Bedingungen ist demnach gleichbedeutend mit der Voraussetzung, dass die willkürlich gegebene Function $f(a)$ durch die Integralwerthe s_ν der nullten bis $(n-1)$ten Ordnung mit einer gewissen Annäherung und durch die unbegrenzte Reihe der Integralwerthe völlig bestimmt werde.

Um diese Function $f(a)$ in ähnlicher Weise wie $F(\alpha)$ angenähert darzustellen, möge eine eindeutige, durchweg positive Function $\varphi(a)$ zur Verfügung stehen, so dass das von $a = -\infty$ bis $a = +\infty$ erstreckte Integral

$$\int (a - b)^\nu \cdot \varphi(a) \cdot da = \sigma_\nu \tag{52}$$

für jeden endlichen ganzzahligen Werth ν endlich ist und durch geeignete Bestimmung eines der Function $\varphi(a)$ angehörenden Parameters mit jedem beliebigen Grade von Genauigkeit durch den über ein bestimmtes kleines Intervall genommenen Integralwerth ersetzt werden kann. Insbesondere sei $\sigma_0 = 1$. Bildet man nun mittelst der geeignet gewählten, reellen Werthe $b_1, b_2 \ldots b_n$ die darstellende Function

$$\Phi(a) = c_1 \varphi(a + b_1) + c_2 \varphi(a + b_2) + \cdots + c_n \varphi(a + b_n) \tag{53}$$

und setzt man für $\nu = 0, 1, 2 \ldots n - 1$

$$\int (a - b)^\nu \cdot \Phi(a) \cdot da = s_\nu ,$$

so können aus den mit Rücksicht auf die Relationen

$$\int (a + b_k - b)^\nu \cdot \varphi(a + b_k) \cdot da = \sigma_\nu$$

$$\int (a - b)^\nu \cdot \varphi(a + b_k) = \sigma_\nu - \binom{\nu}{1} \sigma_{\nu-1} \cdot b_k + \cdots \pm b_k^\nu$$

gebildeten Gleichungen:

$$s_\nu = \sigma_\nu [c] - \binom{\nu}{1} \sigma_{\nu-1} [cb] + \binom{\nu}{2} \sigma_{\nu-2} [cb^2] - \cdots \pm [cb^\nu] \tag{54}$$

$$\nu = 0, 1, 2 \ldots n - 1$$

wo $[c] = c_1 + c_2 + \cdots + c_n$; $[cb] = c_1 b_1 + c_2 b_2 + \cdots + c_n b_n$; u. s. w., die Coefficienten $c_1, c_2 \ldots c_n$ berechnet werden. Die Function $\Phi(a)$ erhält somit die nämlichen Integralwerthe $s_0, s_1 \ldots s_{n-1}$ wie die willkürlich gegebene Function $f(a)$. Zugleich existirt ein in seiner Lage und Größe durch $b_1, b_2 \ldots b_n$ und den Parameter von $\Phi(a)$ bestimmtes Intervall J', so dass für $\nu = 0, 1, \ldots n - 1$

$$\left.\begin{aligned} \int_{J'} (a - b)^\nu \cdot \Phi(a) \cdot da + \delta_\nu &= \int (a - b)^\nu \cdot F(a) \\ |\delta_\nu| &\leqq k_\nu \end{aligned}\right\} \tag{55}$$

Es wird daher auf Grund der Uebereinstimmung in den Integralwerthen nullter bis $(n-1)$ter Ordnung $f(a)$ durch $\Phi(a)$ mit einer gewissen Annäherung dargestellt, wenn sich der Parameter von $\varphi(a)$ und die Werthe $b_1, b_2 \ldots b_n$ so bestimmen lassen, dass das zur Bestimmung der Integralwerthe von $\Phi(a)$ ausreichende Intervall J' mit dem zur Bestimmung der Integralwerthe von $F(a)$ genügenden Intervall J sich deckt.

Setzt man

$$\varphi(a+b_k)=\varphi(a)+\frac{b_k}{1}\varphi'(a)+\frac{b_k^2}{1\cdot 2}\varphi''(a)+\cdots,$$

so erhält man die darstellende Function (53) in der Form

$$\Phi(a)=\gamma_0\varphi(a)+\gamma_1\varphi'(a)+\gamma_2\varphi''(a)+\cdots \tag{56}$$

wo $\gamma_0=[c]\,;\quad \gamma_1=[cb]\,;\quad \gamma_2=\frac{1}{1\cdot 2}[cb^2]\,;\ \ldots$

Eine zur Anwendung dieser Methode geeignete Hülfsfunction ist die das gewöhnliche Fehlergesetz darstellende

$$\varphi(a)=\frac{h}{\sqrt{\pi}}\,exp(-h^2a^2)\,. \tag{57}$$

Die mittels dieser Function gebildete Reihe (56) ist zuerst von Bruns[1]) auf einem anderen Wege entwickelt worden.

Um eine exacte Darstellung der willkürlich gegebenen Function $F(\alpha)$ oder $f(a)$ zu erhalten, muss — von Ausnahmefällen abgesehen — in (41) oder (53) eine unbegrenzte Reihe von Gliedern vorausgesetzt und demzufolge eine unbegrenzte Anzahl von Coefficienten $c_1, c_2, c_3 \ldots$ bestimmt werden. Die so resultirende darstellende Function $\Phi(\alpha)$ oder $\Phi(a)$ ist aber im allgemeinen nur bedingt convergent. Denn die Werthe von $c_1, c_2, c_3 \ldots$ können theils positiv, theils negativ sein, so dass — wenn auch ihre Summe endlich, nämlich gleich s_0, sein muss — die positiven und die negativen Werthe, für sich allein zusammengefasst, einen unendlich großen Betrag ergeben können und nur auf Grund einer bestimmten Aufeinanderfolge zu dem endlichen Betrage s_0 führen. Dann lässt sich aber — allgemein geredet — auch von der Function $\Phi(\alpha)$ oder $\Phi(a)$ nur

---

1) »Ueber die Darstellung von Fehlergesetzen« (Astronomische Nachrichten, CXLIII, No. 3429), ferner »Zur Collectivmaßlehre« (Philosophische Studien, XIV). Die letztere Abhandlung gibt Tabellen für φ und die fünf ersten Ableitungen nebst Anweisungen zu ihrer Benutzung. — Berechnungen von C.G. auf Grund dieser Tabellen hat Werner (»Beiträge zur Collectivmaßlehre«, Philosophische Studien, XV) ausgeführt. — Auf dem oben angegebenen Wege habe ich die mittelst (57) gebildete Reihe (56) in den Philosophischen Studien, XIII (»Ueber Fechner's Collectivmaßlehre und die Vertheilungsgesetze der Collectivgegenstände«) hergeleitet.

bedingte, d. h. auf der Anordnung ihrer Glieder beruhende Convergenz behaupten.

d. Die mit der Functionsdarstellung verbundene Ueberbestimmung der C.G.

Die Darstellung einer Vertheilungstafel durch eine mathematische Function auf dem soeben angegebenen oder einem anderen Wege bietet ein werthvolles Hülfsmittel bei der Feststellung möglicher Gesetzmäßigkeiten. Soll aber die Function zur Charakterisirung des C.G. dienen, so ist zu beachten, dass sie ihrem Wesen nach eine zu weit gehende Bestimmung liefert und somit zu illusorischen, durch keine Erfahrung gestützten Ergebnissen führt.

In dieser Hinsicht ist zunächst zu beachten, dass jede als Darstellung einer Vertheilungstafel gefundene Function die Ableitung einer unbegrenzten Reihe von Summen- oder Integralwerthen s_ν gestattet, welche — wenn $s_\nu = s_0 \cdot \varepsilon_\nu^\nu$ gesetzt wird — die Kenntniss einer unbegrenzten Reihe von Mittelwerthen ε_1, ε_2, ε_3 ... in inf. vermitteln. Es kann aber jedenfalls nur eine begrenzte Anzahl von Werthen s_ν oder ε_ν der empirisch gegebenen Vertheilungstafel zur Bestimmung der Function entnommen werden. Somit wird aus einer beschränkten Anzahl empirisch gefundener Werthe eine unbeschränkte Anzahl von Werthen abgeleitet; und diese Erweiterung der Erkenntniss beruht nicht ausschließlich auf der Erfahrung, sondern wesentlich auf der Form der gerade gewählten Function. Demgemäß wird durch die Wahl der die Functionsdarstellung vermittelnden Hülfsfunction ein willkürlicher und fremdartiger Bestandtheil in die Charakterisirung der C.G. eingeführt.

Sodann ist in Rücksicht zu ziehen, dass die den Häufigkeitswerthen der Vertheilungstafel im allgemeinen nothwendig anhaftende Unsicherheit eine entsprechende Unsicherheit für die auf Grund jener Häufigkeitswerthe bestimmte Function im Gefolge hat. Der einem Argumentwerthe zugehörige Functionswerth kann daher nur hypothetische Geltung beanspruchen, als möglicher Repräsentant eines mehr oder minder ausgedehnten Bereiches von Werthen. Eine in aller Strenge gültige Functionsdarstellung kann darum gar nicht erstrebt werden. Insbesondere kann man nicht hoffen, durch die Function eine Ausgleichung der unsicher schwankenden, empirisch gefundenen Wahrscheinlichkeitswerthe zu erhalten. Vielmehr kann nur auf Grund dieser empirischen Werthe festgestellt werden, wie weit der Functionsdarstellung eine hypothetische Geltung zuerkannt werden kann.

Solche willkürliche und hypothetische Elemente vermeidet die in § 3 entwickelte Bestimmungsweise der C.G.

§ 5. Die Methode der Mittelwerthe und das Gauß'sche Princip des mittleren Fehlers.

Eine wesentliche Stütze findet die Methode, die C.G. durch eine — je nach Bedürfniss mehr oder minder weit auszudehnende — Reihe von Mittelwerthen zu charakterisiren, durch die »Theoria combinationis observationum erroribus minimis obnoxiae« von Gauß. In derselben wird die Fehlertheorie unabhängig von einem hypothetischen Fehlergesetz auf das Princip des mittleren Fehlers gestellt. Dies wird von Gauß selbst durch folgende Worte (Art. 17) klar hervorgehoben:

»In Theoria motus corporum coelestium ostendimus, quomodo valores incognitarum maxime probabiles eruendi sint, si lex probabilitatis errorum observationum cognita sit; et quum haec lex natura sua in omnibus fere casibus hypothetica maneat, theoriam illam ad legem maxime plausibilem applicavimus, ubi probabilitas erroris x quantitati exponentiali e^{-hhxx} proportionalis supponitur, unde methodus a nobis dudum in calculis praesertim astronomicis, et nunc quidem a plerisque calculatoribus sub nomine methodi quadratorum minimorum usitata demanavit.

»Postea ill. Laplace, rem alio modo aggressus, idem principium omnibus aliis etiamnum praeferendum esse docuit, quaecunque fuerit lex probabilitatis errorum, si modo observationum multitudo sit permagna. At pro multitudine observationum modica res intacta mansit, ita ut si lex nostra hypothetica respuatur, methodus quadratorum minimorum eo tantum nomine prae aliis commendabilis habenda sit, quod calculorum concinnitati maxime est adaptata.

»Geometris itaque gratum fore speramus, si in hac nova argumenti tractatione docuerimus, methodum quadratorum minimorum exhibere combinationem ex omnibus optimam, non quidem proxime, sed absolute, quaecunque fuerit lex probabilitatis errorum, quaecunque observationum multitudo, si modo notionem erroris medii non ad mentem ill. Laplace, sed ita ut in artt. 6 et 7 a nobis factum est stabiliamus«.

An jener Stelle definirt Gauß den mittleren Fehler m durch das von $x = -\infty$ bis $x = +\infty$ erstreckte Integral

$$m^2 = \int x^2 \cdot \varphi(x) \cdot dx \;,$$

wo x die Größe eines Fehlers und $\varphi(x) \cdot dx$ die Wahrscheinlichkeit eines zwischen x und $x + dx$ liegenden Fehlers angibt, und sagt: ». . . integrale

$$\int xx\varphi(x)\,dx$$

ab $x = -\infty$ usque ad $x = +\infty$ extensum (seu valor medius quadrati xx) aptissimum videtur ad incertitudinem observationum in genere definiendam et dimetiendam, ita ut e duobus observationum systematibus, quae quoad errorum facilitatem inter se differunt, eae praecisione praestare censeantur, in quibus integrale $\int xx\varphi(x)dx$ valorem minorem obtinet. Quodsi quis hanc rationem pro arbitrio, nulla cogente necessitate, electam esse objiciat, lubenter assentiemur. Quippe quaestio haec per rei naturam aliquid vagi implicat, quod limitibus circumscribi nisi per principium aliquatenus arbitrarium nequit [1]).«

Da nun vorausgesetzt wird (Art. 17), dass die Fehler von einem etwa vorhandenen constanten Betrag befreit worden seien und demgemäß

$$\int x\varphi(x) \cdot dx = 0$$

sei, so sind die Fehler x als Abweichungen der einzelnen Beobachtungen von dem arithmetischen Mittel aller Beobachtungen aufzu-

1) Die ausführliche Mittheilung des Grundgedankens der Gauß'schen Theorie erschien mir nicht überflüssig, da das Princip des mittleren Fehlers z. B. bei der ausführlichen Behandlung der Fehlertheorie in Meyer's »Vorlesungen über Wahrscheinlichkeitsrechnung« (1879) Cap. VIII, S. 243—328, Anhang S. 441—509 keinen Ausdruck findet, hingegen die Ableitung des Fehlergesetzes aus der Hypothese vom arithmetischen Mittel als Methode von Gauß bezeichnet wird.

fassen. Die Fehlerreihe wird somit nach Gauß unabhängig von einem Fehlergesetz und von der Anzahl der Beobachtungen durch das arithmetische Mittel und den mittleren Fehler oder — mit andern Worten — durch den das arithmetische Mittel definirenden Mittelwerth erster Ordnung $\varepsilon_1 = 0$ und durch den auf das arithmetische Mittel als Ausgangswerth bezogenen Mittelwerth zweiter Ordnung ε_2, der mit dem Gauß'schen mittleren Fehler identisch ist, charakterisirt. Es ist dies aber nichts anderes als die Anwendung der in § 3 entwickelten Methode der Mittelwerthe auf die durch Fehlerreihen dargestellte besondere Classe von C.G. unter Beschränkung auf die Mittelwerthe erster und zweiter Ordnung.

Demnach kann die Methode der Mittelwerthe als eine Verallgemeinerung der Gauß'schen Methode, eine Reihe von Beobachtungen durch das arithmetische Mittel und den mittleren Fehler zu charakterisiren, aufgefasst werden. Zugleich erscheint aber das Princip des mittleren Fehlers in einem neuen Lichte. Es ist keineswegs mit der Willkürlichkeit behaftet, die Gauß selbst in den angeführten Worten zuzugeben bereit ist. Vielmehr wird der mittlere Fehler als Mittelwerth zweiter Ordnung zu einem niemals entbehrlichen Gliede in einer Kette von Bestimmungsstücken, deren Ausdehnung von den jeweiligen Bedürfnissen abhängt und deren Beschränkung auf die beiden ersten Glieder zwar bei den Fehlerreihen vielfach genügen mag, aber nicht etwa in der Natur der Fehlerreihen begründet ist.

§ 6. Collectivgegenstände mit Doppelreihen von Varianten.

Die Methode, einen aus nur einer Variantenreihe bestehenden C.G. durch die Mittelwerthe $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_\nu$ zu bestimmen, bietet den Vortheil, ohne weiteres auf C.G. mit mehrfach zusammenhängenden Variantenreihen übertragbar zu sein. Es genügt, dies für eine Doppelreihe von Varianten nachzuweisen.

Werden die Varianten, wie in § 1 angegeben wurde, durch Zahlenpaare (α, β) markirt, so lassen sich stets zwei endliche Reihen aufeinanderfolgender Zahlen $\alpha_1, \alpha_2 \ldots \alpha_r$ und $\beta_1, \beta_2 \ldots \beta_s$ abgrenzen, deren $n = r \cdot s$ Combinationen $(\alpha_\varkappa, \beta_\lambda)$ für $\varkappa = 1, 2 \ldots r$; $\lambda = 1, 2 \ldots s$ den Bereich der Zahlenpaare darstellen, dem die beobachteten Varianten angehören. Man erhält dann, wenn $z_{\varkappa\lambda}$ die absolute Häufigkeit der Variante $(\alpha_\varkappa, \beta_\lambda)$ angibt, die Vertheilungstafel in der Form

$$\begin{array}{c|cccc}
 & \alpha_1 & \alpha_2 & \ldots & \alpha_r \\
\hline
\beta_1 & z_{11} & z_{21} & \ldots & z_{r1} \\
\beta_2 & z_{12} & z_{22} & \ldots & z_{r2} \\
\cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot \\
\beta_s & z_{1s} & z_{2s} & \ldots & z_{rs}
\end{array} \tag{58}$$

in der an Stelle der z die relativen Häufigkeiten

$$p_{\varkappa\lambda} = \frac{z_{\varkappa\lambda}}{m}, \text{ wo } m = \sum z_{\varkappa\lambda}$$

treten können.

Will man indessen den Varianten die stetige Mannigfaltigkeit reeller Zahlenpaare (a, b) zuordnen, so theile man diese Mannigfaltigkeit in aneinandergrenzende Bereiche in der Weise, dass der Ordnungszahl $\alpha_\varkappa$ die von $a_\varkappa - \frac{1}{2} i_\varkappa$ bis $a_\varkappa + \frac{1}{2} i_\varkappa$ sich erstreckenden Zahlen, der Ordnungszahl β_λ die zwischen $b_\lambda - \frac{1}{2} j_\lambda$ und $b_\lambda + \frac{1}{2} j_\lambda$ liegenden Zahlen und somit der Combination $(\alpha_\varkappa, \beta_\lambda)$ der Bereich $\left(a_\varkappa \pm \frac{1}{2} i_\varkappa, b_\lambda \pm \frac{1}{2} j_\lambda\right)$ mit der Mitte $(a_\varkappa, b_\lambda)$ zugehört. Alsdann tritt an Stelle von (58) die Form

$$\begin{array}{c|cccc}
 & a_1 & a_2 & \ldots & a_r \\
\hline
b_1 & z_{11} & z_{21} & \ldots & z_{r1} \\
b_2 & z_{12} & z_{22} & \ldots & z_{r2} \\
\cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot \\
\cdot & \cdot & \cdot & & \cdot \\
b_s & z_{1s} & z_{2s} & \ldots & z_{rs}
\end{array} \tag{58a}$$

Bezieht man jetzt $(a_\varkappa, b_\lambda)$ auf ein beliebig, aber fest gewähltes Werthenpaar (c, d), um aus den ϱ-ten und σ-ten Potenzen der Abweichungswerthe $a_\varkappa - c$ und $b_\lambda - d$ die, über $\varkappa = 1, 2 \ldots r$ und $\lambda = 1, 2 \ldots s$ erstreckten Summen

$$m \cdot \varepsilon_{\varrho\sigma}^{\varrho+\sigma} = \sum z_{\varkappa\lambda}(a_\varkappa - c)^\varrho (b_\lambda - d)^\sigma$$

zu bilden, so gelangt man nach Division mit m zu

$$\varepsilon_{\varrho\sigma}^{\varrho+\sigma} = \sum p_{\varkappa\lambda}(a_\varkappa - c)^\varrho (b_\lambda - d)^\sigma \tag{59}$$

und durch Ausziehen der $(\varrho + \sigma)$ten Wurzel zu dem reellen Werthe $\varepsilon_{\varrho\sigma}$, den ich als den auf (c, d) bezogenen Mittelwerth von der Ordnung (ϱ, σ) bezeichne. Es ist dann nur das Uebereinkommen zu treffen, bei geradzahligem Werthe $\varrho + \sigma$ die Wurzel dem absoluten Betrage nach zu nehmen und insbesondere, wenn sowohl ϱ als auch σ ungerade und zugleich die Summe $\varepsilon_{\varrho\sigma}^{\varrho+\sigma}$ negativ ist, den in diesem Falle auftretenden Factor $\sqrt[\varrho+\sigma]{-1}$ von $\varepsilon_{\varrho\sigma}$ abzutrennen und $\varepsilon_{\varrho\sigma}\sqrt[\varrho+\sigma]{-1}$ zu schreiben.

Setzt man nun in (59) $\varrho = 0, 1 \ldots r - 1$; $\sigma = 0, 1 \ldots s - 1$, so erhält man ein System von $n = r \cdot s$ linearen Gleichungen, aus denen die $n = r \cdot s$ Werthe $p_{\varkappa\lambda}$ gefunden werden können, wenn die $n = r \cdot s$ Werthe $\varepsilon_{\varrho\sigma}^{\varrho+\sigma}$ bekannt sind. Denn die Determinante des Systems ist stets von Null verschieden. Sind aber nur $\varrho \cdot \sigma$ Werthe bekannt, deren Ordnungsmarken die Combinationen von $0, 1 \ldots \varrho - 1$ mit $0, 1 \ldots \sigma - 1$ sind (wo $\varrho \cdot \sigma < r \cdot s$), so kann man aus den zur Verfügung stehenden $\varrho \cdot \sigma$ Gleichungen eine angenäherte Kenntniss der in passender Weise zusammengefassten p-Werthe gewinnen. Denn es treten alsdann die Ueberlegungen, welche in § 3 zu dem entsprechenden Ergebniss bezüglich der Bestimmung von n Wahrscheinlichkeitswerthen durch ν Mittelwerthe $(\nu < n - 1)$ führten, wieder in Kraft.

Es gilt daher der Satz:

Der Verlauf der Wahrscheinlichkeitswerthe $p_{\varkappa\lambda}$ innerhalb der Vertheilungstafel einer aus $n = r \cdot s$ Gliedern bestehenden Doppelreihe von Varianten $(a_1 b_1) \ldots (a_r b_s)$ wird durch die $n - 1$ Mittelwerthe

$$\begin{array}{cccc} & \varepsilon_{10} & \cdots & \varepsilon_{r-1,0} \\ \varepsilon_{01} & \varepsilon_{11} & \cdots & \varepsilon_{r-1,1} \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \varepsilon_{0,s-1} & \varepsilon_{1,s-1} & \cdots & \varepsilon_{r-1,s-1} \end{array}$$

vollständig und durch die $\nu = \varrho \cdot \sigma - 1$ Mittelwerthe ($\varrho \cdot \sigma < r \cdot s$)

$$\begin{array}{cccc} & \varepsilon_{10} & \cdots & \varepsilon_{\varrho-1,0} \\ \varepsilon_{01} & \varepsilon_{11} & \cdots & \varepsilon_{\varrho-1,1} \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \cdot & \cdot & & \\ \varepsilon_{0,\sigma-1} & \varepsilon_{1,\sigma-1} & \cdots & \varepsilon_{\varrho-1,\sigma-1} \end{array}$$

mit mehr oder minder großer Annäherung bestimmt.

Den Mittelwerthen ε_1 und ε_2, welche das Minimum an Bestimmungsstücken für einen, aus nur einer Variantenreihe bestehenden C.G. darstellen, treten somit bei einer Doppelreihe von Varianten die Mittelwerthe ε_{10}, ε_{01}, ε_{11}, ε_{20}, ε_{02}, ε_{21}, ε_{12}, ε_{22} zur Seite.

Der mittlere Fehler, der bei der Bestimmung von $\varepsilon_{\varrho\sigma}^{\varrho+\sigma}$ aus $m = \Sigma z_{\varkappa\lambda}$ Exemplaren des C.G. zu befürchten ist, lässt sich in derselben Weise wie der mittlere Fehler von ε_ν^ν berechnen und ist gleich

$$\sqrt{\frac{\varepsilon_{2\varrho,\,2\sigma}^{2\varrho+2\sigma} - \varepsilon_{\varrho,\,\sigma}^{2\varrho+2\sigma}}{m}}, \tag{60}$$

wobei jedoch die wegen der Endlichkeit des m corrigirten Werthe von $\varepsilon_{2\varrho,\,2\sigma}^{2\varrho+2\sigma}$ und $\varepsilon_{\varrho,\,\sigma}^{\varrho+\sigma}$ vorauszusetzen sind.

Die Correction von $\varepsilon_{\varrho\sigma}^{\varrho+\sigma}$ wegen der Endlichkeit des m wird ähnlich wie diejenige von ε_ν^ν in folgender Weise gewonnen.

Es möge sich der bei endlichem m beobachtete Werth $p_{\varkappa\lambda}$, welcher der Variante $(a_\varkappa, b_\lambda)$ zugehört, für unendlich großes m auf die Gesammtheit der denkbaren Varianten so vertheilen, dass der Bruchtheil $p_{\lambda\varkappa} \cdot \gamma_{\mu\nu}$ auf die Variante $(a_{\varkappa+\mu}, b_{\lambda+\nu})$ fällt. Dann ist $\sum_{\mu\nu} \gamma_{\mu\nu} = 1$ und man kann

$$\eta_{\varrho\sigma}^{\varrho+\sigma} = \sum_{\mu,\,\nu} \gamma_{\mu\nu}(a_{\varkappa+\mu} - a_\varkappa)^\varrho (b_{\lambda+\nu} - b_\lambda)^\sigma \tag{61}$$

setzen, wo die Summation über alle mit dem Systeme denkbarer Varianten verträglichen Werthenpaare μ, ν zu erstrecken ist. Nun tritt an Stelle von

$$p_{\varkappa\lambda}(a_\varkappa - c)^\varrho (b_\lambda - d)^\sigma$$

die Summe

$$\sum_{\mu\nu} p_{\varkappa\lambda} \cdot \gamma_{\mu\nu}(a_{\varkappa+\mu} - c)^\varrho (b_{\lambda+\nu} - d)^\sigma .$$

Dieselbe lässt sich, da $a_{\varkappa+\mu} - c = (a_{\varkappa+\mu} - a_\varkappa) + (a_\varkappa - c)$ und $b_{\lambda+\nu} - d = (b_{\lambda+\nu} - b_\lambda) + (b_\lambda - d)$ mit Berücksichtigung von (61) in die Form

$$p_{\varkappa\lambda}\Big\{\eta_{\varrho\sigma}^{\varrho+\sigma} + \binom{\varrho}{1}\eta_{\varrho-1,\,\sigma}^{\varrho+\sigma-1}(a_\varkappa - c) + \binom{\sigma}{1}\eta_{\varrho,\,\sigma-1}^{\varrho+\sigma-1}(b_\lambda - d)$$

$$+ \binom{\varrho}{1}\binom{\sigma}{1}\eta_{\varrho-1,\,\sigma-1}^{\varrho+\sigma-2}(a_\varkappa - c)(b_\lambda - d) + \cdots + (a_\varkappa - c)^\varrho (b_\lambda - d^\sigma\Big\}$$

bringen. Es wird somit, wenn $(\varepsilon_{\varrho\sigma})$ den corrigirten Werth $\varepsilon_{\varrho\sigma}$ bezeichnet.

$$\left.\begin{aligned}(\varepsilon_{\varrho\sigma})^{\varrho+\sigma} &= \sum_{\varkappa\lambda}\sum_{\mu\nu} p_{\varkappa\lambda} \cdot \gamma_{\mu\nu}(a_{\varkappa+\mu} - c)^\varrho (b_{\lambda+\nu} - d)^\sigma \\ &= \varepsilon_{\varrho,\,\sigma}^{\varrho+\sigma} + \binom{\varrho}{1}\varepsilon_{\varrho-1,\,\sigma}^{\varrho+\sigma-1}\eta_{10} + \binom{\sigma}{1}\varepsilon_{\varrho,\,\sigma-1}^{\varrho+\sigma-1}\eta_{01} + \\ &\quad \binom{\varrho}{1}\binom{\sigma}{1}\varepsilon_{\varrho-1,\,\sigma-1}^{\varrho+\sigma-2}\eta_{11}^2 + \cdots + \eta_{\varrho,\,\sigma}^{\varrho+\sigma}\end{aligned}\right\} \tag{62}$$

Dies ist die Correctionsformel für jedes Gesetz, welches die Vertheilung der bei endlichem m gefundenen Wahrscheinlichkeitswerthe $p_{\varkappa\lambda}$ auf das System der denkbaren Varianten bei unendlich großem m in Uebereinstimmung mit (61) regelt.

Bezieht man die Mittelwerthe $\varepsilon_{\varrho\sigma}$ auf das durch die Bedingung

$$\sum p_{\varkappa\lambda}(a_\varkappa - c) = \sum p_{\varkappa\lambda}(b_\lambda - d) = 0$$

definirte Werthenpaar (c, d), so dass $\varepsilon_{10} = \varepsilon_{01} = 0$, und setzt man — was wohl als annehmbar gelten darf — in diesem Falle

$$\eta_{\varrho\sigma}^{\varrho+\sigma} = \frac{1}{m} \varepsilon_{\varrho\sigma}^{\varrho+\sigma},$$

so wird

$$(\varepsilon_{\varrho\sigma})^{\varrho+\sigma} = \varepsilon_{\varrho\sigma}^{\varrho+\sigma} + \frac{1}{m}\left\{\binom{\varrho}{1}\binom{\sigma}{1} \varepsilon_{\varrho-1,\,\sigma-1}^{\varrho+\sigma-2} \cdot \varepsilon_{11}^2 + \binom{\varrho}{2} \varepsilon_{\varrho-2,\,\sigma}^{\varrho+\sigma-2} \varepsilon_{20}^2 \right. \tag{62a}$$
$$\left. + \binom{\sigma}{2} \varepsilon_{\varrho,\,\sigma-2}^{\varrho+\sigma-2} \varepsilon_{02}^2 + \cdots + \varepsilon_{\varrho\sigma}^{\varrho+\sigma} \right\}$$

Insbesondere ist

$$(\varepsilon_{10}) = \varepsilon_{10} = 0;\ (\varepsilon_{01}) = \varepsilon_{01} = 0;$$

$$(\varepsilon_{20})^2 = \varepsilon_{20}^2\left(1 + \frac{1}{m}\right);\ (\varepsilon_{02})^2 = \varepsilon_{02}^2\left(1 + \frac{1}{m}\right);\ (\varepsilon_{11})^2 = \varepsilon_{11}^2\left(1 + \frac{1}{m}\right);$$

$$(\varepsilon_{21})^3 = \varepsilon_{21}^3\left(1 + \frac{1}{m}\right);\ (\varepsilon_{12})^3 = \varepsilon_{12}^3\left(1 + \frac{1}{m}\right);$$

$$(\varepsilon_{22})^4 = \varepsilon_{22}^4\left(1 + \frac{1}{m}\right) + \frac{4}{m}\varepsilon_{11}^4 + \frac{2}{m}\varepsilon_{02}^2 \cdot \varepsilon_{20}^2 .$$

Dabei ist zu beachten, dass die Annahme, auf der die Formel (62a) beruht, ebenso zu begründen ist wie die frühere, welche zur Ableitung von (25a) aus (25) führte.