Anschauliche und elementare Darstellung der Lexisschen Dispersionstheorie.

Von Dr. **Georg Pólya**, Privatdozent an der Eidgenössischen Technischen Hochschule, Zürich.

Die Einzelerscheinungen — Geburten, Eheschliessungen, Verbrechen, Todesfälle — deren Gesamtanzahlen die Bevölkerungs- und Moralstatistik notiert, hängen von den verschiedensten psychischen und physiologischen Tätigkeiten der Individuen ab, deren Gesetze wir nur sehr unvollständig kennen. Die ungeheure Mannigfaltigkeit von Willensentschlüssen und physiologischen Prozessen, die sich innerhalb der Grenzen eines Staates abspielen, durchkreuzen und zusammenfinden, ist in ihren Einzelheiten so vollständig unüberblickbar, dass wir genötigt sind, diese Einzelheiten (z. B. ob in einem bestimmten Dorf an einem bestimmten Tag eine Geburt erfolgt oder nicht) glattweg dem „Zufall" zuzuschreiben. Um so bemerkenswerter ist, dass diese unzähligen, zufälligen, unberechenbaren Faktoren sich zu regelmässigen, stabilen Gesamtresultaten zusammenfügen. Die Stabilität gewisser statistischen Zahlen ist derart gross, dass wir aus ihren relativ geringfügigen Änderungen auf irgendeinen tief einschneidenden Wechsel der allgemeinen Lebensbedingungen schliessen können.

Die Dispersionstheorie von Lexis liefert uns ein zahlenmässiges Mass und eine einleuchtende Erklärung der grössern und geringern Grade von statistischer Stabilität. Diese Theorie stützt sich auf die Wahrscheinlichkeitsrechnung. Das heisst letzten Endes nur soviel, dass das rein zufällige Moment an dem statistischen Geschehen durch einen Vergleich mit den einfachsten Glücksspielen aufgeklärt wird. Wie bei Geburt eines Kindes die Alternative: Knabe oder Mädchen, so liegt bei dem Wurf mit einer Münze die Alternative „Wappen oder Schrift" vor. Die langen Serien von Knaben- oder Mädchengeburten, die die Statistik notiert, werden mit langen Serien von Würfen mit einer Münze, oder mit der häufigen Wiederholung von irgendeinem ähnlichen Glücksspiel (Ziehung aus einer Urne, Roulette usw.) verglichen. Dieser Vergleich ermöglicht die an dem einen Gebiet erworbenen Einsichten auf das andere zu übertragen.

Die anschauliche Darstellung der Lexisschen Theorie, die ich im folgenden geben werde, besteht einerseits darin, dass ich ein solches „Glücksspiel" zum Vergleich heranziehe, wo sich der allgemeine Verlauf einer langen Serie unmittelbar übersehen lässt, und anderseits darin, dass ich den Vergleich nicht mit Dazwischenschalten von mathematischen Formeln, sondern direkt ausführe, und zwar konsequent, in alle Einzelheiten ausführe. Diese unmittelbare Darstellung ist jedermann, also insbesondere auch Nichtmathematikern ohne weiteres zugänglich. Aber auch für jemand, der Mathematik wirklich beherrscht, besagt eine solche Darstellung mindestens ebensoviel wie die Formeln, die sich daran anschliessend sofort hinschreiben lassen, und sagt vielleicht noch mehr, denn sie hebt den eigentlichen Inhalt und die wesentlichen Voraussetzungen der Formeln klarer hervor.

Die nachfolgende Abhandlung ist ohne besondere Vorkenntnisse lesbar. Natürlich muss die Erlangung einer allgemeinen Einsicht ein gewisses Minimum an geistiger Energie kosten, und wenn im folgenden nur die geringsten Forderungen an die mathematische Vorbildung des Lesers gestellt werden, so muss dafür seine Phantasie etwas mehr in Anspruch genommen werden. Um den Anschluss an die gewöhnliche Ausdrucksweise der Statistiker und insbesondere das Studium der grundlegenden Arbeiten von Lexis zu erleichtern, werde ich bei jeder wichtigern Ausführung auf die parallele Stelle der Lexisschen Abhandlungen hinweisen [1]).

I. Verschiedene Modelle des normal-zufälligen Geschehens.

Die Anwendung der Wahrscheinlichkeitsrechnung auf Statistik kommt letzten Endes, wie schon gesagt, auf den Vergleich mit einem Glücksspiel heraus [2]). Dieser Vergleich ist die Hauptsache, nicht die Rechnung, und wenn wir den Mut haben, den Vergleich genau in die Augen zu fassen und in alle Einzelheiten auszumalen, so können wir uns eine grosse Menge Rechnung ersparen. Es kann z. B. die Geburt eines Kindes einem

[1]) *Lexis,* Abhandlungen zur Theorie der Bevölkerungs- und Moralstatistik. Jena, 1903.

[2]) Vgl. *Lexis,* S. 41, 102, 222 und passim.

Wurf mit einer Münze verglichen werden. Sagen wir unumwunden heraus, worin besteht die Analogie der beiden Vorgänge. Beide Vorgänge sind komplizierter Natur, aber enden mit einer einfachen Alternative. Das neugeborne Kind ist ein Knabe oder ein Mädchen, die zur Ruhe gelangte Münze zeigt entweder „Wappen" oder „Schrift". Beide Vorgänge sind zufälliger Art, d. h. ihre Komplikation entzieht sich unsrer Beurteilung, und wir können nicht im voraus sagen, welcher Fall der Alternative eintreffen wird. Die einzelnen Geburten einer statistischen Serie sind voneinander unabhängig. Das heisst: das zahlenmässige Verhältnis der Geschlechter unter den Neugebornen des Monats November übt keinen nachweisbaren Einfluss auf die Bestimmung des Geschlechts der Neugebornen im Monat Dezember aus. In ähnlichem Sinne sind die einzelnen Würfe in einer langen Partie von „Wappen- oder -Schrift"-Spiel voneinander unabhängig. Die Münze hat doch kein Gedächtnis, um ihre spätern Entscheidungen nach den frühern richten zu können. Die Münze hat nur immer die gleiche Disposition, auf Wappen oder auf Schrift zu fallen. Wir vermuten, dass auch die Anlage der Menschenrassen, männliche Nachkommen in einem gewissen Prozentsatz zu zeugen, im Laufe der historischen Zeiten merklich unverändert geblieben ist.

Es ist aber nicht zu vergessen der folgende numerische Unterschied: eine „richtige" Münze (wir haben bisher stillschweigend vorausgesetzt, dass unsre Münze „richtig" ist) hat die gleiche Anlage, Wappen oder Schrift zu zeigen, und daher werden die Anzahlen der Würfe beiderlei Art in einer sehr langen Serie von dem numerischen Verhältnis 50 % zu 50 % nicht wesentlich abweichen. Im Gegensatz dazu überwiegen, wie bekannt, in allen längern statistischen Serien die Knabengeburten. Man findet z. B., dass in der Schweiz während der 45 Jahre 1871 bis 1915 auf insgesamt 4,078,979 Neugeborne 2,095,720 Knaben und 1,983,259 Mädchen geboren worden sind. Wollen wir unsern Vergleich so einrichten, dass auch in bezug auf die durchschnittlichen Prozentsätze zwischen den Würfen beiderlei Art und den Neugebornen beiderlei Geschlechts Übereinstimmung herrscht, so müssen wir, um die statistische Wirklichkeit richtig zu wiedergeben, eine etwas asymmetrische Münze betrachten. Wir wollen aber nicht so sehr das spezielle statistische Problem behandeln, das das Geschlecht der Neugebornen betrifft. Es handelt sich hier vielmehr um die allgemeine Frage der statistischen Stabilität. Und es wird uns in Hinblick auf diese Frage vorteilhaft sein, vorerst den einfachsten und anschaulichsten zufälligen Vorgang zu betrachten, den Wurf mit einer vollständig symmetrischen Münze, die beide Seiten mit gleicher Leichtigkeit zeigt. Wollen wir hinter dieser idealen Münze ein ihr genau entsprechendes statistisches Objekt sehen, so müssen wir uns irgendeine Menschen- oder Tierrasse vorstellen, wo das mittlere Zahlenverhältnis der Nachkommenschaft beiderlei Geschlechts 50 % zu 50 % ist. Ich werde öfter von einer solchen fiktiven Population sprechen, in der die Knaben- und Mädchengeburten gleich häufig sind, und ich hoffe, der Leser wird diese Fiktion nicht zu ungeheuerlich finden.

Der Leser soll all diese Einzelheiten nicht verachten. Was uns interessiert, ist wahrlich nicht das primitive Glücksspiel „Wappen oder Schrift". Die Münze ist für uns nur eine konkrete Darstellung des rein zufälligen Geschehens. Und der Statistiker, dessen grosse Zahlen Summen von lauter zufälligen Einheiten sind — ob das einzelne neugeborne Kind Mädchen oder Knabe wird, ist wohl als zufällig zu bezeichnen — sollte den Zufall nicht verachten. Bevor man die Änderungen einer statistischen Grösse auf bestimmte Einflüsse zurückzuführen sucht, sollte man sich doch ernstlich fragen, ob die fraglichen Änderungen über den normalen Betrag von zufälligen Schwankungen hinausgehen oder nicht. Ob die Schwankungen gross oder klein sind, wird noch heutzutage nur allzuoft nach der blossen Schätzung der Zahlen oder nach dem Anblick einer Figur rein gefühlsmässig beurteilt. Man müsste aber eigentlich die Schwankungen mit einem bestimmten normalen Mass womöglich quantitativ vergleichen. Um ein solches Normalmass der grössern oder geringern Zufälligkeit zu finden, müssen wir uns mit unsrer Münze beschäftigen. Die Münze ist eben nur ein konkretes Modell des *normal-zufälligen Geschehens* [1]).

Man könnte natürlich an Stelle der Münze auch ein andres Modell benutzen, z. B. Ziehungen aus einer Urne, Versuche an der Roulette usw. oder überhaupt irgendein genau definiertes, konkretes Glücksspiel. Ich wähle im folgenden ein geometrisches Modell. Wenn der Leser dessen Einzelheiten, die auf den ersten Blick unwichtig scheinen können, mit einiger Geduld verfolgt, kann er sich relativ mühelos, insbesondere ohne längere oder höhere mathematische Ausführungen, die wichtigsten Kenntnisse über das zufällige Geschehen erwerben.

Denken wir uns ein Strassennetz, dessen regelmässig angeordnete Strassen quadratische Häusergruppen umschliessen und von Nordwest nach Südost oder darauf rechtwinklig von Nordost nach Südwest laufen (vgl. Figur 1). Ein Spaziergänger geht von einer Strassenecke *A* aus immer nach Süden, indem er manch-

[1]) Vgl. für die Bezeichnung *Lexis*, S. 183.

mal die Strassen, die nach Südost führen, manchmal die nach Südwest benützt. Diese beiden Richtungen sind ihm völlig gleich, und er entscheidet sich an jeder

Fig. 1.

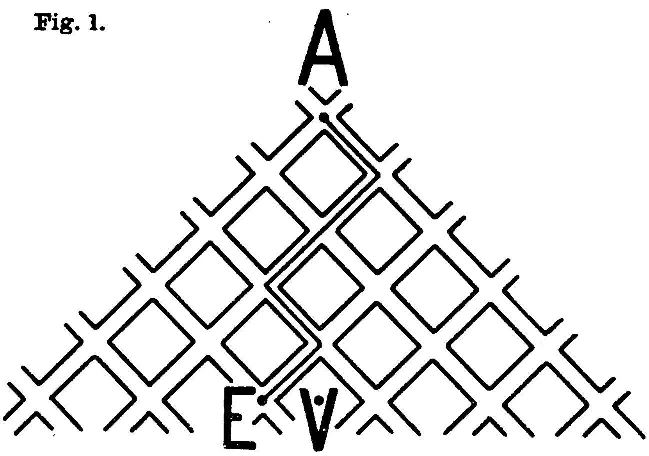
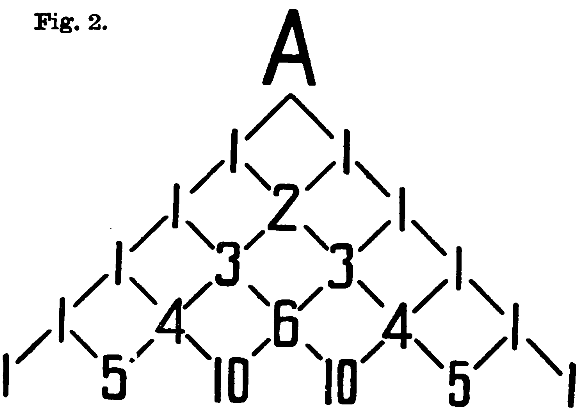


Fig. 2.



neuen Ecke, wo er ankommt, für die eine oder für die andere, nach zufälligem Einfall. Anstatt die Entscheidung nach seinen Grillen zu treffen, könnte sie der Spaziergänger auch einer Münze übertragen und sich nach Osten oder nach Westen richten, je nachdem die Münze Wappen oder Schrift zeigt.

Denken wir jetzt an die fiktive Population, wo Knaben- und Mädchengeburten die gleiche Wahrscheinlichkeit haben. Ist die Entscheidung darüber, ob ein neugebornes Kind Knabe oder Mädchen sein soll, völlig dem Zufall anheimgestellt, so ist diese Entscheidung über das Geschlecht des Kindes dem Ausfall von einem Wurf mit der Münze hinsichtlich der Zufälligkeit vollständig äquivalent. Unser Spaziergänger kann also an Stelle eines Geldstücks sich der Führung eines Zivilstandsregisters anvertrauen, und zwar so, dass er sich bei der ersten, zweiten, dritten usw. Strassenecke, wo er ankommt, sich bzw. nach der ersten, zweiten, dritten Geburtseintragung richtet, und sich nach rechts oder nach links wendet, je nachdem die betreffende Eintragung sich auf eine Knaben- oder auf eine Mädchengeburt bezieht.

Das Gebaren von diesem Spaziergänger ist höchst sonderbar und das Bild scheint gewagt. Ein gewisses Wagnis ist wohl dabei, aber es besteht nur darin, explizite und unumwunden herauszusagen, was die Anwendung von Wahrscheinlichkeitsrechnung auf statistische Probleme eigentlich bedeutet und zugleich diese Anwendung geometrisch zu illustrieren. Möglicherweise nimmt aber der Leser nicht so sehr an der Vorstellung, als vielmehr an der etwas krassen Ausdrucksweise Anstoss. Dann überlege er sich, dass der „Spaziergänger“ eigentlich bloss ein beweglicher Punkt an unsrer Figur ist, der den zeitlichen Verlauf einer statistischen Serie darstellt.

Diese geometrische Darstellung wird sich als sehr nützlich erweisen. Nehmen wir daher keinen Anstand, unser gewagtes Bild in seinen Einzelheiten auszumalen. Unser Spaziergänger geht also ständig nach Süden, pendelt aber nach den zufälligen Weisungen seines Zivilstandsregisters nach Osten oder nach Westen. Der in der Figur 1 eingezeichnete Zickzackweg AE entspricht einer Serie von fünf Geburtseintragungen: Knabe, Mädchen, Mädchen, Knabe, Mädchen. Nehmen wir die halbe Diagonale eines Quadrates als Längeneinheit an. Dann entfernt sich der Spaziergänger für jede Geburtseintragung eine Längeneinheit nach unten in der Figur 1 und geht eine Einheit nach rechts oder nach links, je nachdem die betreffende Eintragung auf einen Knaben oder auf ein Mädchen lautet. Die Gesamtanzahl der Geburten wird somit durch die vertikale Entfernung vom Punkte A, also durch die Strecke AV dargestellt. (In Figur 1 sind bloss die Endpunkte eingetragen, um die Zeichnung nicht zu überladen.) Durch die Abweichung des Spaziergängers von der Vertikallinie AV wird die Differenz zwischen den Anzahlen der Knaben- und Mädchengeburten dargestellt. Diese Differenz ist positiv oder negativ, je nachdem die Knaben- oder die Mädchengeburten in Überschuss sind, und diese beiden Fälle unterscheidend, richtet sich die repräsentative Linie VE (in Figur 1 sind bloss die Endpunkte ersichtlich) nach rechts oder nach links. In dem durch die Figur 1 speziell dargestellten Beispiel ist die fragliche Differenz $2 - 3 = -1$, und die nach links gerichtete Strecke VE stellt einen Überschuss von einer Mädchengeburt dar. Der Überschuss der Knabengeburten über die Mädchengeburten ist doppelt so gross als ihr Überschuss über die Hälfte aller Geburten, wie man sich leicht überlegt. Folglich stellt die Strecke VE durch ihre halbe Länge die Abweichung der Knabengeburten von ihrer durchschnittlichen Anzahl, d. h. in unserm (fiktiven) Falle von der Hälfte aller Geburten dar. Die Richtung der Strecke VE stellt die Richtung der Abweichung dar, und zwar zeigt die Strecke VE nach rechts, wenn unter den Gebornen mehr, und nach links, wenn sich darunter weniger Knaben befinden als es dem durchschnittlichen Verhältnis entsprechen würde.

Der Zickzackweg des Spaziergängers, der aus einer Serie von schrägen Strecken, die nach rechts oder nach links hinabfallen, besteht, ist die geometrische Darstellung des zeitlichen Verlaufs einer statistischen Serie von Knaben- und Mädchengeburten. Die Statistik notiert nur die Gesamtresultate, die geometrisch durch die beiden Strecken AV und VE angegeben werden können. Den kapriziösen Zickzackweg, der zu dem Endpunkte E führt, notiert die Statistik nicht, aber der Statistiker hat sich auch darum zu kümmern, was für ein Weg zu dem aufgezeichneten Endresultat geführt hat. Wenn insbesondere die Entscheidung darüber, ob das neugeborne Kind Knabe oder Mädchen sein soll, völlig dem Zufall überlassen ist, dann gibt die grillenhafte Irrfahrt eines Spaziergängers, der an jeder Strassenecke eine Münze nach der nächsten Richtung befragt, ein in allen wesentlichen Punkten getreues Bild von dem Verlauf einer statistischen Serie.

Ich glaube, dass der Leser seine Zeit nicht verliert, wenn er versucht, seine üblichen Begriffe und vielleicht auch einige ganz kurze konkrete Beispiele von Serien in die geometrische Sprache der Figur 1 zu übersetzen. Denn wenn einmal das erläuterte geometrische Modell des normal-zufälligen Geschehens gründlich begriffen ist, dann ist der Übergang leicht sowohl zu dem üblichen mathematischen wie auch zu dem physikalischen Modell, das ich alsbald beschreiben werde.

Das mathematische Modell des normal-zufälligen Geschehens wird in jedem Lehrbuch der Wahrscheinlichkeitsrechnung ausführlich behandelt. Hier darauf näher einzutreten, wäre unangebracht. Jedoch dass eine mathematische Behandlung dieser Fragen überhaupt möglich ist, hat für uns ein so einschneidendes prinzipielles Interesse, dass ich mindestens die Möglichkeit einer solchen Behandlung zeigen will. Das tue ich um so mehr, weil das Nötige sich an unserm geometrischen Modell besonders schön und einsichtig erläutern lässt. Tragen wir in jede Strassenecke der Figur 1, die in Betracht kommt, die Anzahl der besprochenen Zickzackwege ein, die von dem Ausgangspunkte zu der betreffenden Ecke hinführen! So entsteht die Figur 2. Man kann sich durch unmittelbares Ausprobieren aller fraglichen Zickzackwege leicht überzeugen, dass die Zahlen in den obersten Horizontalreihen der Figur 2, die dem Punkt A am nächsten liegen, richtig sind. Man kann diese Eintragung mit grösster Leichtigkeit und ganz mechanisch beliebig weit fortsetzen. Denn die Wege, die zu einer bestimmten Strassenecke führen, kommen von oben und von einer nächstliegenden Strassenecke, entweder von links oder von rechts. Folglich muss man in jede Strassenecke die Summe der beiden nächsten Zahlen eintragen, die nach oben links und nach oben rechts schon eingetragen sind [1]).

Die Summe *aller* Zahlen in der

ersten, zweiten, dritten, vierten, fünften . . .,

Horizontalreihe ist bzw.

2, 4, 8, 16, 32 . . .

Dass die Anzahl aller Zickzackwege, die zu einer gegebenen Horizontalreihe führen, doppelt so gross ist als die Anzahl derjenigen nach der vorangehenden Horizontalreihe, erklärt sich ohne weiteres daraus, dass jeder Zickzackweg sich an jeder Ecke in zwei neue Wege gabelt.

Mit Hülfe der Figur 2 können wir verschiedene Aufgaben über die fiktive Population lösen, wo das Geschlechtsverhältnis der Neugebornen 50 % zu 50 % ist. Wie häufig wird es da vorkommen, dass unter fünf auf Geratewohl zusammengelesenen Geburten, z. B. unter fünf Geburten, die zufälligerweise hintereinander in ein Zivilstandsregister eingetragen sind, drei Mädchengeburten und zwei Knabengeburten sich befinden? Das wird auf 32 Fälle durchschnittlich zehnmal vorkommen, lautet die Antwort. In der Tat, fünf Geburtseintragungen führen unsern bekannten Spaziergänger von dem Punkte A in die fünfte Zeile der Figur 1, und insbesondere die Eintragung von drei Mädchen- und zwei Knabengeburten führt ihn in den Punkt E. Auf die insgesamt 32 Wege, die von dem Punkte A in die fünfte Zeile führen, kommen, nach Figur 2, zehn solche, die in dem Punkte E enden. Daher die gegebene Lösung.

Die Anwendung der Wahrscheinlichkeitsrechnung auf statistische Probleme hat nichts Mystisches an sich. Sie besteht aus der Lösung von Aufgaben, die sich von der eben gelösten wohl in grösserer Komplikation unterscheiden, schon darum, weil sie sich auf Serien von einigen tausend oder hunderttausend Einzelfällen und nicht bloss auf eine Serie von fünf Fällen beziehen. Aber ihrem Wesen, ihrem intellektuellen Werte nach sind irgendwelche Anwendungen der Wahrscheinlichkeitsrechnung auf Statistik weder besser noch schlechter einzuschätzen als die oben gelöste primitive Aufgabe.

Wir wollen aber unsre Einsicht in das Spiel des Zufalls nicht durch mathematische Formeln hindurch,

---

[1]) Die Figur 2 stellt also das aus der Mittelschule bekannte *Pascalsche Dreieck* dar. Es wäre schon auf dieser Stufe des Unterrichts möglich, die geometrische und wahrscheinlichkeitstheoretische Bedeutung von diesem Gegenstande zu erörtern, nur leider geschieht das gewöhnlich nicht.

sondern durch unmittelbare Beobachtung erlangen. Wir könnten auch die Zahlen der Figur 2 durch Experiment und Beobachtung ermitteln. Wir könnten in dem Punkte A eine grosse Anzahl von Spaziergängern loslassen, mit dem einzigen Auftrag, eine gewisse Horizontallinie der Figur, südlich vom Punkte A, zu erreichen, und ihren zufälligen Einfällen überantworten, wie oft sie die Richtung Südost, wie oft sie die Richtung Südwest bevorzugen. Wir müssten dann bloss an jeder Strassenecke genau beobachten, welcher Teil der Spaziergänger dort vorbeigeht. Wir können aber das Experiment viel bequemer und sicherer mit lebloser Materie vornehmen. Wir denken uns die Hausblöcke der Figuren 1 und 2 abgerundet, verjüngt und verkleinert, bis sie die Dicke und Höhe von eisernen Nägeln haben. Die Nägel sind regelmässig in Quadraten angeordnet und in ein Brett eingeschlagen

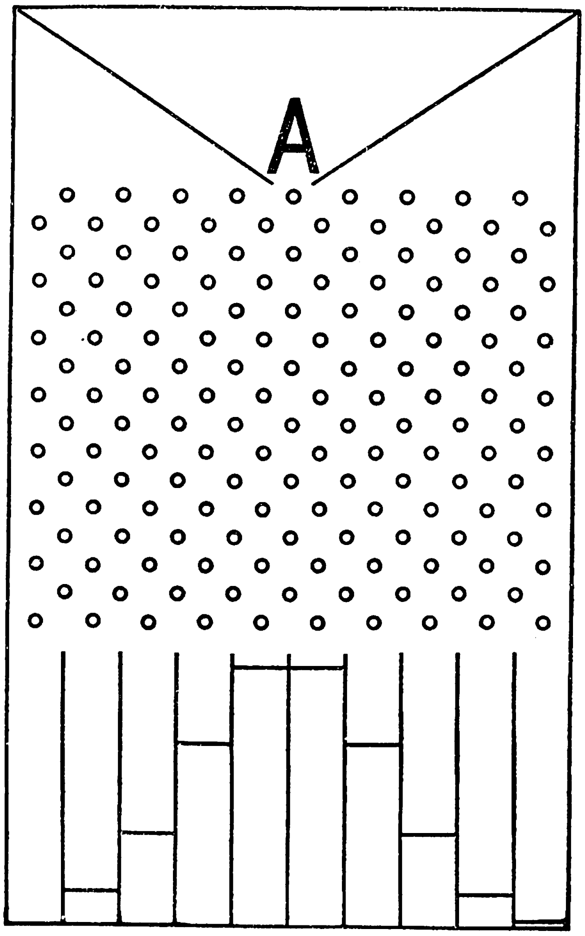


Fig. 3.

(vgl. Figur 3). Wir lassen im Punkte A an Stelle von Spaziergängern eine grosse Anzahl von Schrotkörnern los, mit Hülfe von zwei trichterförmig geneigten Leisten. Um den Schrotkörnern Auftrag nach einer bestimmten Richtung zu erteilen, stellen wir das Brett schief, indem wir es um seinen untern horizontalen Rand neigen, derart, dass die vorherige Richtung Nord-Süd, die Richtung AV der Figur 1, zur Richtung des steilsten Falles wird. Die Schrotkörner, die aus dem Trichter beim Punkte A herausfallen, streben in der Richtung des steilsten Falles nach unten, werden aber an jedem im Wege stehenden Nagel abgelenkt, nach rechts oder nach links, wie es der Zufall trifft. Um zu wissen, wo sie eine gewisse Horizontallinie passieren, arretieren wir sie sofort nach deren Übertritt in Fächern von gleicher Breite, wo sie leicht gezählt werden können. Diese Fächer sind durch vertikale Leisten abgegrenzt und am untersten Teil des Brettes angebracht. Man kann annehmen, dass die Anzahl der Schrotkörner in einem Fach der Höhe der Füllung annähernd proportional ist. Der Apparat ist in Figur 3 mit leerem Trichter und mit in den

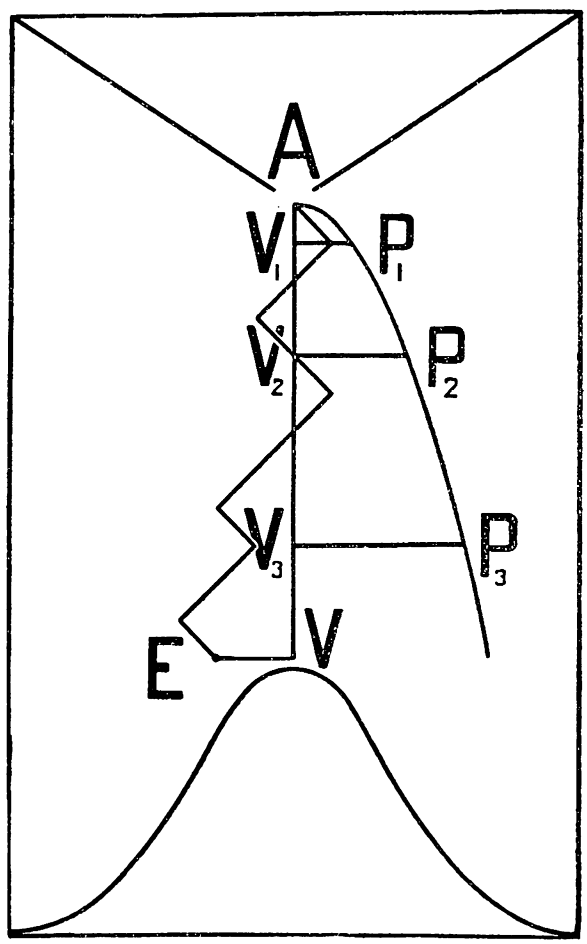


Fig. 4.

ungleich gefüllten Fächern zur Ruhe gelangten Schrotkörnern schematisch abgebildet. Ähnliche Bretter werden auch zu einem gewissen Glücksspiel benützt. Auf die Bedeutung dieses Apparates für die Theorie des Zufalls hat der englische Biologe Galton aufmerksam gemacht. Der Apparat ist unter dem Namen „Galtonsches Brett“ bekannt [1]).

Das nach ihm benannte Brett hat Galton zur Veranschaulichung der fluktuierenden Variabilität der Organismen erfunden. Es wird nicht überflüssig sein, auf diesen Punkt mindestens ganz kurz einzugehen. Denken wir uns eine Reihe von Glasröhren gleicher Dicke nebeneinander aufgestellt, worin Bohnen aus der Versuchskultur von einer Bohnenrasse nach zunehmendem Gewicht sortiert wurden [2]). Die Verteilung der einzelnen Bohnen auf die Eprouvetten hat die grösste äussere und innere Ähnlichkeit mit der Verteilung der Schrotkörner auf die Fächer des Galtonschen Apparates. Die äussere Ähnlichkeit muss nicht lange erörtert werden. Bohnen und Schrotkörner füllen die ihnen beschiedenen Fächer verschieden hoch auf, die mittlern höher, die äussern niedriger. Die innere Ähnlichkeit besteht darin, dass beiden Arten von Körnern der Zufall ihre Fächer angewiesen hat. Die Schrotkörner kamen alle aus derselben Öffnung, fielen unter der Einwirkung derselben Gravitationskraft, und es waren die zahlreichen zufälligen kleinen Zusammenstösse, die sie nach links und nach rechts pufften und in die verschiedenen Fächer warfen. Die Bohnen entwickelten sich aus derselben erblichen Anlage, waren unter merklich gleichen äussern Bedingungen an demselben Versuchsfeld auferzogen, und es waren die zahllosen kleinen Zufälligkeiten ihres Lebenslaufes, die die Entwicklung ihres Gewichtes mal hemmten, mal förderten und sie in verschiedene Gewichtsklassen verteilten. Für die fluktuierende Variabilität der Menschenrassen, mit der sich die dem Statistiker näherliegende Anthropometrie befasst, hat das Galtonsche Brett selbstverständlich eine ähnliche Bedeutung. Es führt uns das normale Schema der Variation um einen mittlern Typus vor die Augen [3]).

Wir werden an dem beschriebenen Galtonschen Brett die hauptsächlichsten Gesetze des normal-zufälligen Geschehens kennen lernen.

[1]) Der Apparat pflegt in mathematischen oder naturhistorischen Sammlungen vorhanden zu sein. Vgl. für eine andere Abbildung und Beschreibung z. B. *O. Hertwig,* Das Werden der Organismen (Jena, 1916), S. 342. Man würde einen bessern Anschluss an das Pascalsche Dreieck erreichen, wenn man die Nägel durch eng zusammenschliessende, nicht in Quadraten, sondern in regelmässigen Dreiecken angeordnete, dicke, runde Pflöcke ersetzen würde. Diese gelegentlich von mir konstruierte Modifikation des Galtonschen Brettes will ich an andrer Stelle näher beschreiben.

[2]) Vgl. eine passende Abbildung bei *O. Hertwig,* a. a. O., S. 331.

[3]) Vgl. *Lexis,* S. 101—129.

II. Die Gesetze des normal-zufälligen Geschehens.

Die Bedeutung des Galtonschen Brettes für unsre Frage wird noch mehr in die Augen springen, wenn wir das bisher Besprochene in etwas allgemeinern Ausdrücken formulieren.

Jede statistische Aufzeichnung bezieht sich auf *Serien* von bestimmten *Ereignissen.* In Betracht kommen z. B. solche *Ereignisse,* wie die Geburt eines Knaben, die Wiederverheiratung eines Witwen usw. Eine *Serie* besteht aus der Gesamtheit von allen Ereignissen, die sich innerhalb bestimmter zeitlichen und örtlichen Grenzen abspielen, z. B. in einem gewissen Lande in einem gewissen Kalenderjahr. Mit *Grundzahl* einer statistischen Serie bezeichne ich die Gesamtzahl aller Individuen, die für das betreffende Ereignis überhaupt in Frage kommen. Bei Witwenheiraten ist diese Grundzahl die Gesamtzahl aller Witwen. Die Grundzahl ist die Anzahl aller Neugebornen, wenn es sich um die Knabenquote, oder die Gesamtzahl der Individuen einer Altersklasse, wenn es sich um die Mortalitätsziffer der betreffenden Altersklasse handelt usw. Unter *Ereigniszahl* verstehe ich die Anzahl aller fraglichen Ereignisse, der Knabengeburten, Witwenheiraten, Todesfälle usw. Die Ereigniszahl ist immer kleiner als die Grundzahl, weil der Teil immer kleiner ist als das Ganze. Zum Beispiel machen die neugebornen Knaben nur einen Teil aller Neugebornen aus.

Das Hinunterrollen eines Schrotkörnchens im Galtonschen Apparat ist das mechanische Modell einer statistischen Serie, wo über das Eintreten oder Nichteintreten des Ereignisses in jedem einzelnen Falle der reine Zufall entscheidet. Unter „Ereignis“ verstehen wir dabei eine Ablenkung des Schrotkörnchens an einem Nagel *nach rechts.* Die „Grundzahl“ ist die Anzahl der horizontalen Nägelreihen in dem Apparat, da das Schrotkörnchen an jeder Reihe abgelenkt werden kann. Die Grundzahl ist also durch die Linie AV der Figur 4 dargestellt. Die Figur 4 stellt ein Schema des Galtonschen Apparates dar, mit einigen eingezeichneten Kurven, die in den nachfolgenden Auseinandersetzungen eine Rolle spielen. Insbesondere stellt die Zickzacklinie von A nach E den Weg eines Schrotkörnchens schematisch dar, d. h. die betrachtete Serie von Ablenkungen. Die „Ereigniszahl“ ist in der Figur 4 nicht unmittelbar dargestellt. Aber die Strecke VE stellt (bis auf einen Faktor 2, worauf jetzt nicht sehr ankommt) die Abweichung der Ereigniszahl von der durchschnittlichen Ereigniszahl dar, wie oben an der Figur 1 erläutert wurde. Ist die Grundzahl die Anzahl aller Geburten und die Ereigniszahl die Anzahl der darunter vorgekommenen Knabengeburten wie in den an die Figur 1 angeknüpften Erläuterungen, so stellt die Linie AV, wie man sich erinnert, eben die Anzahl aller Geburten dar, und die Strecke VE die (doppelt genommene) Abweichung der Knabengeburten von ihrem mittlern Wert. Dass in der Figur 4 nicht die Ereigniszahl selber, sondern nur ihre Abweichung vom Durchschnittswert zum Ausdruck kommt, passt für unsern Zweck gerade gut, da wir doch eben diese Abweichungen studieren wollen.

Die grosse Menge von Schrotkörnchen stellt eine grosse Anzahl von statistischen Serien dar, die alle unter denselben Grundbedingungen zustande gekommen sind und alle dieselben Grundzahlen aufweisen, da doch alle Körnchen dieselbe Anzahl von Nägelreihen passierten. Die Serien unterscheiden sich nur in den Ereigniszahlen, indem der eine Schrotkorn mehr Ablenkungen nach rechts, der andre mehr nach links erfuhr und so mehr oder weniger rechts oder links von dem Punkte V angekommen ist, der vertikal unter dem Ausgangspunkt A liegt. Die Schrotkörner, die in der Nähe des Punktes V eingetroffen sind, haben etwa ebensoviel Ablenkungen nach rechts wie nach links erfahren. Die Ereigniszahlen der entsprechenden statistischen Serien weichen also nur gering von der durchschnittlichen Ereigniszahl ab, die eben durch den Punkt V dargestellt wird. Wieviel statistische Serien unter den betrachteten eine so geringe Abweichung von der mittlern Ereigniszahl aufzeigen, macht das mittelste Fach im untern Teile des Galtonschen Brettes durch die Höhe seiner Füllung sichtbar. In den verschiedenen Fächern aufgehalten, registrieren die Schrotkörner automatisch, ob die ihnen widerfahrenen Ablenkungen nach rechts, ob die nach links überwiegen und auch das, wieviel der Überschuss über den durchschnittlichen Wert beträgt. Dass die mittlern Fächer stärker gefüllt sind als die seitwärts gelegenen, entspricht der geläufigen Erfahrung, dass kleine Abweichungen von dem Durchschnittswert sich häufiger einstellen als grosse. Die Symmetrie der Füllungen nach den beiden Seiten bestätigt die verbreitete Ansicht, dass rein zufällige Abweichungen vom Mittelwert ebenso leicht nach rechts wie nach links erfolgen können.

Die Stösse, die die einzelnen Schrotkörner in ihrem Laufe von rechts und links empfangen, und die kapriziösen Zickzackwege, die sie beschreiben, stellen die Wechselfälle und die verschlungenen Wege des Zufalls dar. Die regulär abgestufte Füllung der Fächer, in die die Schrotkörner schliesslich hineingestreut werden, gibt uns ein konkretes Bild von dem ausgleichenden Zufall, der nicht nur feste Mittelwerte entstehen lässt, sondern auch den Oszillationen um den Mittelwert und den Abweichungen davon eine gewisse eigentümliche Art von Regelmässigkeit aufprägt. In der Figur 4 ist die eckige Streuungsfigur der Figur 3 zu einer glatten Kurve ausgeglichen, die, von kleinen Besonderheiten und Zufälligkeiten befreit, die regelmässige Verteilung der kleinen und grossen, positiven und negativen Abweichungen von dem Mittelwert in grösster Reinheit zum Ausdruck bringt. Diese ist die berühmte *Gausssche Fehlerkurve*, von der ich nur sagen musste, wie sie sich in diese Betrachtung einfügt, aber die ich hier nicht näher untersuchen will[1]).

Viel wichtiger ist für uns zu lernen, wieso man eine Streuungsfigur, ähnlich wie die des Galtonschen Apparates, auf Grund gewöhnlicher statistischer Daten konstruieren kann. Ich nehme als Beispiel die Statistik der Knaben- und Mädchengeburten in der Schweiz im Zeitraum 1871—1915 (vgl. Tabelle I). Diese Statistik erstreckt sich also auf 45 Serien, wenn wir die Geburten innerhalb eines Kalenderjahres zu je einer Serie zusammenfassen. Wenn wir die Knabenquote untersuchen wollen, so müssen wir die jährlichen Zahlen von Knabengeburten als Ereigniszahlen ansehen, d. h. die Zahlen in Spalte (3). Die Grundzahlen sind in der Spalte (2) zusammengestellt. Diese Grundzahlen sind alle voneinander verschieden, und damit stossen wir auf die erste Schwierigkeit in der Konstruktion unsrer Streuungsfigur. Damit die Abweichungen vom Mittelwert gleichberechtigt seien, sollten sich alle Ereigniszahlen auf eine und dieselbe Grundzahl beziehen, wie das im Galtonschen Brett tatsächlich der Fall war, wo alle Schrotkörnchen dieselbe Höhe hinunterrollten. Wir heben diese Schwierigkeit mindestens zum grössten Teil dadurch, dass wir nicht die absoluten Ereigniszahlen in Spalte (3), sondern die in Promillen ausgedrückten relativen Ereigniszahlen in Spalte (4) untersuchen. Das heisst wir betrachten die von Jahr zu Jahr ein wenig variierende, in Promillen aller Geburten ausgedrückte Knabenquote. Die mittlere Knabenquote berechnet sich aus der Gesamtzahl aller Geburten und aus der der darunter befindlichen Knabengeburten in der ganzen 45jährigen Periode zu 513.8‰. [Vgl. Tabelle I, letzte Zeile, Spalten (2), (3), (4)]. Die Abweichungen der Knabenquote in den einzelnen Jahren von der mittlern Knabenquote sind in Spalte (5) zusammengestellt. Diese Abweichungen sind positiv oder negativ, je nachdem die Knabenquote den Wert 513.8‰ übertroffen oder nicht erreicht hat. Diese Spalte (5) ist in der Figur 5 auf die übliche Weise durch Stäbe dargestellt. Der Verlauf der Zeit ist durch die vertikale Richtung von unten nach oben wiedergegeben, entsprechend der Anordnung der Tabelle 1. Die Stäbe sind horizontal gelegt, damit ihre enge Analogie zu der Strecke VE der Figuren 1 und 4

[1]) Vgl. *Lexis*, S. 101—129, die rechte Hälfte der Abbildung S. 88 und passim.

noch mehr in die Augen springt. Die Stäbe zeigen durch ihre Richtung nach rechts oder links von der vertikalen Mittellinie eine Abweichung von dem Mittelwert in demselben Sinne an wie die Strecke *VE*.

Jetzt bleibt nur noch übrig, die Abweichungen nach Richtung und Grösse zu gruppieren. Dies geschieht in der folgenden kleinen Tabelle, die man leicht der Spalte (5) der Tabelle 1 entnimmt.

Grenzen	— 5		— 4		— 3		— 2		— 1		0		+ 1		+ 2		+ 3		+ 4		+ 5
Anzahl der Abweichungen dazwischen		1		1		3.5		7.5		12		8		5		4		2		1	

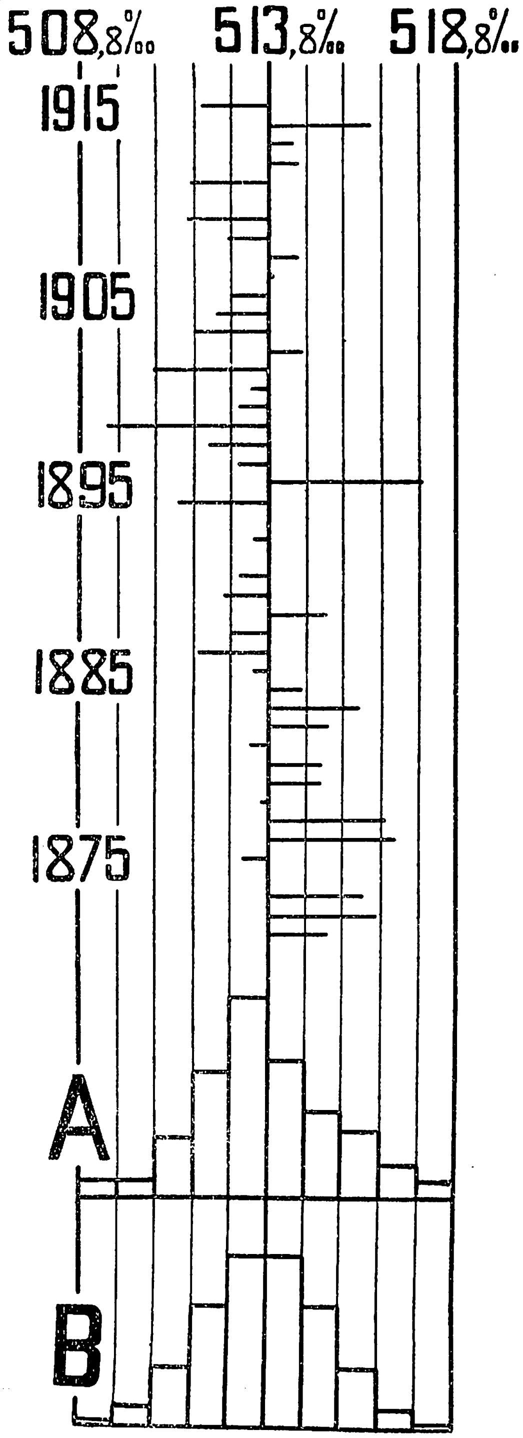


Fig. 5.

Solche Abweichungen, die genau auf die Grenze von zwei Klassen fallen, sind zur Hälfte zur einen, zur Hälfte zur andern Klasse gezählt. So können in der zweiten Zeile der Tabelle gebrochene Zahlen vorkommen. Zwischen den Grenzen — 3 ‰ und — 2 ‰ befinden sich 3 Abweichungen (Jahre 1894, 1909, 1911). Dazu kommt noch die Abweichung im Jahre 1903, die genau 2.0 ‰ ausmacht. So rechne ich in der Klasse von — 3 ‰ bis — 2 ‰ eben 3.5 Abweichungen.

Die Einteilung der Abweichungen in 10 verschiedene Klassen ist in der Figur 5 durch 11 äquidistante vertikale Linien vollzogen. Zwischen je zwei Linien enden soviel Stäbe, als Abweichungen in die betreffende Klasse fallen. Solche Endpunkte, die genau auf eine Linie fallen, zählen zur Hälfte zu beiden benachbarten Streifen. Die Rechtecke *A* in der Figur 5 stellen die obige Häufigkeitstabelle der Abweichungen auf die übliche Weise graphisch dar. Die Höhe jedes Rechtecks ist proportional der Anzahl der Abweichungen unter den betreffenden Grenzen, d. h. der Anzahl von Stäben, die in dem Streifen, wo sich das Rechteck befindet, enden. Die Rechteckfigur, die die hinuntergerollten Schrotkörner in den untern Fächern des Galtonschen Apparates bilden, zeigt zu der Rechteckfigur *A* eine weitgehende Analogie. Diese Analogie wäre überraschend, wenn sie nicht bloss ein Teil der durchgehenden Analogie wäre, die wir seit Anfang dieser Betrachtung verfolgen.

Ich muss noch auf einen Punkt aufmerksam machen. Die Streuungsfigur *A* stellt die Verteilung von 45 Abweichungen vom Mittelwert dar, ist also einer solchen Streuungsfigur im Galtonschen Apparat zu vergleichen, die bloss durch 45 hinuntergerollte Körner gebildet wird. Man begreift, dass die Zahl 45 noch viel zu klein ist, um dem Zufall volle Gelegenheit zu geben, seine regularisierende Wirkung zu entfalten. Die Abweichungen, die durch die Rechtecksfigur *A* in Figur 5 dargestellt sind, verteilen sich noch etwas unregelmässig. Zum Vergleich habe ich die Rechtecksfigur *B* berechnet und gezeichnet, die eine regelmässige Streuung darstellt, wie sie an Stelle von

45 aus sehr vielen zufälligen Abweichungen sich ergeben würde[1]).

Das mechanische Bild, das das Galtonsche Brett von einer durch die reine Zufälligkeit regierten statistischen Serie entwirft, haben wir somit durch einen wesentlichen Zug ergänzt. Wir werden Gelegenheit haben, das Bild noch weiter auszumalen und unsern ständigen Vergleich durch weitere Übereinstimmungspunkte zu vervollständigen. Wir wollen jetzt das Galtonsche Brett zur Ableitung der wichtigsten Gesetze des normal-zufälligen Geschehens benutzen. Die wenigen Formeln und die dazu führenden kurzen, klein gedruckten Auseinandersetzungen können solche Leser, die mit der Algebra an ganz schlechtem Fuss stehen, ruhig überschlagen, da ich die Resultate auch in Worten aussprechen und an Beispielen erörtern werde. Ich werde übrigens langsam, von Punkt zu Punkt, fortschreiten[2]).

a) Die durchschnittliche Abweichung. Unser Bestreben geht dahin, ein quantitatives Mass für die grössere oder geringere Schwankung statistischer Grössen aufzustellen. Die in den einzelnen Serien auftretenden Werte sind mehr oder weniger weit um einen mittlern Wert ausgestreut, zeigen grössere oder kleinere Abweichungen davon. Als einfachstes Mass der Schwankung stellt sich somit die *durchschnittliche Abweichung* ein, d. h. *das arithmetische Mittel aus den Beträgen aller Abweichungen.* Nehmen wir z. B. die Knabenquote unter den Neugebornen in der Schweiz in den Jahren 1871—1915. Die jährlichen Abweichungen von dem mittlern Wert 513.8 ‰ sind in der Spalte (5) der Tabelle I zusammengestellt. Addiert man alle diese Abweichungen, *ohne Rücksicht auf das Vorzeichen,* so erhält man die Summe 64.9 in Tabelle I, Spalte (5), letzte Zeile. Diese Summe der Abweichungsbeträge, dividiert durch 45, durch die Anzahl der fraglichen Serien, ergibt 1.44 ‰ als durchschnittliche Abweichung der jährlichen Knabenquote von der mittlern Knabenquote 513.8 ‰.

Man bezeichne mit
n die Anzahl der fraglichen Serien
$\delta_1, \delta_2, \delta_3, \cdots \delta_n$ die absoluten Beträge der Abweichungen von dem Mittelwert in der ersten, zweiten, ... n-ten Serie bzw. So ist

$$\text{durchschnittliche Abweichung} = \frac{\delta_1 + \delta_2 + \delta_3 + \cdots + \delta_n}{n}$$

[1]) Viel ausgedehntere Vergleiche ähnlicher Art, nur nicht graphisch, sondern rechnerisch ausgeführt, finden sich bei *Lexis*, S. 143—159.

[2]) Die nun folgenden Formeln und Begriffe werden bei *Lexis* ständig benutzt, insbesondere aber in den beiden Abhandlungen, S. 130—212. Formeln und Begriffe, die *Lexis* ebenfalls benutzt, wie wahrscheinliche Abweichung, Präzision usw., die aber für den wesentlichen Zweck nicht unbedingt nötig sind, werde ich nicht besprechen.

b) Die mittlere Abweichung. Die eben erklärte durchschnittliche Abweichung ist ohne Zweifel das einfachste und natürlichste Mass für die Grösse der Schwankung. Man soll sich aber über einen sehr wesentlichen Punkt nicht täuschen: sie ist nicht etwa das einzig mögliche Mass der Schwankung. An Stelle der absoluten Beträge könnten wir z. B. auch die Quadrate der Abweichungen betrachten und das arithmetische Mittel dieser Quadrate als Mass der Schwankung einführen. In der Tat, bei dem Quadrieren kommt das Vorzeichen der Abweichungen in Wegfall, und daher tragen die grossen positiven und die grossen negativen Abweichungen einmütig zur Vergrösserung des arithmetischen Mittels ihrer Quadrate bei. Nur hat dieses Mass nicht den Charakter einer Abweichung, sondern den des Quadrates von einer Abweichung. Wir erhalten daher eine Art von Abweichung, die *mittlere Abweichung,* wie man sie nennt, indem wir die *Quadratwurzel aus dem arithmetischen Mittel der Abweichungsquadrate* ausziehen.

Z. B. sind in der Tabelle I die Quadrate der Abweichungen in der Spalte (6) notiert. Ihre Summe ist 146.65 (vgl. letzte Zeile der besagten Spalte). Durch Division mit 45 erhält man 3.26 als arithmetisches Mittel der Abweichungsquadrate. Die Wurzel daraus, d. h. 1.80 ‰ ist die mittlere Abweichung.

Mit den oben eingeführten Bezeichnungen ist

$$\text{mittlere Abweichung} = \sqrt{\frac{\delta_1^2 + \delta_2^2 + \delta_3^2 + \cdots + \delta_n^2}{n}}.$$

Die Ausdrücke durchschnittliche und mittlere Abweichung geniessen in der Wissenschaft Gewohnheitsrecht, obzwar sie insoweit schlecht gewählt sind, dass sie auf das gewöhnliche Sprachgefühl nicht genügend unterscheidend wirken. Man muss sich daher einigermassen Mühe geben und nicht die Begriffe nach den Worten, sondern die Worte nach den Begriffen richten, wie sie hier definiert worden sind.

Bei einer regulären normalen Verteilung von Abweichungen, wie sie durch eine grosse Anzahl von Schrotkörnern hervorgebracht wird, die viele Stufen in dem Galtonschen Apparat hinunterrollten, sind die durchschnittliche und die mittlere Abweichung einander proportional. Dies ist an und für sich einleuchtend. Die mathematische Theorie gestattet auch den Proportionalitätsfaktor zu berechnen. Ich führe vollständigkeitshalber an, dass nach der Theorie

$$\text{mittlere Abweichung} = 1.253 \times \text{durchschnittliche Abweichung}$$

ist im Falle von regulärer Streuung. Um diese Beziehung an unsrem Beispiel zu prüfen, müssen wir die Werte

$$1.80 \text{ und } 1.253 \times 1.44 = 1.81$$

einander gegenüberstellen. Der Vergleich fällt so günstig wie nur möglich aus. Das heisst dass in der betrachteten Beziehung die Schwankungen der Knabenquote in den schweizerischen Geburten von 1871—1915 sich von rein zufälligen Schwankungen nicht merklich unterscheiden[1]).

c) Das Anwachsen der Schwankungen absoluter Ereigniszahlen mit wachsender Grundzahl. Es ist eine allgemeine und sozusagen selbstverständliche Erfahrung, die jedem Statistiker geläufig sein dürfte, dass wenn die Grundzahlen zunehmen, die Schwankungen der *absoluten* Ereigniszahlen ebenfalls zunehmen, hingegen die Schwankungen der *relativen,* d. h. in Prozenten oder Promillen ausgedrückten Ereigniszahlen abnehmen. Betrachten wir die Anzahl der Todesfälle oder Geburten oder die von irgendwelchen andern Ereignissen, die die Statistik notiert, in einer Reihe von sukzessiven Jahren, u. zw. das eine Mal in einem kleinern Kanton, das andre Mal in der ganzen Schweiz. Bilden wir die Differenzen zwischen je zwei sukzessiven Jahreswerten, so sind selbstverständlich die Differenzen in dem Kanton bedeutend kleiner. Betrachten wir hingegen eine relative statistische Grösse, die in Prozenten der Bevölkerung ausgedrückt wird, etwa die Geburts- oder die Todesziffer, so sind die Abweichungen zwischen zwei sukzessiven Jahreswerten im allgemeinen viel geringer in der Schweiz als in dem Kanton. Werden zwei entsprechende Kurven, die absolute Zahlen darstellen, die eine für das Land, die andre für den Kanton, nebeneinander gestellt, so zeigt die erste viel grössere Ordinaten und auch viel grössere Schwankungen. Wenn hingegen die zwei Kurven Prozentzahlen darstellen, so zeigen sie im grossen ganzen ähnlichen Verlauf, nur ist die auf das ganze Land bezügliche viel glatter als die andre. Aber die grossen Zahlen gleichen nur die Schwankungen der Relativzahlen aus. Die Schwankungen der absoluten Ereigniszahlen schwellen an, wenn die Grundzahl anwächst.

In *welchem Masse* findet jedoch dieses Anwachsen statt? Offenbar kann die Zunahme der Schwankung der absoluten Zahlen der Zunahme der Grundzahl nicht einfach proportional sein, denn sonst könnte doch die Schwankung der relativen Zahlen nicht abnehmen. Hier stehen wir vor einer Frage von fundamentaler Wichtigkeit. Wir werden sie für rein zufällige Schwankungen mit Hülfe des Galtonschen Apparates beantworten.

[1]) Die Zahl 1.253 ist aus der Kreiszahl $\pi = 3.1416$ durch Division mit 2 und Quadratwurzelziehen abgeleitet. Die Relation zwischen durchschnittlicher und mittlerer Abweichung verifiziert *Lexis* an umfangreichem Stoff, S. 161—163.

Der Vorgang an dem Galtonschen Apparat ist ein physikalischer Vorgang, und zwar eine Fallerscheinung. Es müssen also darin die einfachen physikalischen Gesetze des Falles durch alle Zufälligkeiten hindurch letzten Endes doch zur Gültigkeit gelangen. Sie müssen insbesondere dann zum Ausdruck kommen, wenn wir Mittelwerte betrachten, wo die vielen kleinen entgegengesetzten Störungen sich aufheben können. Betrachten wir z. B. den Durchschnittswert (d. h. das arithmetische Mittel) aller positiven Abweichungen. Da bei rein zufälligen Änderungen positive Abweichungen sich ebenso leicht einstellen können wie negative, hat der Durchschnittswert der positiven denselben Betrag als der der negativen Abweichungen, oder auch denselben, den das arithmetische Mittel aller Abweichungen besitzt, das wir unter *a)* kurz als durchschnittliche Abweichung bezeichnet haben. Der Durchschnittswert der positiven Abweichungen hat selber den Charakter von einer Abweichung und wird daher durch eine von der vertikalen Mittellinie nach rechts gerichtete horizontale Strecke dargestellt. Wir sollen die Veränderung der Grundzahl studieren. Die Grundzahl in dem Galtonschen Apparat wird jedoch durch die Anzahl der Stufen horizontaler Nägelreihen dargestellt, die die Schrotkörner hinunterrollen (vgl. oben). Sind die verschiedenen Grundzahlen durch die wachsenden vertikalen Strecken AV_1, AV_2, AV_3 dargestellt, so werden die dazugehörigen durchschnittlichen Abweichungen durch die horizontalen Strecken V_1P_1, V_2P_2, V_3P_3 dargestellt (vgl. Figur 4). Diese Abweichungen wachsen mit wachsender Grundzahl, da doch die hinunterströmende Menge der Schrotkörner sich in der Tiefe immer mehr und mehr zerstreut.

Was ist das mathematische Gesetz von dieser anwachsenden Streuung? Wir stellen diese Frage schon das zweite Mal, jetzt ist sie nur etwas anders gefasst. Jedes Schrotkörnchen, das an dem Galtonschen Brett hinunterrollt, erhält von dessen Nägeln eine grosse Anzahl horizontaler Impulse, die teils nach rechts, teils nach links gerichtet sind und sich wohl zum guten Teil aufheben. Die Schrotkörner, die sich rechts von der Linie AV befinden, haben offenbar mehr Impulse in dieser Richtung erhalten. Die Punkte P_1, P_2, P_3 bezeichnen die Bahn von einer Art durchschnittlichen Bewegung der Schrotkörner, die in der rechten Hälfte des Apparates hinunterrollen. Diese durchschnittliche Bewegung wird offenbar viel einfacher sein als die Bewegung der einzelnen Schrotkörner. Wenn sich diese in grillenhaften Zickzackwegen bewegen, so ist anzunehmen, dass die durchschnittliche Bewegung, worin sich die kleinen Zufälligkeiten der Einzelbewegungen aufheben, in einer annähernd glatten Linie erfolgt. Es ist sogar anzunehmen, dass diese Bahnkurve

$A\,P_1, P_2, P_3$ den einfachen Fallgesetzen folgt und dass die durchschnittliche Bewegung annähernd so aussieht, als ob der dominierende, nach rechts gerichtete Impuls einem materiellen Teilchen schon im Ausfallspunkte A erteilt wäre und dann dieses materielle Teilchen ungehindert von Nägeln, sogar von Reibung und Luftwiderstand, hinunterfallen würde.

Wie die Physik lehrt, beschreibt das fallende Teilchen unter diesen Bedingungen eine Parabel, d. h. das Gesetz seiner Bahn ist das folgende: die horizontalen Entfernungen wachsen so wie die Quadratwurzeln der vertikalen Entfernungen. In unsrer Figur 4 verhalten sich die vertikalen Entfernungen

$$A V_1 : A V_2 : A V_3 = 1 : 4 : 9.$$

Darum verhalten sich die horizontalen Entfernungen

$$V_1 P_1 : V_2 P_2 : V_3 P_3 = 1 : 2 : 3.$$

Das Galtonsche Brett ist nur eine konkrete Darstellung des normal-zufälligen Geschehens. Unser Resultat bedeutet also folgendes: *In dem normal-zufälligen Geschehen wachsen die Schwankungen der absoluten Ereigniszahlen wie die Quadratwurzeln aus den Grundzahlen.*

Dieses Resultat ist von grösster prinzipieller Wichtigkeit. Wir wollen es an einem schematischen Beispiel erläutern. Vergleichen wir drei Statistiken, die sich auf dasselbe Ereignis beziehen, nur in Gebieten von verschiedener Grösse aufgenommen sind, etwa in einem Dorf, in einer Stadt und in einem Land. Dementsprechend sind die Grundzahlen verschieden. Sie seien etwa

100 10,000 1,000,000.

Die Quadratwurzeln aus diesen Zahlen sind

10 100 1000.

Daher sind in den drei Gebieten Abweichungen der jährlichen Ereigniszahl von deren mittlerm Wert, die bzw.

5 50 500

Einheiten betragen, als gleichzufällig zu betrachten. Sollte der mittlere Wert der Ereigniszahl die Hälfte der Grundzahl sein, so ist also in den drei Gebieten mit dem gleichen Recht zu erwarten, dass das Ereignis sich bzw.

45 bis 55 4950 bis 5050 499 500 bis 500 500

Mal einstellt. Man kann sich unter Grundzahl die Gesamtzahl der Geburten, unter Ereigniszahl die Anzahl der Knabengeburten vorstellen, wenn man sich wieder die Fiktion eines Landes gefallen lässt, wo die Neugeburten beiderlei Geschlechts im Mittel gleich häufig sind.

Die Schwankungen der absoluten Ereigniszahlen wachsen also nicht der Grundzahl proportional, sondern viel weniger stark, nämlich nur der Quadratwurzel proportional. Dadurch erklärt sich die jetzt zu besprechende

d) Abnahme der Schwankungen relativer Ereigniszahlen mit wachsender Grundzahl. Unter relativer Ereigniszahl verstehe ich die in Promillen (oder Prozenten) der Grundzahl ausgedrückte Ereigniszahl, d. h. die tausend- (hundert-) fache Ereigniszahl dividiert durch die Grundzahl[1]). Solche relative Ereigniszahlen sind z. B., wie schon oben erwähnt, die Knabenquote (in Promillen aller Geburten), die Todesziffer (in Promillen der Gesamtbevölkerung) usw. Da die Quadratwurzel dividiert durch den Radikanden die reziproke Quadratwurzel ergibt, folgt aus der Begriffsbestimmung nach dem Resultat unter *c)*, dass die *Schwankungen relativer Ereigniszahlen mit wachsender Grundzahl abnehmen, und zwar proportional der reziproken Quadratwurzel aus der Grundzahl.*

Greifen wir auf das eben behandelte schematische Beispiel am Ende der Ausführungen unter *c)* zurück. In den drei dort erwähnten Gebieten kann also mit der gleichen Wahrscheinlichkeit erwartet werden, dass die relative Ereigniszahl bzw.

450 bis 550 ‰, 495 bis 505 ‰, 499.5 bis 500.5 ‰

beträgt. Die Abweichungen dieser Promillen von dem mittlern Wert 500 sind bzw.

50 ‰ 5 ‰ 0.5 ‰.

Diese Abweichungen verhalten sich ebenso wie

$$\frac{1}{10} \qquad \frac{1}{100} \qquad \frac{1}{1000}$$

d. h. wie die reziprok genommenen Quadratwurzeln der Grundzahlen

100 10,000 1,000,000.

Ich muss wohl nicht mit vielen Worten erörtern, dass das Gesagte den eigentlichen Grund der allgemeinen Bestrebung enthüllt, statistische Verhältniszahlen aus möglichst grossen Grundzahlen abzuleiten. Wir haben doch eben das mathematische Gesetz gefunden, nach dem die zufälligen, die Auffindung des wahren mittlern Wertes gefährdenden Schwankungen der statistischen Verhältniszahlen abnehmen, wenn die Grundzahl zunimmt.

[1]) „Bei der Aufstellung statistischer Verhältniszahlen ist vorzugsweise die Form des einfachen Wahrscheinlichkeitsbruches zu empfehlen, der allenfalls noch mit 100 oder 1000 multipliziert werden mag." Diese Empfehlung von *Lexis* (S. 174) ist wohl begründet in der grössren Einfachheit der nötigen Formeln. Um ihr zu folgen, betrachte ich in meinen Beispielen nicht die traditionelle Grösse: Anzahl der Knabengeburten auf 1000 Mädchengeburten, sondern diese Anzahl auf 1000 Geburten *überhaupt* berechnet.

Ich bezeichne die Grundzahl mit g. Dann ist, nach dem unter *c)* Gesagten, die durchschnittliche Abweichung der absoluten Ereigniszahlen proportional $\sqrt{g}$ und folglich die durchschnittliche Abweichung der in Promillen ausgedrückten relativen Ereigniszahlen proportional

$$\frac{1000\sqrt{g}}{g} = \frac{1000}{\sqrt{g}},$$

oder, mit einem andern Proportionalitätsfaktor, einfach $\frac{1}{\sqrt{g}}$ proportional.

e) Theoretische Berechnung der mittlern Abweichung aus der mittlern relativen Ereigniszahl. Wir haben unter *a)* und *b)* zwei verschiedene Masse der Schwankung kennen gelernt, die wir als durchschnittliche und mittlere Abweichung voneinander unterschieden haben. Man könnte natürlich beliebig viele andre ähnliche Masse der Schwankung einführen, wie sich der Leser leicht denken kann. Sie müssten nur immer komplizierter werden, und die Nützlichkeit würde mit der Komplikation nicht Schritt halten. Wie eine nähere Untersuchung zeigt, ist die mittlere Abweichung das geeignetste Schwankungsmass. Das grundlegende Gesetz, das wir unter *c)* gefunden haben, und das die Grösse der Abweichungen mit der Grösse der Grundzahl in Zusammenhang bringt, ist für die mittlere Abweichung viel genauer erfüllt als für irgendeine andre Abweichung. Das lehren schwierige mathematische Untersuchungen, die nicht hierher gehören. Ferner kann die mittlere Abweichung mit ziemlicher Sicherheit vorausgesagt werden, wenn nur der mittlere Wert der relativen Ereigniszahl bekannt ist. Darüber kann man sich mit den ersten Elementen der Algebra überzeugen.

Man muss sich nur zuerst in einen äussersten Fall versetzen, die die Statistik nie explizite betrachtet: man muss zuerst Serien untersuchen, deren Grundzahl Eins is. Es soll auf 1000 Fälle durchschnittlich V Mal das Ereignis eintreten. In unsrer Terminologie ist also V die mittlere in Promillen ausgedrückte relative Ereigniszahl. In 1000 „Serien" von der Grundzahl Eins (so kann man noch die Sache formulieren) wird das Ereignis V Mal vorkommen und $1000 - V$ Mal unterbleiben.

In einer „Serie" von der Grundzahl Eins können nur zwei verschiedene „relative Ereigniszahlen" sich einstellen, nämlich

$$1000\,‰ \quad \text{und} \quad 0\,‰,$$

je nachdem das Ereignis eintrifft oder nicht eintrifft. Daher gibt es nur zwei mögliche Abweichungen von der mittlern relativen Ereigniszahl V, nämlich

$$1000 - V \quad \text{und} \quad 0 - V.$$

Diese Abweichungen werden bei tausendfacher Wiederholung der Serie

$$V \quad \text{bzw.} \quad 1000 - V$$

Mal vorkommen. Das arithmetische Mittel der Abweichungsquadrate, von der die Regel unter *b)* spricht, ist daher

$$\frac{V\,(1000 - V)^2 + (1000 - V)\,(0 - V)^2}{1000} = V\,(1000 - V),$$

wie man nach leichter algebraischen Umformung findet. Man erhält daraus, ebenfalls nach der zitierten Regel, dass die mittlere Abweichung der relativen Ereigniszahl für die Grundzahl 1 gleich

$$\sqrt{V\,(1000 - V)}$$

ist. Die mittlere Abweichung befolgt das unter *d* erörterte Gesetz so genau, dass wir aus dem für die Grundzahl 1 gültigen Ausdruck ohne weiteres die mittlere Abweichung der relativen Ereigniszahl für die Grundzahl g ableiten können durch einfache Division mit $\sqrt{g}$. Letztere ergibt sich also zu

$$\frac{\sqrt{V\,(1000 - V)}}{\sqrt{g}} = \sqrt{\frac{V\,(1000 - V)}{g}}.$$

Das Resultat der elementaren Rechnung lässt sich so in Worte fassen: *Man multipliziere die mittlere relative Ereigniszahl mit der mittlern relativen Anzahl der Fälle, wo das Ereignis ausbleibt (beide Anzahlen müssen in Promillen ausgedrückt sein und sie ergänzen einander zu 1000 ‰). Man dividiere das Produkt durch die Grundzahl und man ziehe aus dem entstehenden Quotienten die Quadratwurzel aus. So wird der mutmassliche Wert der mittlern Abweichung erhalten* [1]).

Bei der Anwendung dieser ausserordentlich wichtigen Vorschrift auf wirkliche statistische Zahlen darf man nicht aus den Augen verlieren, dass die verschiedenen Serien einer und derselben Statistik im allgemeinen verschiedene Grundzahlen haben. Greifen wir auf unser vorheriges Beispiel, auf die schweizerische Statistik der Knabenquote von 1871 bis 1915 zurück. In den 45 Serien dieser Statistik variiert die Grundzahl von 77,931 bis 100,635. (Das Maximum ist in 1901, das Minimum in 1915 erreicht worden.) Wir nehmen also als Grundzahl die durchschnittliche Grundzahl an, d. h. die Anzahl von allen Geburten der Periode 1871—1915 dividiert durch 45. Das ist 4,078,979 : 45 = 90,644. Die mittlere relative Ereigniszahl ist in unserm Falle die Knabenquote der Periode 513.8 ‰. Das „Ereignis", das wir untersuchen, ist eine Knabengeburt. Die mittlere relative Anzahl der Fälle, wo dieses Ereignis ausbleibt, ist also die Mädchenquote der Periode 486.2 ‰. Somit ist der vorschriftsmässige Wert der mutmasslichen mittlern Abweichung die Quadratwurzel aus

$$\frac{513.8 \times 486.2}{90644} = 2.756,$$

d. h. 1.66 ‰. Dieser Wert stimmt mit dem vorangehend unter *b)* ausgerechneten Wert 1.80 ‰ nicht genau überein. Eine vollständige Übereinstimmung hätte man verständigerweise auch nicht erwarten können. Aber die gefundene näherungsweise Übereinstimmung muss als sehr merkwürdig bezeichnet werden, wenn wir uns vorstellen, was für eine lange, und wir können wohl sagen, kühne Gedankenreihe uns zum Vergleich der beiden Zahlen 1.66 ‰ und

[1]) Diese Berechnungsweise nennt *Lexis* die „statistische" (vgl. S. 134—135) im Gegensatz zu der unter *b)* erörterten unmittelbaren Berechnungsweise, die er als die „physikalische" bezeichnet (vgl. S. 136—137).

1.80 ‰ geführt hat. Übrigens gehört eine nähere Kritik dieser Übereinstimmung in die folgenden Abschnitte. Die ausgeführte Rechnung hat für uns vorderhand nur Wert als Exemplifizierung der gefundenen Rechenvorschriften.

Das für uns wichtigste Gesetz des normal-zufälligen Geschehens (ob dieses Gesetz auf das sozialstatistische Geschehen anwendbar ist, bleibt vorderhand dahingestellt) besagt, dass die Resultate der beiden unter *b)* und unter *e)* vorgeschriebenen wesentlich verschiedenen Rechnungen näherungsweise übereinstimmen müssen. Wir könnten unsre Entwicklung an diesem Höhepunkte eigentlich abbrechen. Ich will nur kurz eingehen auf eine für uns nicht so sehr wichtige, aber allgemein benutzte

f) Verfeinerung der Rechnung unter b). Würden wir mit einer noch so idealen symmetrischen Münze mehrere längere Serien von Würfen unternehmen, so würde wohl in den meisten Serien das Vorkommen von Wappen und Schrift von dem idealen Verhältnis 500 ‰ zu 500 ‰ etwas abweichen. Wenn wir die relative Ereigniszahl der Schriftwürfe aus den zusammengefassten Würfen aller Serien ausrechnen würden, so würden wir höchst wahrscheinlich wieder nicht genau 500 ‰ dafür bekommen, sondern einen davon etwas verschiedenen Durchschnittswert. Dieser *experimentelle* Durchschnittswert würde jedoch mit den Werten in den einzelnen Serien besser übereinstimmen als der *wahre* Durchschnittswert 500 ‰. Folglich würden die einzelnen Abweichungen von dem experimentellen Durchschnittswert etwas kleiner ausfallen als von dem wahren. Zieht man also nicht die Abweichungen von dem wahren, sondern die von dem experimentellen Durchschnittswert in die Rechnung, der unter *b)* gegebenen Vorschrift gemäss, so erhält man einen etwas zu kleinen Wert für die mittlere Abweichung.

Diesem Übel abzuhelfen wäre um so notwendiger, weil wir in der Wirklichkeit die „wahren" Durchschnittswerte nirgendwo ermitteln können und nur mit experimentellen Durchschnittswerten operieren. Der Gefahr, dass die mittlere Abweichung zu klein ausfällt, kann man nun, wie die Theorie lehrt, mit Vorteil auf die folgende Weise begegnen: wir ziehen nicht das arithmetische Mittel der Abweichungsquadrate in die Rechnung, d. h. die Summe dieser Quadrate dividiert durch die Anzahl aller Serien, sondern eine etwas grössere Zahl: die Summe der fraglichen Quadrate, dividiert durch die *Anzahl der Serien weniger Eins.* In unserm Beispiel erhielten wir früher unter *b)*

$$\sqrt{\frac{146.65}{45}} = 1.80\ ‰.$$

Wir erhalten jetzt

$$\sqrt{\frac{146.65}{44}} = 1.82\ ‰.$$

Der Unterschied ist unwesentlich, wie in allen Fällen, wo die Anzahl der Serien nicht ganz gering ist.

Die Formel unter *b)* ist also folgendermassen abzuändern:

$$\text{Mittlere Abweichung} = \sqrt{\frac{\delta_1^2 + \delta_2^2 + \delta_3^2 + \cdots + \delta_n^2}{n-1}}.$$

III. Die Entstehung der übernormalen Dispersion.

Der Einzelfall einer statistischen Serie, z. B. der einzelne Todesfall, ist durch ein komplexes Ursachensystem bedingt, das die Statistik nicht vollständig entwirren kann und nicht vollständig zu entwirren hat. Vielmehr werden zwei Arten von bedingenden Umständen unterschieden: die ständigen und die unregelmässigen. Die ständigen Ursachen sind diejenigen, die systematisch, in allen Fällen gleichmässig, durchgehends in derselben Richtung wirken wie z. B. die wirtschaftliche Lage, die Gesamtheit der Gesetze und deren mehr oder weniger strenge Handhabung, die erbliche Anlage der Rasse usw. Man kann die Gesamtheit der ständigen Ursachen kurz als die *Anlage* bezeichnen. Die unregelmässigen Ursachen variieren von Fall zu Fall, wirken unsystematisch mal in der einen, mal in der andern Richtung, und man pflegt ihre wirre, unfassbare, vag abgegrenzte Gesamtheit als den *Zufall* zu bezeichnen. In konkreten Fällen kann es sehr zweifelhaft sein, was zur Anlage gehört und was auf die Rechnung des Zufalls zu setzen ist. Dem Sinne nach ist aber die Unterscheidung ziemlich klar und wird auch immer gemacht, nur mehr oder weniger bewusst, mehr oder weniger explizite[1]). Die Statistik kümmert sich nicht um die individuell variablen Ursachen, sondern sucht den konstanten Teil der bedingenden Umstände zu isolieren, der in allen Fällen auf ähnliche Weise wirkt, d. h. sie sucht die Anlage kennen zu lernen. Es gehört zu den vornehmsten Aufgaben der Statistik, die Änderungen der statistischen Zahlen von Serie zur Serie verfolgend festzustellen, in welchem Masse diese Änderungen als rein zufällige Schwankungen aufzufassen, in welchem Masse sie bestimmten, unterscheidbaren Einflüssen, d. h. in der vorgeschlagenen Terminologie, den Änderungen der Anlage zuzuschreiben sind. So ist z. B. die durchgehende Abnahme der Sterblichkeit, die vor dem Krieg in den meisten Ländern beobachtet wurde, wohl nicht auf den Zufall, sondern auf den Fortschritt der hygie-

[1]) Vgl. *Lexis*, S. 131–132 und passim.

nischen Einrichtungen, auf die Ausdehnung und die strengere Handhabung der Vorschriften über ansteckende Krankheiten usw., kurz auf solche Einflüsse zurückzuführen, die ich hier nicht aufzuzählen brauche, die man aber wohl als eine Änderung der Sterblichkeitsanlage innerhalb der betreffenden statistischen Gesamtheiten bezeichnen darf.

Der Begriff des normal-zufälligen Geschehens hängt mit dem eben erörterten statistischen Begriff der Anlage eng zusammen. Denken wir das normal-zufällige Geschehen wieder durch eine Münze konkret dargestellt. Eine wirkliche Münze würde sich nach einigen hunderttausend Würfen abnutzen, verbiegen, beschädigen und ihre Anlage, Wappen oder Schrift hervorzubringen, merklich ändern. Nicht so die ideale Münze, mit der wir in Gedanken experimentieren; diese wird nach Millionen von Würfen genau dieselbe Anlage besitzen, wie vor dem allerersten und ihre beiden Seiten mit gleicher Wahrscheinlichkeit zeigen. Ebenso ist irgendeinem idealen Glücksspiel, mit dem wir den Begriff des normal-zufälligen Geschehens illustrieren können, eine gewisse absolute Konstanz der Disposition eigen, die am Ende über die unregelmässig veränderlichen äussern Bedingungen obsiegt und feste Mittelwerte hervortreten lässt.

Das normal-zufällige Geschehen ist also ein idealer Grenzfall des wirklichen statistischen Geschehens. In diesem Grenzfall ist eine völlige Konstanz der Anlage, z. B. die unveränderliche Symmetrie der Münze, mit einem völlig regellosen Wechsel der übrigen Bedingungen verbunden. In dem Fall der Münze z. B. müssen wir uns vorstellen, dass die Impulse, die ihr bei den sukzessiven Würfen erteilt werden, nach Grösse und Richtung völlig unsystematisch variieren. Die Schwankungen im normal-zufälligen Geschehen werden eben nur darum vollständig auf die Rechnung des Zufalls gesetzt, weil sie nicht aus irgendwelchen Veränderungen der Anlage, die doch als völlig unveränderlich supponiert wird, erklärt werden können.

Die wirklichen statistischen Grössen können sich diesem idealen Grenzfall des normal-zufälligen Geschehens mehr oder weniger nähern, sie können ihn nie erreichen. Die den statistischen Serien zugrundeliegenden Anlagen ändern sich von Jahr zu Jahr, oder eigentlich stetig im Laufe der Zeit. Das Zahlenverhältnis der Geschlechter unter den Neugebornen ist vielleicht diejenige statistische Grösse, die die grösste Konstanz der Anlage und demzufolge die grösste Ähnlichkeit mit einem idealen Zufallsspiel zeigt, aber auch die Schwankungen von dieser statistischen Verhältniszahl können nicht als rein zufällig aufgefasst werden, wie es später besprochen wird. Die wirklichen statistischen Grössen können, wie gesagt, sich nur in grösserm oder geringerm Grade dem idealen Schema der reinen Glücksspiele nähern, und wir sind eben im Begriffe, ein exaktes Mass zur zahlenmässigen Beurteilung dieser Annäherung aufzustellen. Wir wollen also einen quantitativen Anhaltspunkt zur Entscheidung der Frage gewinnen, in welchem Masse die Schwankungen statistischer Grössen dem Zufall, in welchem Masse sie den Änderungen der Anlage, d. h. bestimmten, durchgehenden, unterscheidbaren Einflüssen zuzuschreiben sind. Die Wichtigkeit dieser Aufgabe kann nicht übersehen werden. Es ist doch eine alltägliche und doch die vornehmste Aufgabe der Statistik, in den Schwankungen ihrer Zahlen bestimmbare Einflüsse von dem reinen Spiel des Zufalls zu unterscheiden.

Das Galtonsche Brett, das uns die Gesetze des normal-zufälligen Geschehens vor die Augen geführt hat, kann uns auch die jetzt in Frage kommenden Verhältnisse erklären, und zwar anschaulicher und einsichtiger, als irgend eine Rechnung es tun könnte.

Die dem normal-zufälligen Geschehen zugrundeliegende invariable Anlage wird an dem Galtonschen Brett durch die unveränderliche Richtung der Schwerkraft dargestellt, die die Schrotkörner immer nach unten richtet, während die zufälligen Zusammenstösse mit den Nägeln des Brettes unregelmässig variieren, mal nach links, mal nach rechts wirken. Denken wir uns, dass, entsprechend den einzelnen statistischen Serien, die in der Zeit aneinander folgen, die einzelnen Schrotkörner nacheinander hinuntergelassen werden. Jeder Schrotkorn soll also in dem Augenblick im Punkte A des Apparates losgelassen werden, wo der vorangehende Schrotkorn in einem der untern Fächer zur Ruhe gelangt. Die nacheinander hinunterrollenden Schrotkörner erfahren dieselbe ständig wirkende Ursache, sie legen ihren Weg unter dem Einfluss der unveränderlich gerichteten Schwerkraft zurück. Der sozialstatistischen Wirklichkeit entspricht das allerdings nicht; die sozialen Anlagen variieren von Jahr zu Jahr. Wie sollen wir also diese Variation der Anlage, die von Serie zu Serie immer etwas anders gerichtet ist, an unserm Modell darstellen? Offenbar so, dass wir *die Richtung der Schwerkraft relativ zum Galtonschen Brett variieren* lassen, indem wir das Brett mehr oder weniger schief stellen, für jeden einzelnen Schrotkorn etwas anders. (Es ist die Grundlinie des Galtonschen Brettes, woran die Fächer angeordnet sind, die in unsrer ursprünglichen Anordnung horizontal war, die jetzt mehr oder weniger geneigt zu stellen ist, damit die Richtung des steilsten Falles an der Brettfläche variiert.) Wenn die soziale Anlage eine gerichtete Evolution durchmacht, so müssen wir das Brett immer mehr und mehr nach der einen Seite neigen. Wenn die soziale Anlage sich undulatorisch ändert, in mal anschwellender, mal abflachender Wellenbewegung, so müssen wir das Galtonsche Brett hin- und herpendeln lassen. Es gibt statistische Grössen, die auf die Änderung der sozialen Lebensbedingungen ziemlich empfindlich reagieren, wie z. B. die Heirats- oder die Geburtsziffer. Solche Grössen darstellend muss sich das Galtonsche Brett in grossen heftigen Ausschlägen bewegen. Andre Grössen, wie z. B. die Knabenquote, entspringen mehr physiologischen Bedingungen, die sich im Laufe der Zeit kaum ändern. Zur Darstellung solcher Grössen müssen wir auf fast unmerkliche leise Verschiebungen des Galtonschen Brettes denken.

Ob die Veränderung der Anlage undulatorisch, ob sie evolutorisch ist, macht in einer Beziehung keinen Unterschied aus: in beiden Fällen werden die Schwankungen der Ereigniszahlen grösser. Das sehen wir am Galtonschen Brett. Wir können das Galtonsche Brett hin- und herpendeln lassen, oder auch immer mehr auf die eine Seite neigen, die hinunterströmende Schrotkornmasse wird sich in beiden Fällen, wie man sich leicht denken kann, mehr zerstreuen als an dem ruhendem Brett. Es entsteht in beiden Fällen eine *übernormale Dispersion.* In der Tat: an dem ruhenden Brett rührt die Streuung der Schrotkörner nur von den zufälligen Hindernissen, von den Nägeln her. Am bewegten Brett gesellt sich zu dieser *normalzufälligen Schwankungskomponente* noch eine *„physikalische“ Schwankungskomponente,* die von den Bewegungen der ganzen Anlage des Apparates herrührt[1]).

Jetzt kommen wir zu der fundamentalen Frage: wie sollen wir die nicht zufällige, die von den Änderungen der Anlage herrührende Schwankung einer statistischen Grösse an den Ereigniszahlen der sukzessiven Serien ansehen? Um die Antwort auf diese Frage vorzubereiten, haben wir eben die Gesetze des normal-zufälligen Geschehens vorher entwickelt. Unser Hauptresultat war, dass für das normal-zufällige Geschehen die mittlere Abweichung aus der Grundzahl und aus der mittlern relativen Ereigniszahl vorausberechnet werden kann. Nun muss in dem wirklichen Geschehen, wo zu den rein zufälligen Schwankungen nichtzufällige Schwankungen der Anlage hinzukommen, das Schwankungsmass, d. h. die mittlere Abweichung, grösser ausfallen als in dem normal-zufälligen Geschehen. In wirklichen Fällen wird also die unter *b)* bzw. unter *f)* vorgeschriebene unmittelbare Berechnung der mittlern Abweichung im allgemeinen grössere Resultate liefern als die unter *e)* erläuterte theoretische Berechnung.

[1]) Vgl. *Lexis*, S. 177.

Man vergleicht die beiden Werte gewöhnlich durch Quotientenbildung. Die unmittelbar berechnete mittlere Abweichung [vgl. unter *b)* und unter *f)*] wird durch die theoretisch berechnete [vgl. unter *e)*] dividiert. Der entstehende Quotient heisst in der theoretisch-statistischen Literatur der Divergenzkoeffizient[1]).

Mit den Bezeichnungen, die im vorangehenden Abschnitt unter *a)* bis *f)* erläutert worden sind, schreibt sich der

$$\text{Divergenzkoeffizient} = \sqrt{\frac{g\,(\delta_1^2 + \delta_2^2 + \delta_3^2 + \cdots + \delta_n^2)}{V\,(1000 - V)\,(n - 1)}}$$

Der Divergenzkoeffizient ist ein Bruch, dessen Zähler den Nenner um so mehr übertrifft, je mehr die Schwankungen der Anlage neben den rein zufälligen Schwankungen in den Vordergrund treten. Der Divergenzkoeffizient wird also um so grösser, je mehr das untersuchte Geschehen von dem normal-zufälligen abweicht, *divergiert.* Wenn die Schwankungen in der Hauptsache zufällig sind, d. h. von unregelmässig wirkenden, von Fall zu Fall variierenden Ursachen herrühren, so wird der Divergenzkoeffizient um Eins herum sein. Je mehr die Änderungen der Anlage die rein zufälligen Schwankungen überwiegen, um so grösser wird im allgemeinen der Divergenzkoeffizient ausfallen. Das heisst das Vorhandensein von Ursachen, die in allen Fällen einer statistischen Serie in dem ganzen untersuchten Gebiet gleichmässig wirken, die aber von Serie zu Serie, von Jahr zu Jahr sich ändern, wird sich an der Grösse des Divergenzquotienten erkennen lassen. Wenn die Grösse des Divergenzquotienten ein Anzeichen für die Veränderung der Anlagen ist, so zeigt seine Kleinheit ihre Unveränderlichkeit, ihre Stabilität an. Der Divergenzkoeffizient kann also auch zur Feststellung der grössern und geringern Grade von statistischer Stabilität benutzt werden. Genau genommen liefert uns der Divergenzkoeffizient kein absolutes Mass der Stabilität. Der Divergenzkoeffizient kommt, wie schon gesagt, der Einheit nahe, wenn die Schwankung der Anlage im Vergleich zu den rein zufälligen Schwankungen klein ist. Der Divergenzkoeffizient wird gross in dem umgekehrten Falle. Was also der Divergenzkoeffizient durch seine Grösse und Kleinheit anzeigt, ist nicht die Kleinheit und Grösse der Stabilität an sich genommen, sondern im Verhältnis zur normal-zufälligen Schwankungskomponente. Bei grossen Grundzahlen ist die rein zufällige Schwankung statistischer Verhältniszahlen geringer als bei mässigen (vgl. unter *d)* im vorangehenden Abschnitt). Bei grossen Grundzahlen kommen somit die Veränderungen der Anlage

[1]) Die Bezeichnung „Divergenz“ ist auf Dormoy zurückzuführen. Vgl. *Lexis*, S. 130. Für die Sache selber vgl. die beiden Abhandlungen S. 130—212, insbesondere S. 183, ferner S. 227—232.

leichter zum Ausdruck als bei mässigen. (Das ist doch eben der Grund, warum man möglichst grosse Grundzahlen zur Feststellung solcher Veränderungen heranziehen muss.) Dieses leichtere und schwierigere Hervortreten der Änderungen der Anlage lässt uns voraussehen, dass bei grossen Grundzahlen der Divergenzkoeffizient im allgemeinen grösser ausfällt als bei kleinen. Wesentlich kleiner als Eins kann hingegen der Divergenzkoeffizient nie ausfallen, mindestens nie bei unverbundenen Massenerscheinungen, die für alle in der Grundzahl aufgezählten Individuen gleichmässig in Betracht kommen [1]).

Man kann die Sache noch etwas anders formulieren. Wenn die Ereigniszahlen der verschiedenen Serien den Gesetzen des normal zufälligen Geschehens entsprechend um ihren mittlern Wert ausgestreut sind, so ist der Divergenzkoeffizient Eins, und man spricht von *normaler Dispersion.* In Wirklichkeit sind die Ereigniszahlen, den Änderungen der Anlage entsprechend, *übernormal* dispergiert — eben diesen Angelpunkt der ganzen Theorie haben wir an dem Galtonschen Brett veranschaulicht — und der Divergenzquotient ist grösser als Eins. *Unternormale* Dispersion gibt es nicht bei homogenen, unverbundenen Massenerscheinungen. Das heisst bei solchen kann der Divergenzkoeffizient nie, oder mindestens nicht andauernd, unter Eins sinken.

IV. Beispiele und Schlusswort.

Die allgemeinen Aussagen des vorangehenden Abschnitts müssen natürlich an ausgedehntem statistischen Material verifiziert werden. Das ist durch Lexis und seine Schüler auf sehr befriedigende Weise geschehen. Die Lexissche Dispersionstheorie, deren Grundzüge ich hier dargelegt habe, ist heute durch eine ausgedehnte Erfahrung gestützt. Ich will hier nur einige Stichproben machen, die nur zum Beispiel und zur Veranschaulichung dienen sollen. Ich entnehme mein Material der schweizerischen Geburtsstatistik der Jahre 1871—1915 [2]). Ich habe drei statistische Verhältnisgrössen betrachtet: die Knabenquote unter sämtlichen Neugebornen (Tabelle I), die Knabenquote unter den unehelich Gebornen (Tabelle II), und die Quote der unehelich Gebornen (Tabelle III). Die hauptsächlichsten Ergebnisse dieser Tabellen sind in Tabelle IV rekapituliert.

Die Ausführung der nötigen Rechnungen habe ich an der Tabelle I eingehend erläutert. Die Behandlung der Tabellen II und III ist identisch. Ich habe die Tabellen auch äusserlich ähnlich angelegt, bis auf die Weglassung einiger Spalten, die nur eine Wiederholung vorangehender gewesen wären. Die richtige Stelle der weggelassenen Spalten ist an der Numerierung der übriggebliebenen ersichtlich. Bevor der Leser die Theorie auf ein neues Beispiel anwendet, kann er mit Vorteil die Handhabung der Begriffe an diesen Tabellen mittels einiger Stichproben einüben. Sie sind eben zu diesem Zwecke so ausführlich gehalten.

In der Tabelle IV vergleiche ich zunächst die Spalten (4) und (5). Wir finden, dass die den Gesetzen des normal-zufälligen Geschehens entnommene Relation zwischen der durchschnittlichen und der mittlern Abweichung, die unter *b)* im Abschnitt II erläutert wurde, auch bei unsern statistischen Grössen in ziemlicher Annäherung zutrifft. Die Zahlen der Spalte (7) sind kleiner als die der Spalte (6), und dem entsprechend sind alle drei in der Spalte (8) verzeichneten Divergenzkoeffizienten grösser als Eins. Dies ist eine Bestätigung der Hauptthese der Lexisschen Dispersionstheorie, wonach in dem wirklichen Geschehen die Schwankungen immer grösser sind als in dem normal-zufälligen.

Zu den einzelnen untersuchten Grössen ist folgendes zu bemerken: die Knabenquote unter den Neugebornen zeigt den kleinsten Wert des Divergenzkoeffizienten: 1.10, also eine von der normalen kaum unterscheidbare Dispersion. Dieses Verhalten der Knabenquote hat *Lexis* an einem sehr grossen Zahlenmaterial, den verschiedensten Ländern entnommen, nachgewiesen. Danach sind die Schwankungen der Knabenquote als vorwiegend zufällig zu betrachten. Dieser Ausspruch bedeutet nichts Unfassbares, sondern nur folgendes: an den Knabenquoten, die die 45 betrachteten sukzessiven Jahrgänge ergeben haben, ist es kaum möglich, eine Veränderung der Anlage, d. h. eine deutliche Entwicklung nachzuweisen [1]).

[1]) Auf die hier angedeutete wichtige Frage der *Homogeneität* der betrachteten Serien können wir nicht näher eingehen. Ich verweise für diese Frage, die die Grenzen der Anwendbarkeit der *Lexisschen* Theorie betrifft, auf *Lexis*, S. 227—232, und *L. v. Bortkiewicz,* Homogeneität und Stabilität in der Statistik (Skandinavisk Aktuarietidskrift, 1918, Heft 1—2). Auch die wichtige Frage der Schwankungsgrenzen des Divergenzkoeffizienten kann hier nicht behandelt werden. Vgl. darüber *L. v. Bortkiewicz,* Der mittlere Fehler des zum Quadrat erhobenen Divergenzquotienten (Jahresbericht der Deutschen Mathematikervereinigung, Bd. 27, 1918, S. 71—126). Beide Gegenstände sollen auch in einem Lehrbuch der Wahrscheinlichkeitsrechnung behandelt werden, das der Verfasser vorbereitet.

[2]) Vgl. Schweizerische Statistik, die Lieferungen 112, 170 und S. 4 der Lieferung 205.

[1]) Danach muss die „anhaltende verhältnismässige Abnahme der Knabengeburten“ in der Periode 1871—1908 (vgl. Schweizerische Statistik, Lieferung 170, Seite 54*—55*) zum grössten Teil als ein Werk des Zufalls und nicht als eine auf der Veränderung der Lebensbedingungen beruhende Evolution betrachtet werden. In der Tat, schon die nächsten Jahre 1909—1915 bringen einen Umschlag. Es wird a. a. O., S. 56*, über die Knabenquoten in den verschiedenen grössern Kantonen gesagt: „sie sind unter sich zu wenig verschieden, als dass bestimmte Ursachen und Einflüsse angenommen werden könnten.“ Ähnliches sollte man, nach der Berechnung des Divergenzkoeffizienten, über die Knabenquoten der verschiedenen *Jahre* sagen.

Die Knabenquote unter den unehelichen Geburten scheint etwas weniger stabil zu sein. Trotz der mässigen Grundzahl ist der Divergenzquotient grösser als vorher, nämlich 1.22. Ob die Anlage zur Knabenquote in den unehelichen Geburten wirklich stärker veränderlich ist, müsste durch eine ausgedehnte Untersuchung, insbesondere durch Vergleich von Daten aus verschiedenen Ländern entschieden werden. Dass aber die Entscheidung so ausfällt, wie unsre Stichprobe andeutet, nämlich, dass die Knabenquote unter den unehelichen Geburten stärker veränderlich ist, würde sehr gut zu einer Ansicht stimmen, die *Lexis* gelegentlich geäussert hat[1]). Ich will es kurz ausführen, wieso. Das tue ich nicht so sehr wegen der Sache selbst, als um zu zeigen, wie man überhaupt sich die „Veränderung der Anlage“ bei der Knabenquote, diesem klassischen Beispiel der statistischen Stabilität, zu denken hat, und zu welchen Überlegungen solche Beobachtungen über die Grösse von Divergenzkoeffizienten Anlass geben können.

Gewisse antiquierte Hypothesen darf man wohl heutzutage nach blosser Erwähnung übergehen. Zum Beispiel, dass das Geschlecht des Kindes sich nach der ältern Hälfte des Ehepaares richtet, oder nach derjenigen Hälfte, die bei der Zeugung „an Körper- und Willenskraft überlegen ist“. Es ist viel wichtiger, etwas von Versuchen an verschiedenen Tieren und Pflanzen zu wissen, bei denen durch tiefgehende Eingriffe in die Lebensbedingungen und in den Befruchtungsvorgang das zahlenmässige Verhältnis zwischen der Nachkommenschaft beiderlei Geschlechts beträchtlich und systematisch verschoben werden konnte[2]). Aber auch ohne so weit herauszuholen, kann man sich die Sache an geläufigem statistischem Material klarmachen. Geläufig ist, dass die Sterblichkeit der Knaben im ersten Lebensmonat bedeutend grösser ist als die der Mädchen. Dasselbe muss der Fall in dem letzten Stadium des embryonalen Lebens sein, weil die Knaben unter den Totgebornen in beträchtlich grösserm Prozentsatz auftreten als unter den Lebendgeborenen. Es ist anzunehmen, dass die grössere Sterblichkeit der Knaben sich auch auf die frühern Stadien des embryonalen Lebens erstreckt und dass etwa durch Früh- und Fehlgeburten vor dem normalen Fälligkeitstermin relativ mehr Knaben- wie Mädchenleben verloren gehen. Irgendein Einfluss, der die Kindersterblichkeit in dem frühern embryonalen Stadium erhöht, wird also, wie anzunehmen ist, relativ mehr Opfer an Knaben- wie Mädchenleben fordern, und wird folglich das normale Verhältnis der Geschlechter unter den in der richtigen Zeit zur Geburt gelangenden Embryonen verschieben, und zwar zugunsten der Mädchengeburten. Die Statistik der unehelichen Geburten ist für diese Überlegung wie eine Probe auf das Exempel. Unter den unehelichen Neugebornen gibt es relativ mehr Totgeborne als unter den ehelichen, was aus der Situation der Mütter ohne weiteres verständlich ist, und relativ weniger Knaben. In unserm Material z. B. gab es 513.8 ‰ Knaben unter allen, und nur 508.7 ‰ Knaben unter den unehelich Gebornen. Dieser Unterschied könnte also auch auf die Rechnung der frühembryonalen Sterblichkeit gesetzt werden, die unter unehelichen Kindern sicher grösser ist.

Wir lernen aus dieser Betrachtung zweierlei: Wenn auch die Anlage der Menschenrasse, Knaben oder Mädchen zu *zeugen*, im Laufe der Zeiten für normale Verhältnisse ganz unverändert geblieben wäre, die Anlage, Knaben oder Mädchen zu *gebären*, ändert sich mit jeder Lebensbedingung, die die frühembryonale Sterblichkeit beeinflusst. Diese Lebensbedingungen bestehen, wenn es sich um aussereheliche Geburten handelt, zum Teil auch aus sozialen Anschauungen, also aus etwas leicht Veränderlichem, und so wäre ganz gut möglich, dass es allgemein zutrifft, was wir in unserm Fall beobachtet haben, dass die Dispersion der unehelichen Knabenquote das normale Mass mehr übertrifft als die der allgemeinen Knabenquote. Die Entscheidung muss natürlich einer eingehenden Untersuchung überlassen werden.

Noch einige Worte über unser letztes Beispiel. Der Prozentsatz der Unehelichen unter den Geborenen ist selbstverständlich stärker variabel als der Prozentsatz der Knaben, da doch eigentliche soziale Anlagen im allgemeinen wohl stärker variieren als überwiegend physiologische. Das kommt dadurch zum Ausdruck, dass die Dispersion stark übernormal ist: der Dispersionskoeffizient ist viel grösser als Eins, nämlich 3.78. Man könnte in diesem Falle auch ohne das Heranziehen des Divergenzkoeffizienten einsehen, dass die Anlagen sich verändert haben: in den 22 ersten Beobachtungsjahren liegen 16 Zahlen über der mittlern Quote der Periode, in den 22 letzten Beobachtungsjahren nur vier. Man sieht daraus ohne jede Theorie, dass die Quote der unehelich Gebornen im Laufe der Zeit im allgemeinen abgenommen hat. Dass überhaupt eine Veränderung stattgefunden hat, war schon aus der Grösse des Divergenzkoeffizienten zu ersehen, und so dient auch diese Beobachtung der Lexisschen Theorie zur Bestätigung.

Werfen wir einen Blick auf den zurückgelegten Weg. Wir haben in Berechnung der Divergenzkoeffi-

[1]) Vgl. *Lexis*, S. 164—169.

[2]) Vgl. etwa das Referat von *C. Correns*, Die Konkurrenz der Keimzellen und das Geschlechtsverhältnis (Die Naturwissenschaften, 6. Jahrgang, 1918, S. 277—280).

Anmerkungen.

Tabellen I, II, III. Die letzte Zeile gibt die Summe der vorangehenden 45 Zeilen in den Spalten (2), (3), (6), die Summe der Beträge in der Spalte (5) und die mittlere relative Ereigniszahl der Periode in der Spalte (4).

Tabelle I.

Jahr	Geborne insgesamt	Darunter Knaben		Abweichung von 513.8 ‰	Quadrat der Abweichung
		absolut	in ‰		
(1)	(2)	(3)	(4)	(5)	(6)
1915	77 931	39 899	512.0	— 1.8	3.24
1914	90 128	46 555	516.5	+ 2.7	7.29
1913	92 603	47 634	514.4	+ 0.6	0.36
1912	95 171	48 979	514.6	+ 0.8	0.64
1911	94 185	48 191	511.7	— 2.1	4.41
1910	96 669	49 667	513.8	0.0	0.00
1909	97 296	49 772	511.6	— 2.2	4.84
1908	99 468	50 994	512.7	— 1.1	1.21
1907	97 696	50 272	514.6	+ 0.8	0.64
1906	98 971	50 861	513.9	+ 0.1	0.01
1905	98 057	50 287	512.8	— 1.0	1.00
1904	98 300	50 365	512.4	— 1.4	1.96
1903	97 119	49 709	511.8	— 2.0	4.00
1902	99 993	51 474	514.7	+ 0.9	0.81
1901	100 635	51 390	510.7	— 3.1	9.61
1900	97 695	50 148	513.3	— 0.5	0.25
1899	97 894	50 223	513.0	— 0.8	0.64
1898	95 184	48 494	509.5	— 4.3	18.49
1897	93 369	47 821	512.2	— 1.6	2.56
1896	91 674	47 029	513.0	— 0.8	0.64
1895	88 184	45 668	517.9	+ 4.1	16.81
1894	87 317	44 665	511.4	— 2.4	5.76
1893	88 100	45 262	513.8	0.0	0.00
1892	86 265	44 289	513.4	— 0.4	0.16
1891	86 721	44 559	513.8	0.0	0.00
1890	81 620	41 867	513.0	— 0.8	0.64
1889	84 279	43 205	512.6	— 1.2	1.44
1888	84 444	43 515	515.3	+ 1.5	2.25
1887	84 661	43 404	512.8	— 1.0	1.00
1886	84 142	43 073	511.9	— 1.9	3.61
1885	83 579	42 908	513.4	— 0.4	0.16
1884	84 794	43 644	514.7	+ 0.9	0.81
1883	85 197	43 975	516.2	+ 2.4	5.76
1882	85 987	44 319	515.4	+ 1.6	2.56
1881	88 503	45 444	513.3	— 0.5	0.25
1880	87 413	45 037	515.2	+ 1.4	1.96
1879	89 692	46 213	515.2	+ 1.4	1.96
1878	91 426	46 959	513.6	— 0.2	0.04
1877	92 861	48 000	516.9	+ 3.1	9.61
1876	94 595	48 924	517.2	+ 3.4	11.56
1875	91 806	47 102	513.1	— 0.7	0.49
1874	86 918	44 657	513.8	0.0	0.00
1873	84 495	43 629	516.3	+ 2.5	6.25
1872	84 313	43 567	516.7	+ 2.9	8.41
1871	81 629	42 071	515.4	+ 1.6	2.56
Total	4 078 979	2 095 720	513.8	64.9	146.05

Geburtenstatistik 1871—1915.

Anmerkungen.

Tabelle II. Die fehlende Spalte (1) ist identisch mit der Spalte (1) der Tabelle I.
Tabelle III. Für die fehlende Spalte (1) vgl. Tabelle I, Spalte (1)
" " " " (2) " " I, " (2)
" " " " (3) " " II, " (2).

Tabelle II.

Unehelich Geborne	Darunter Knaben absolut	Darunter Knaben in ‰	Abweichung von 508.7 ‰	Quadrat der Abweichung
(2)	(3)	(4)	(5)	(6)
3 627	1 874	516.7	+ 8.0	64.00
4 556	2 394	525.5	+ 16.8	282.24
4 390	2 209	503.2	— 5.5	30.25
4 592	2 354	512.6	+ 3.9	15.21
4 372	2 263	517.6	+ 8.9	79.21
4 417	2 217	501.9	— 6.8	46.24
4 456	2 289	513.5	+ 4.8	23.04
4 558	2 287	501.8	— 6.9	47.61
4 382	2 251	513.5	+ 4.8	23.04
4 364	2 200	504.1	— 4.6	21.16
4 469	2 276	509.3	+ 0.6	0.36
4 215	2 112	501.1	— 7.6	57.76
4 188	2 102	501.9	— 6.8	46.24
4 422	2 267	512.7	+ 4.0	16.00
4 465	2 265	507.3	— 1.4	1.96
4 463	2 278	510.4	+ 1.7	2.89
4 518	2 306	510.4	+ 1.7	2.89
4 363	2 184	500.6	— 8.1	65.61
4 424	2 317	523.7	+ 15.0	225.00
4 318	2 235	517.6	+ 8.9	79.21
4 009	2 095	522.6	+ 13.9	193.21
4 107	2 079	506.2	— 2.5	6.25
4 114	2 090	508.0	— 0.7	0.49
4 148	2 171	523.4	+ 14.7	216.09
4 066	2 106	517.9	+ 9.2	84.64
3 855	1 947	505.1	— 3.6	12.96
3 923	1 985	506.0	— 2.7	7.29
4 061	2 069	509.5	+ 0.8	0.64
4 048	1 997	493.3	— 15.4	237.16
4 158	2 136	513.7	+ 5.0	25.00
4 191	2 149	512.8	+ 4.1	16.81
4 222	2 172	514.4	+ 5.7	32.49
4 226	2 068	489.4	— 19.3	372.49
4 282	2 190	511.4	+ 2.7	7.29
4 279	2 202	514.6	+ 5.9	34.81
4 121	2 111	512.3	+ 3.6	12.96
4 157	2 138	514.3	+ 5.6	31.36
4 381	2 195	501.0	— 7.7	59.29
4 573	2 348	513.4	+ 4.7	22.09
4 771	2 332	488.8	— 19.9	396.01
4 084	2 106	515.7	+ 7.0	49.00
4 190	2 125	507.5	— 1.5	2.25
4 323	2 156	498.7	— 10.0	100.00
4 377	2 108	481.6	— 27.1	734.41
4 643	2 348	505.7	— 3.0	9.00
192 868	98 103	508.7	323.1	3 793.91

Tabelle III.

Uneheliche auf 1000 Geborne	Abweichung von 47.3 ‰	Quadrat der Abweichung
(4)	(5)	(6)
46.5	— 0.8	0.64
50.6	+ 3.3	10.89
47.4	+ 0.1	0.01
48.4	+ 1.1	1.21
46.4	— 0.9	0.81
45.7	— 1.6	2.56
45.8	— 1.5	2.25
45.8	— 1.5	2.25
44.8	— 2.5	6.25
44.1	— 3.2	10.24
45.6	— 1.7	2.89
42.9	— 4.4	19.36
43.1	— 4.2	17.64
44.2	— 3.1	9.61
44.6	— 2.7	7.29
45.7	— 1.6	2.56
46.2	— 1.1	1.21
45.8	— 1.5	2.25
47.4	+ 0.1	0.01
47.1	— 0.2	0.04
45.5	— 1.8	3.24
47.0	— 0.3	0.09
46.7	— 0.6	0.36
48.1	+ 0.8	0.64
46.9	— 0.4	0.16
47.2	— 0.1	0.01
46.5	— 0.8	0.64
48.1	+ 0.8	0.64
47.8	+ 0.5	0.25
49.4	+ 2.1	4.41
50.1	+ 2.8	7.84
49.8	+ 2.5	6.25
49.6	+ 2.3	5.29
49.8	+ 2.5	6.25
48.3	+ 1.0	1.00
47.1	— 0.2	0.04
46.3	— 1.0	1.00
47.9	+ 0.6	0.36
49.2	+ 1.9	3.61
50.4	+ 3.1	9.61
44.5	— 2.8	7.84
48.2	+ 0.9	0.81
51.2	+ 3.9	15.21
51.9	+ 4.6	21.16
56.8	+ 9.5	90.25
47.3	84.9	286.93

Tabelle IV.

Nummer der Tabelle	Mittlere Grundzahl	Mittlere relative Ereigniszahl in ‰	$1.253 \times$ durchschnittliche Abweichung	Mittlere Abweichung berechnet			Divergenz-koeffizient
				unmittelbar nach *b)*	unmittelbar nach *f)*	theoretisch aus Spalten (2) (3) nach *e)*	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
I	90 644	513.8	1.81	1.80	1.82	1.66	1.10
II	4 286	508.7	9.00	9.18	9.28	7.63	1.22
III	90 644	47.3	2.36	2.52	2.55	0.67	3.78

Der Begriff und die Berechnung der vorkommenden Grössen ist im Abschnitt II erklärt.

zienten ein wichtiges Mittel zur Lösung der vornehmsten Aufgabe der Statistik gewonnen. Diese besteht nämlich darin, aus den registrierten Zahlen die Änderungen der sozialen Anlagen festzustellen. Der Divergenzkoeffizient gibt zwar kein absolutes Kriterium, aber auf alle Fälle einen zahlenmässigen, exakten Anhaltspunkt, um mindestens über das Vorhandensein oder Nichtvorhandensein von solchen Veränderungen der Grundbedingungen urteilen zu können. Wir haben auch eine gewisse anschauliche Einsicht in die Schwankungen der sozialen Massenerscheinungen gewonnen durch unmittelbaren Vergleich mit den Schwankungen physikalischer Massenerscheinungen, wie sie an dem Galtonschen Brett vor sich gehen. Ich glaube, dass dieser Vergleich wertvoll ist, weil er in sehr weitgehende Einzelheiten zutrifft. Ich bemerke noch, dass die prinzipielle Seite von allen wichtigen Anwendungen der Wahrscheinlichkeitsrechnung sich ebenfalls durch einen unmittelbaren Vergleich mit dem Galtonschen Brett beleuchten lässt, vielleicht einsichtiger und treffender als durch blosse Formeln, die letzten Endes nur Mathematiker überzeugen und nur Mathematikern zugänglich sind.