Logik der Statistik

Von Dr. *Arnold Schwarz*, Bern

Inhalt:

Die folgenden Ausführungen haben einen doppelten Zweck: den alten, reichen Erfahrungsschatz der Logik als wissenschaftliche Denklehre auf ein Gebiet zu übertragen, das noch ein undurchdringlicher Urwald von Lehrmeinungen bedeckt; also für die Statistik in bescheidenem Rahmen das zu leisten, was die Untersuchungen eines Wundt[1]) für die exakten und Geisteswissenschaften im gesamten geleistet haben.

Aus den theoretischen Darlegungen[2]) sollen sich aber auch praktische Anwendungen ergeben. Die starke Beeinflussung der sogenannten statistischen Praxis durch methodische Untersuchungen wird heute niemand mehr leugnen wollen. Es ist daher ein Gebot der Praxis, Theorie zu treiben.

I. Die statistische Begriffsbildung

Wenn nicht klar ist, was man zählen will, kann man nicht zählen[3]). Das liegt auf der Hand. Mit Recht hat daher Zizek die scharfe begriffliche Umgrenzung der Zähleinheit als Grundlage jeder statistischen Auszählung gefordert. Ist diese Forderung erfüllbar? Und erschöpft sich die statistische Begriffsbildung im Aufstellen der Zähleinheit? Das kann schon deswegen nicht sein, weil diese oft genug ein Provisorium ist. Erst durch die Erhebung wird festgestellt, was die Zähleinheit inhaltlich bedeutet, und oft wird noch nach der Zählung ihr Umfang eingeschränkt.

Die Besonderheiten der statistischen Begriffsbildung können am besten an dem Beispiel des Betriebsbegriffs gezeigt werden, der allen Statistikern, welche

[1]) Logik, Bd. II und III, Stuttgart 1907 und 1908.

[2]) Untersuchungen über die Zusammenhänge der Wahrscheinlichkeitstheorie mit der Statistik sind hier nur gestreift worden, weil ich in einem andern Aufsatz («Philosophie der Statistik», Allgemeines Statistisches Archiv 1931, Heft 2) die erkenntnistheoretischen Grundlagen der Statistik zu skizzieren gesucht habe. Diese Schrift möchte ich hier zwar nicht als bekannt voraussetzen. Ohne ihre Kenntnis muss aber das Folgende recht fragmentarisch erscheinen. — Untersuchungen über die Ursachenforschung in der Statistik bilden ein Kapitel der statistischen Methodik, das vielleicht bei anderer Gelegenheit folgen wird.

[3]) Zizek, Grundriss der Statistik, München 1921, S. 62: «Es muss klar sein, wer zu zählen ist, bzw. welche Einzelfälle in die Erhebung einzubeziehen sind.»

eine Betriebszählung durchzuführen haben, stets so grosse Schwierigkeiten bereitet. Offenbar deckt sich der Sprachgebrauch keineswegs mit den Bedürfnissen des Statistikers. Auch die nationalökonomischen und juristischen Begriffsbestimmungen des Betriebes [1]), so z. B. Sombarts Definition als einer «Veranstaltung zum Zwecke fortgesetzter Werkverrichtung», reichen bei weitem nicht hin, um die Durchführung einer Betriebszählung zu ermöglichen. Der Statistiker ist neuerdings zu dem Schluss gekommen, es sei am besten, den Betrieb überhaupt nicht zu definieren. Er begnügt sich damit, eine ganze Anzahl von wirtschaftlichen Kategorien aufzuführen, welche in die Zählung einzubeziehen sind. Am weitesten in der Aufzählung des Inhalts des Betriebsbegriffes ist die Österreichische Betriebszählung von 1902 gegangen, welche dem Zähler eine umfangreiche Druckschrift in die Hand gab, die in alphabetischer Anordnung eine grosse Zahl von Betriebsarten und Betriebsbenennungen enthielt und dazu dienen sollte, im Zweifelsfalle nachgeschlagen zu werden, ob ein vorliegendes wirtschaftliches Gebilde unter den Begriff des Betriebes falle und in die Zählung einzubeziehen sei. Die Logiker bezeichnen diese Art der Definition als die primitivste. Sie nennen sie bekanntlich «definitio per enumerationem simplicem».

Ergibt sich schon aus diesen wenigen Andeutungen, dass die strenge Umgrenzung eines statistischen Begriffes, der zur Zählung dienen soll, oft gar nicht möglich ist, so zeigt die Erfahrung andererseits, dass zur Verarbeitung der Zählung ein so umfassender Begriff wie der des Betriebes gar nicht eindeutig festgelegt sein muss.

Die Lehre vom Begriff hat in der jahrtausendealten Geschichte der Logik eine merkwürdige Wandlung erfahren. In der Logik des Aristoteles beruht das syllogistische Verfahren durchweg auf dem Vorhandensein von streng umrissenen allgemeingültigen Begriffen. Durch Zusammenfügen von Begriffen, so lehrte man noch in der scholastischen Logik und bis in die Neuzeit hinein, entstehen die Urteile, durch Zusammenfügen der Urteile die Schlüsse[2]). Dass aber der Begriff ein Niederschlag aus Urteilen ist, etwas Wandlungsfähiges und Fliessendes, ein Gefäss, das sich stets wieder mit anderm Inhalt füllt, dass er eine Abkürzung ist, ein Namensschild für eine ganze Reihe von komplizierten, miteinander verbundenen Vorstellungen, die in uns zum Anklingen kommen und mehr oder weniger deutlich im Lichtkegel des Bewusstseins auftauchen, diese Anschauung ist verhältnismässig jung. Aus ihr folgt unmittelbar die Unmöglichkeit, gewisse Begriffe überhaupt einwandfrei zu definieren[3]). Nicht nur die Grenzen der Begriffe sind fliessend, was man schon oft bemerkt hatte, sondern auch ihr Kern: je nach Umständen und Zwecken muss er wieder in anderm Licht erscheinen.

Wir erkennen Dinge nur durch Merkmale (Kant)[4]). Und wir greifen aus der Masse der Merkmale eines Begriffs nur jene heraus, die für uns wichtig sind. Das häufigste Objekt der Statistik, der Mensch, «ist ein komplexer Gegenstand. Doch

---

[1]) Passow, Betrieb, Unternehmung, Konzern, Jena 1925, gibt eine reichhaltige Übersicht.

[2]) Donat, S. J., Logica, Innsbruck 1914, gibt in leicht lesbarem Latein eine ausgezeichnete Darstellung der scholastischen Logik.

[3]) Besonders schön ausgeführt bei v. Kries, Logik, Tübingen 1916, S. 563 ff.

[4]) Logik, Leipzig 1904, S. 64.

aus der Fülle der Merkmale hebt ein Armeelieferant seine Eigenschaft heraus, so viele Pfund im Tag zu essen; ein General, so viele Meilen zu marschieren; der Stuhlfabrikant, eine solche Form zu haben; der Redner, auf die und die Gefühle zu reagieren; der Theaterdirektor, den und den Preis, und nicht mehr, für einen unterhaltenden Abend zu bezahlen[1].» Der Statistiker sollte sich von der Auffassung freimachen, einen Begriff immer nur starr in einem einzigen Sinn zu brauchen. «Stadt» und «Land», «Gross-» und «Kleinstadt», «Gross-» und «Kleinbetrieb», «Handwerk» und «Industrie» erfordern, je nach der Untersuchung und ihrem Zweck, wechselnde quantitative Abgrenzungen[2].

Einen wirklichen Inhalt bekommen die statistischen Begriffe wie «Fabriken», «Ausländer», «Lohneinkommen» natürlich erst durch den statistischen Urteilsprozess, dessen Niederschlag sie sind und den wir im folgenden Abschnitt näher betrachten wollen.

II. Das statistische Urteil

ist eine besondere, noch nicht beschriebene Spezies der umfangreichen Gattung der Urteile, die in den dickleibigen Erbauungsbüchern des Verstandes, den Handbüchern der Logik, aufgezählt werden. Die Lehre vom Urteil steht seit langem im Mittelpunkt der logischen Theorie. Sie wird als das Quellgebiet der Logik bezeichnet[3]. Merkwürdigerweise ist es aber noch keineswegs gelungen, dieses Quellgebiet zu erforschen und die Geheimnisse des Urteilsprozesses aufzuklären. Auf die einfache Frage: «Was ist ein Urteil?» weiss kein Logiker eine Antwort[4].

Nur dies hat man erkannt, dass die früheren Erklärungen des Urteils im Grunde gar keine sind, so vor allem die reinen Worterklärungen, dass ein Urteil eine Aussage (Aristoteles)[5], eine Verbindung von Subjekt und Prädikat[6], die Behauptung einer Beziehung (Windelband)[7] sei. Höchst verschiedene Denkformen nehmen die äussere Form des Urteils an. Deshalb wird es wohl kaum je gelingen, eine Einheitsformel für «das» Urteil, eine befriedigende Definition zu finden.

Jedenfalls brauchen wir nicht auf sie zu warten — so wenig ungezählte Generationen auf sie gewartet, aber doch fleissig Urteile gefällt haben. Für unsere Absicht wird es genügen, den eigentlichen Zweck des Urteilsprozesses klar vor Augen zu haben. Das Urteil dient der Begriffsbildung, der Anreicherung und Klärung der Begriffe. Und reiche, richtige Begriffe sind das Ziel jeder Forschung. Durch das Urteil nun wird einem Begriff ein Merkmal zugesprochen oder abgesprochen. Dem Subjektsbegriff S wird der Prädikatsbegriff P zugeordnet («S ist P»), oder er wird als unvereinbar mit ihm erklärt («S ist nicht P»).

---

[1]) William James, Psychology, New York 1890, II, S. 332.

[2]) Man wird einwenden, dann seien die abgegrenzten Massen nicht homogen. Das sind sie ohnedies nie. Ihre Gleichartigkeit kann sich immer nur auf wenige gleiche Merkmale beschränken. Das macht uns blind und «ungerecht» (James) gegen die andern, oft ebenso wichtigen und meistens nicht gleichmässig vertretenen Merkmale.

[3]) Lask, Die Lehre vom Urteil, Tübingen 1912.

[4]) Stöhr, Die Vieldeutigkeit des Urteils, Leipzig 1895, S. 69: «Es wird... vergeblich sein, das Wesen „des“ Urteils ergründen zu wollen.»

[5]) I. Analytikon, übersetzt von Kirchmann, Leipzig 1877.

[6]) Welton, An intermediate Logic, London 1911, S. 113.

[7]) Die Prinzipien der Logik, in Enc. der phil. Wiss., Tübingen 1912, S. 23.

Hand in Hand mit diesem Vorgang geht ein meist völlig unbewusstes Prüfen der Begriffsumfänge von S und P, ob nämlich S und P wirklich aufeinander bezogen werden können, ob alle S P sind oder ob nur einige S P sind, ob keine S P seien oder ob einige S nicht P seien. So gelangt man zu den vier Arten quantitativ bestimmter kategorischer Urteile: «Alle S sind P», «einige S sind P», «kein S ist P», «einige S sind nicht P».

Die Subsumtionstheorie des Urteils ist heute noch die am meisten verbreitete[1]), obwohl sie immer wieder heftigen Angriffen ausgesetzt war. Es wird ihr entgegengehalten, sie sei einseitig und rein formal. Mit Recht hat Stöhr[2]) hervorgehoben, das Urteil z. B.: «Horch, die Lerche singt im Ätherblau», wolle keineswegs besagen: «Die Lerche fällt in den logischen Umfang dessen hinein, was jetzt im Ätherblau singt.» Sicherlich ist die Einordnung zweier Begriffe, die Untersuchung ihrer gegenseitigen Umfangslagerung, nicht das Wesentliche des Urteils. Aber sie muss jede Urteilsbildung begleiten. In der Tat beruht die ganze Syllogistik nur auf ihr, und alle Neuerungen, z. B. die Erdmannsche Immanenztheorie des Urteils, kommen bei der Darstellung der hergebrachten Schlussformen nicht um das Studium der Begriffsumfänge herum. Der neuerdings besonders gepflegte Zweig der Logik, die sogenannte Algorithmik, beruht auf nichts anderem als auf der Quantifikation des Urteils. Die quantitative Seite ist nicht die Seite, wohl aber eine wichtige Seite der Urteilsbildung.

Dass die quantitative Seite des Urteils in der Statistik von grosser Bedeutung ist, lässt sich von Anfang an vermuten. Die Statistik hat in der Tat eine Form des Urteils ausgebaut, die bisher von der logischen Theorie stark vernachlässigt war: das sogenannte partikuläre Urteil (von der Form: «einige S sind P»). Wegen seiner Unbestimmtheit[3]) galt dieses Urteil in der logischen Theorie nie als voll, so dass überhaupt der Wert der sogenannten i-Schlüsse der Schullogik stark in Zweifel gezogen wurde. Hermann Cohen[4]) schreibt:

> «„Einige S" treten jetzt auf den Plan. Man sollte denken, die Logik mache sich selbst damit den Garaus; so augenfällig, so selbstverräterisch ist dieser Verfall in die unheilbare Unbestimmtheit. Aller Sinn für Bestimmtheit und Genauigkeit muss abgestumpft und abgestorben sein, wenn man in der Logik sich mit Einigen abspeisen lassen kann. Und diese Einige figurieren vollends noch unter der Fahne der Quantität, während sie doch das gerade Widerspiel derselben sind, und ernstlich nur eine quantité négligeable vorstellen.»

Es ist nun allerdings richtig, dass das partikuläre Urteil «einige S sind P» äusserst dehnbar ist; da es ebensogut «einige wenige S» als auch «nahezu alle S», ja sogar «alle S[5])», bedeuten kann; dass das Urteil «einige S sind nicht P» eventuell

---

1) Benno Erdmann, Logische Elementarlehre, II. Aufl., Halle 1907, gibt S. 343 ff. die beste Darstellung der Einwendungen; eine vorzügliche Übersicht der Urteilstheorien bei Case im Artikel «Logic» der Enc. brit. 11. Aufl. nebst Verteidigung der Subsumtionstheorie. Jerusalem (Lehrbuch der Psychologie, 4. Aufl. Wien 1907), S. 127, schreibt, die Prüfung der Umfangsverhältnisse sei viel geeigneter als jene der Inhaltsverhältnisse.

2) A. a. O., S. 63.

3) In der Logik versteht man gewöhnlich unter «unbestimmten» Urteilen nicht die partikulären, sondern solche von der Form: «Metalle sind nützlich.» Jevons, Leitfaden der Logik, übersetzt von Kleinpeter, Leipzig 1913, S. 66.

4) Logik der reinen Erkenntnis, 2. Aufl., Berlin 1914, S. 542.

5) Jevons, a. a. O., S. 67 f.

auch bedeuten kann «kein S ist P». Doch geht aus dem Zusammenhang meist mit ziemlicher Bestimmtheit hervor, wieviel ungefähr zum Begriffsumfang der S gehört. Der gewöhnliche Sprachgebrauch bereits drängt zu einer nähern Bestimmung, wie aus der Häufigkeit der Ausdrücke: «so gut wie alle», «ein grosser Teil», «sehr viele», «weit über die Hälfte», «nur wenige» usw. hervorgeht[1]).

Die Statistik ist aber noch weiter gegangen und hat sich des unbestimmten Urteils bemächtigt, um es für ihre Zwecke brauchbar zu gestalten, um es zu präzisieren. Aus dem unbestimmten Urteil «einige S sind P» wird das ganz bestimmte: «so und so viele S (von allen bekannten S) sind P». Der Umfang des Begriffes S gegenüber dem P ist dadurch zahlenmässig festgestellt[2]).

Es kann aber auch der Begriff P gegenüber dem S unbestimmt gelagert sein, wie dies im sogenannten divisiven Urteil: «S ist zum Teil P_1, zum Teil P_2» der Fall ist. Hier präzisiert der Statistiker[3]), indem er sagt: «S ist zu so und so vielen Teilen P_1 und zu so und so vielen Teilen P_2». Z. B. verwandelt sich das ganz wertlose unbestimmte Urteil: «Die Berufstätigen sind teilweise in der Landwirtschaft, teilweise in der Industrie tätig» in das statistische Urteil: «Die Berufstätigen sind zu 27 % in der Landwirtschaft, zu 45 % im Gewerbe beschäftigt[4]).» Diesen Vorgang nenne ich «Partition».

Die traditionelle Logik hat bis in die neueste Zeit hinein als wissenschaftlich brauchbare Urteile nur jene bezeichnet, die allgemein gültig, oder gar nur jene, die umkehrbar sind[5]). Wer auf allgemein gültige Beziehungen zwischen den S und P ausgeht, wer wissen will, welche Merkmale allen S zukommen, oder gar, welche Merkmale nur den S zugesprochen werden können, der wird auch mit dem präzisierten Urteil, wie wir es nennen wollen, «so und so viele S von allen S sind P», nicht viel anfangen können. Heute geht freilich ein wachsender Teil der wissenschaftlichen Erkenntnis auf statistische Urteile zurück. Man hat eingesehen, dass es bereits eine wesentliche Bereicherung unserer Erkenntnis bedeutet, zwar nicht zu wissen, dass einigen S das Merkmal P, wohl aber, wie vielen S dieses Merkmal zukommt. Es hat sich nämlich gezeigt, dass dieses Zahlenverhältnis in sehr vielen Fällen kein völlig schwankendes, einmaliges oder zufälliges ist, sondern unter gewissen Bedingungen als mehr oder weniger beständig oder doch als bezeichnend gelten darf.

Wenn wir uns einen Augenblick vergegenwärtigen, welche geistigen Vorgänge sich bei der Erfassung der Zahlen einer grössern statistischen Erhebung abspielen, wie der Konsument der Statistik auf sie reagiert, so bemerken wir deutlich, dass es ihm im Grunde nicht auf die Höhe der absoluten Zahlen ankommt, sondern vielmehr auf ein Abtasten von Grössenverhältnissen, genau von der Art, wie wir sie als Urform des statistischen Urteils hingestellt haben. Es handelt sich um ein Auswägen der Massen gegeneinander, ein Abschätzen ihrer Schrumpfungen oder

---

1) Erdmann, Logik S. 480.

2) Z. B. aus: «Viele Schweizer sind verheiratet» wird «$1{,}_2$ Millionen Schweizer (von $3{,}_5$ Millionen) sind verheiratet»; oder wenn S = 100 gesetzt: 35 % der Schweizer sind verheiratet.

3) Erdmann, a. a. O. S. 555: «Unser Denken drängt zur spezifischen Disjunktion.»

4) Es ist Quételets Verdienst, neben der Variationsbreite als erster ihre Frequenz studiert zu haben.

5) Welton and Monahan, An intermediate Logic, London 1911, S. 366.

Ausdehnungen seit früheren Erhebungen, ein Einteilen in Beziehungssysteme und Grössenklassen, alles mit dem Ziel, einen noch leeren Begriff, etwa den des Betriebes oder der Wohnbevölkerung, durch Anreicherung mit den verschiedensten Merkmalen mit Leben zu füllen.

Es darf uns nicht verwundern, dass diese Art der Vervollkommnung der partikulären Urteile in der Schullogik noch keinen Einzug gehalten, dass die neue Anschauungsweise auf einen kurzen Zeitabschnitt zurückzublicken hat. Erst in den 40er Jahren des vergangenen Jahrhunderts begann man über die zunehmende Zahlenmode der Statistik zu klagen, und bald begann in den Witzblättern der Statistiker als komische Figur aufzutauchen. Dass es früher schon Vorläufer gegeben hat, so z. B. einen der geistreichsten Nationalökonomen, den Abbé Galiani, der seine volkswirtschaftlichen Schlüsse rein aus der Kenntnis der Zahlenverhältnisse und der Sammlung von Tatsachen in den sogenannten Staatskalendern zog, sei natürlich nicht in Abrede gestellt.

Worin besteht die neue Anschauungsweise, die dazu berufen war, ähnlich wie der Entwicklungsgedanke einen andern Abschnitt in der Geistesgeschichte einzuleiten? Sie glaubt, wie gesagt, nicht mehr an die festen Verbindungen zwischen den Merkmalen der Begriffe, sie geht nicht mehr auf «unzerreissbare Zusammenhänge» aus[1]), worin bisher alles Forschen bestanden hatte. An Stelle der starren gedanklichen Verbindung in den Urteilen «alle S sind P» tritt jetzt das unbestimmte Urteil «S ist manchmal P», «S kann P sein» — aber was hier an Bestimmtheit verloren geht, gewinnt man wieder an Bestimmtheit durch die Feststellung, in wie vielen Fällen S das Merkmal P aufweist und wie stark vertreten die verschiedenen Unterarten von P gegenüber dem S sind.

Damit ist aber die Frage noch keineswegs beantwortet, warum ein solches Präzisieren des unbestimmten Urteils wertvoll sei. Zunächst kann es oft genügen zu wissen, ob in vorwiegendem Masse oder aber nur in verschwindendem Ausmass gewisse Merkmale an gewisse Erscheinungen gebunden sind. Auf Grund dieser Kenntnis allein schon sind weitere Schlüsse möglich. Dazu kommt aber noch ein anderes. Durch die quantitative Betrachtungsweise solcher Urteile: in wie vielen Fällen S P ist und in wie vielen Fällen nicht, wurde man dazu geleitet, zu studieren, wie oft, bei wiederholten Beobachtungen, Abweichungen von den festgestellten Häufigkeiten vorkamen. Man sammelte Erfahrungen über die Variationsbreite und -häufigkeit einer Erscheinung. Man achtete auf die Maxima und Minima, die auf sehr vielen Gebieten an sich schon von grösstem Interesse sind. Soweit aber die Betrachtungsweise in der Hauptsache stets auf die Masse gerichtet blieb, verschwand die Vorliebe für extreme, sonderbare Fälle und machte dem zahlenmässigen Studium des Typischen Platz.

Wo Massenerscheinungen als solche die Aufmerksamkeit des Forschers gefangen hielten, wie in der Soziologie, da musste das Präzisieren von vagen, nur gefühlsmässig bestimmten Urteilen von ganz besonderem Werte sein.

Aber nicht nur das zahlenmässige Präzisieren des Subjekts-, sondern auch des Prädikatbegriffs, die Partition erwies sich von grösstem Nutzen: Das Aufspalten

[1]) Tschuprow, Grundbegriffe der Korrelationstheorie, Nordisk Statistisk Tidskrift 1924, S. 174.

eines komplizierten und widerspruchsvollen Ganzen in verschieden stark vertretene Teile zeitigte ganz neue Erkenntnisse. Wo es nicht möglich war, für eine grosse Gesamtmasse spezifische Merkmale aufzufinden, da war es dagegen oft möglich, die Masse in kleine Blocks verschiedener Grösse zu zerschlagen[1]), die aber jeder für sich ein ganz charakteristisches Gepräge zeigten. Diese Gebilde, wie z. B. bestimmte Gewerbszweige oder gewisse Berufsgruppen der Industrie oder des Handwerks, konnten ihrerseits wieder nach quantitativ bestimmten einheitlichen Merkmalen gegliedert werden und offenbarten dann weitere bezeichnende Unterschiede. Auch diese Angaben verglich man von Zählung zu Zählung. Man wurde so auf Entwicklungstendenzen aufmerksam. Ja durch Nebeneinanderstellen von vielen präzisierten Urteilen, weil diese sich durch das Zahlensystem, «einer Nomenklatur von unendlicher Feinheit und Ausdehnung» (Mach), darstellen liessen, in Reihen, deren Ondulationen von einer ideellen Ebene ausgehend gedacht wurden, ergaben sich besondere Konfigurationen, das Strahlenbündel, die Gleich-, Gegen- und Folgebewegung (Wagemann). — Untersuchungen, die in einem aufs Praktische gerichteten Teil der statistischen Tätigkeit, in der Konjunkturbeobachtung wie in der Meteorologie, besondere Verbreitung fanden.

Vielleicht wird manchem die hier gegebene Grundformel des statistischen Urteils als ärmlich und inhaltsleer erscheinen, weil er nicht die zahlreichen Folgerungen für die statistische Methodik übersieht und auch nicht daran denkt, dass dem logischen Symbol eine stillschweigende Voraussetzung zugrunde liegt, die wir zwar schon oben erwähnt haben, deren Wichtigkeit jedoch nicht stark genug betont werden kann. Diese Voraussetzung ist die Überzeugung, dass die präzisierten partikulären Urteile konstant seien, wenigstens bis zu einem gewissen Grade, sei es nun, wie manche glauben, weil sie nur eine unvollkommene Erscheinungsform vollkommener allgemeiner Urteile sind, die hinter ihnen stehen und daher eine streng kausale Grundlage besitzen, sei es, weil wahrscheinlichkeitstheoretische Erwägungen dazu führen. Aber auch wer weder an das eine noch andere glaubt: die relative Konstanz, das Beharrungsvermögen der Zahlen als eine Erfahrungstatsache, kann er nicht leugnen. Eine Erklärung besitzen wir allerdings nicht für sie[2]). Die Erfahrung aber genügt jedenfalls, um aus ihr Nutzen zu ziehen.

---

[1]) Stets handelt es sich um ein Zerlegen, nicht um ein «Zusammenfassen von Zählblättchen», wie ich an anderer Stelle (Zahlenfetischismus, diese Zeitschrift 1926, S. 310) gezeigt habe.

[2]) Dies ist die Meinung von Lexis (Handw. der Staatsw. 1910, VII, S. 825) für die Erscheinungen auf bevölkerungs- und moralstatistischem Gebiet. Die vielen Versuche, das Beharrungsvermögen der Zahlen zu erklären, sind wenig befriedigend, vor allem die stochastische. Ich muss hier auf meinen Aufsatz im Allgemeinen Statistischen Archiv «Philosophie der Statistik» verweisen. Eine sehr geistreiche, aber nicht allseitig begründete Theorie hat Kammerer in seinem «Gesetz der Serie», Stuttgart 1919, aufgestellt. In der Tat weisen zahllose Erscheinungen in der Natur und auch in der menschlichen Wirtschaft auf einen serienweisen Rhythmus hin, d. h. auf mehr oder weniger gleichmässiges Oszillieren bis zu einem gewissen Zeitpunkt, wonach dann ein anderer Rhythmus einsetzt — so in der Meteorologie, aber auch in der Konjunkturforschung, die den umfassendsten Gegenbeweis gegen die zufallsbedingten Schwankungen empirischer Zahlenreihen liefert.

Dass ein gewisser Ausgleich in den statistischen Massen eintritt und ähnliche Erscheinungen zeitigt, wird niemand leugnen wollen. Er erklärt sich einerseits durch Fehler, z. B. beim Abzählen von statistischen Zählkarten wird einmal zuviel, einmal zuwenig abgezählt. Anderseits führen

Wenn wir sie nicht hätten, müssten wir über die Zahlenmassen, die wir aufstapeln, erschrecken. Sie sind aus «Momentaufnahmen» entstanden, aber sie gelten für Zeit und Ewigkeit. Man hat sie als Zahlenfriedhöfe bezeichnet. Aber ebenso wie die in Friedhöfen versenkten Generationen wirken die in Zahlen niedergelegten Verhältnisse weiter im lebendigen Geschehen, weil aus ihnen die heute bestehenden herausgewachsen sind.

III. Die statistischen Schlüsse

Jeder Schluss ist nach der traditionellen Logik ein Urteil aus andern Urteilen[1]). Aber, hat man gesagt, dann bringt kein Schluss etwas Neues, weil er nur aus gegebenen Urteilen ein neues Urteil bildet. Der Schluss aus den Urteilen: alle Menschen sind sterblich, Sokrates ist ein Mensch, also ist er sterblich, liege bereits im Obersatz: alle Menschen sind sterblich, vor. Ein Schluss sei also höchstens dann neu, wenn er falsch ist.

Diese schon von den Skeptikern aufgebrachten Einwände gegen die Syllogistik sind unzählige Male wiederholt[2]), aber auch widerlegt worden[3]). Sie hätten nur dann ihre Berechtigung, wenn es keine Schlüsse gäbe, welche über die bisherige Erfahrung hinausgreifen.

Die der Statistik eigentümlichen Schlüsse, denen wir uns jetzt zuwenden, tun dies in mehrfacher Hinsicht.

1. Generalisierende Schlüsse. Sie kommen natürlich im täglichen Leben wie in der Wissenschaft überall vor, ihre eigentliche Heimat liegt aber zweifellos auf statistischem Gebiet. Aus dem unbestimmten Urteil «Einige S sind P» wird durch Erweiterung des Subjektbegriffes das bestimmte, allgemein gültige Urteil «Alle S sind P». Dies ist das Schema der Induktion. Auch in ihr liegt offenbar, ebenso wie im Präzisieren der partikulären Urteile, eine Verbesserung dieser unvollkommenen, unbestimmten Urteile. Nur erfordert der Prozess der Induktion einen Hilfssatz, ein weiteres Urteil. Aus eben diesem Grunde nennen wir das neu gewonnene allgemeine Urteil einen Schluss.

Das Hinausgreifen über die Erfahrung kann erstens darin bestehen, dass die Konstanz eines statistischen Urteils von der Form «So und so viele S von allen S sind P», die wir oben nur als relativ und problematisch dargestellt haben, als ständig erwiesen wird (So und so viele S von allen S sind immer P); zweitens kann eine Ausdehnung des partikulären Urteils «Einige S sind P» auf neue Fälle,

---

innere Bewegungen in den Massen oft zum Ausgleich. Die Statistik bewies z. B., dass in den schweizerischen Hochtälern keine Abnahme, sondern eine Zunahme der Bevölkerung zu beobachten war, trotz der notorischen und erschreckenden Abnahme der Bevölkerung einzelner Hochtalgemeinden. Sie war in andere Hochtalgemeinden, in Kurorte und Wintersportplätze, abgewandert.

[1]) Jerusalem, Lehrbuch der Psychologie. 4. Auflage. Wien 1907, S. 126: «Das Schliessen ist somit nichts anderes als ein Urteilen.»

[2]) Besonders geistreich von F. C. S. Schiller, Formal Logic, London 1912; auch von Mauthner, Beitr. z. Kritik der Sprache, Bd. III, S. 371 f.

[3]) Am besten bei Sigwart, Logik; ferner bei Wundt, Logik I, 1919, S. 303, der darauf aufmerksam macht, der Satz «Alle Menschen sind sterblich» werde nicht auf die gestorbenen, sondern auf die noch lebenden Menschen angewendet. «Wie anders würde es in der Welt aussehen, wenn nicht unsere ganze Lebensführung unter der Herrschaft dieses Syllogismus stünde!»

eine örtliche oder zeitliche Verallgemeinerung stattfinden (Auch andere S sind P; oder: Alle S sind P); drittens endlich ermöglicht sie nicht nur eine Erweiterung, sondern oft eine Verbesserung der Daten unserer Erfahrung[1]).

Der erste Fall, die Festigung eines partikulären Urteils, geschieht auf Grund von wahrscheinlichkeitstheoretischen Erwägungen und soll uns später noch kurz beschäftigen. Der zweite Punkt, die Ausdehnung des partikulären Urteils auf neue Fälle, wenn nur wenige bekannt sind, ist Sache der Analogie oder Induktion; die Verbesserung der Erfahrung, worauf wir hier nicht eintreten wollen, ist Sache der Interpolation.

Das Gebiet der Induktion ist weit ausgedehntes, jungfräuliches Land[2]). Es kann gar keine Rede davon sein, in einem Zeitschriftenaufsatz dieses Problem ausführlich zu behandeln. Für den Statistiker freilich ist wohl kaum eines in der ganzen wissenschaftlichen Methodenlehre wichtiger. Die Frage lautet für ihn: Was berechtigt uns dazu, von einigen uns bekannten Fällen auf andere ähnliche Fälle zu schliessen?

Die Masse allein tut es nicht. Die grosse Zahl allein vermag uns keine Sicherheit zu geben, von ihr auf eine noch grössere Zahl zu schliessen. Diese Auffassung ist zwar auch verschiedentlich vertreten worden: stark eingeschränkt in Laplaces Gesetz der Sukzession[3]) oder, wie Mill es genannt hat, im Schluss auf die angrenzenden Fälle[4]). Das oft zitierte Beispiel von der Sonne, die eine Million Mal aufgegangen ist und daher wohl noch ein weiteres Mal aufgehen wird, zeigt uns jedoch schlagend das Gegenteil dessen, was es beweisen soll. Wir schliessen auf ein weiteres Aufgehen dieses Gestirnes auf Grund nicht von Abzählungen, sondern von Beobachtungen und Hypothesen über die Gesetzmässigkeit der planetarischen Bewegung.

Der Hilfssatz, auf den wir uns in der Statistik, wo Experimente unmöglich sind, stets bewusst oder unbewusst stützen, wenn wir von einigen bekannten S auf weitere unbekannte oder auf alle S schliessen, besteht darin, dass wir Beobachtungen über die Variationsbreite und -häufigkeit der fraglichen Erscheinung heranziehen. Sie ist auf den verschiedensten Wissensgebieten ausserordentlich verschieden, weil jede Erscheinung ihre besondere, nur ihr eigentümliche statistische Konfiguration besitzt[5]). Daher ist mehr oder weniger Vorsicht geboten. Allgemeine Regeln lassen sich nicht aufstellen. Ich muss hier auf meine Ausführungen im Allgemeinen Statistischen Archiv 1931, Heft 2, verweisen, um nicht dort Gesagtes zu wiederholen.

---

[1]) Z. B. wenn man Punkte durch eine Kurve verbindet, die Kurve aber oft neben den beobachteten Punkten vorbeiführt. «On ne se borne pas à généraliser l'expérience, on la corrige.» Poincaré, La science et l'hypothèse, Paris 1912, S. 170.

[2]) Kants Wort: «Noch hat kein Logiker die Analogie und Induktion gehörig bearbeitet» (Philos. Hauptvorlesungen nach Kollegheften, München 1924, S. 490), gilt noch heute.

[3]) Abraham Wolf (nicht zu verwechseln mit Hellmuth Wolff, was die ausgezeichneten Arbeiten des Erstgenannten über Wissenschaftslehre und Logik nicht verdienen würden) in der Enc. brit. 14. Aufl., Bd. 12, S. 271, wo er die Schwächen dieses Gesetzes aufzeigt.

[4]) System der deduktiven und induktiven Logik, übersetzt von Schiel, Braunschweig 1877, II, S. 86 f.

[5]) Czuber, Die stat. Forschungsmethoden, Wien 1921, S. 47: «wie der Habitus für eine Pflanze».

Nur darauf möchte ich hinweisen, dass die sogenannte Repräsentative Methode in der Statistik auf einem ganz andern Grundsatz beruht. Hier ist der Grundgedanke massgebend, dass die Gesamtmasse Zufallsgesetzen gehorcht, d. h. dass beliebig herausgegriffene kleinere Teilmassen dasselbe Gefüge, dasselbe Konglomerat aufweisen wie die Gesamtmasse: eine Voraussetzung, welche bei sozialen und wirtschaftlichen Erscheinungen mit ihrer grossen Abhängigkeit der Einzelfälle voneinander und ihrer wechselnden Verflechtung niemals zutrifft. Daher können die Ergebnisse der repräsentativen Statistik auch nur in grober Annäherung ein Bild der unbekannten Gesamtmasse, von der sie ein Teil sind, liefern. Die neuesten Bestrebungen der repräsentativen Statistik laufen übrigens darauf hinaus, nicht mehr viele Muster (samples) nach irgendwelchen Zufallskriterien, wie alphabetische oder numerische Sukzessionen, aus der Masse herauszuschöpfen, sondern weniger umfangreiche Muster ganz bewusst daraufhin auszulesen, ob sie eine gewisse Gewähr dafür bieten, andere Teile der Masse einigermassen zu vertreten[1]). Aus dem Induktionsschluss wird damit ein solcher der Analogie.

2. Wahrscheinlichkeitsschlüsse. Laplace glaubte, aus dem Ungewissen das Gewisse, aus dem Zufälligen und Schwankenden das Feste, aus dem Dunkel des Unbekannten die «Wahrheit» hervorbrechen zu sehen — und zwar durch wahrscheinlichkeitstheoretische Berechnungen[2]). Diese Vorstellung, so sehr sie sich in der Folge als trügerisch erwies, spukt heute noch in den Köpfen mancher mathematischer Statistiker in etwas veränderter Gestalt, in dem Glauben an eine «Wesensform» der Erscheinung[3]), die sich in den Zahlenreihen, von den Schlacken des Zufalls gereinigt, vorfindet. Vergeblich haben sich die Biologen bemüht, die Statistiker zu überzeugen, dass weder eine Rasse noch eine Art oder gar eine Familie oder Klasse (so wenig wie das Geschlechtsverhältnis der Geborenen) durch statistische Häufigkeitsbeobachtungen oder Korrelationen festgestellt werden könne. Diesen hartnäckig verschlossenen Rätseln des Lebens werden jedenfalls die Statistiker nie beikommen können.

Die Wahrscheinlichkeitsschlüsse erstreben heute ein weit bescheideneres Ziel. Dies gilt vor allem von den nicht numerischen Wahrscheinlichkeitsschlüssen, die so alt sind wie die Logik selbst[4]). «Wenn ich, anstatt zu wissen, alle A sind B, nur weiss, dass die meisten A B seien, C aber A ist, so werde ich aber vermuten, dass C auch B sei. Oder ich weiss, alle A sind B, vermute, die C werden wohl A sein und schliesse, die C werden wohl B sein. Die meisten Gewitter ziehen von Westen nach Osten. Ich vermute dies also auch von dem heutigen. Auf solche Weise sind im täglichen Leben diese Schlüsse in unserem Denken beständig im Spiel[5]).» Kant gründete seine Theorie der Entwicklung des Planetensystems auf einen solchen einfachen Wahrscheinlichkeitsschluss[6]).

[1]) Statt gleichsam aus Hunderten von Urnen mit vermuteter gleicher Zusammensetzung nur je eine Kugel zu ziehen, werden wenige Urnen ausgewählt, aber auf ihren Gesamtinhalt geprüft.

[2]) Laplace, Essai philosophique sur les Probabilités, Paris 1921, I, S. 56 f.

[3]) Siehe meinen Aufsatz «Die statistische Wesensform» im Allg. Stat. Archiv 1928, S. 235.

[4]) Aristoteles, a. a. O. S. 142.

[5]) Fries, System der Logik, 1837, S. 318.

[6]) Wundt, Logik, I, 1919, S. 323.

«Der Wahrscheinlichkeitsschluss», schreibt Wundt[1]), «folgert aus der relativen Häufigkeit gegebener, aus einer Reihe einzelner Fälle bestehender Tatsachen auf die Wahrscheinlichkeit des zukünftigen Eintrittes der nämlichen oder übereinstimmender Fälle... Indem an die Stelle der unbekannten Folgerung der Wahrscheinlichkeit eines Ereignisses deren numerische Schätzung tritt, verwandelt sich nun der gemeine in den numerischen Wahrscheinlichkeitsschluss.»

Als Wahrscheinlichkeit bezeichnet man in diesem Fall bekanntlich das Verhältnis der dem Ereignis günstigen Fälle zu der Gesamtzahl der möglichen (d. h. günstigen und ungünstigen) Fälle. Diese «klassische» Definition der Wahrscheinlichkeit beruht auf zwei Fiktionen: dass nämlich erstens der gesamte Wahrscheinlichkeitsbereich sich in gleiche, zum mindesten gedacht gleiche Teile einteilen lasse, was in Form eines disjunktiven Urteils geschieht[2]), eine Anschauung, die schon Fries[3]) vor der geistreichen Entwicklung der Spielraumstheorie bei v. Kries gehabt hat. Die zweite Fiktion ist die, dass jedes günstige Einzelereignis gleich wahrscheinlich sei[4]). Ein Mann will sein Leben versichern. Man stellt sein Alter fest und sieht nach, dass unter tausend Personen seines Alters 990 das kommende Jahr überleben. Man nimmt daher an, dass er ebenfalls zu dieser Klasse gehöre, man behandelt ihn nicht als Individuum, das vielleicht eine ganz davon abweichende Lebenserwartung besitzt, weil er einen gefährlichen Beruf oder eine ungünstige erbliche Veranlagung hat, sondern als gleichartigen Teil seiner Klasse[5]).

Die zweite Methode, die Wahrscheinlichkeit eines Ereignisses zu bestimmen, besteht darin, dass man die relative Häufigkeit eines Ereignisses ohne weiteres mit seiner Wahrscheinlichkeit gleichsetzt (wenn die Zahl der Beobachtungen sehr gross war)[6]). So und so viele S sind P: dies ist bereits die Wahrscheinlichkeit dafür, dass stets auch neu hinzukommende S im gleichen Verhältnis zu P stehen werden. Fragen wie die folgenden: Wieviele grosse Söhne hatten grosse Väter? oder welches ist die Wahrscheinlichkeit dafür, dass ein blondes Mädchen blaue Augen hat? werden nach dieser Methode der Wahrscheinlichkeit einfach als das numerische Verhältnis der beobachteten Fälle bestimmt[7]), indem das partikuläre Urteil präzisiert wird: So und so viele S von allen beobachteten S sind P.

Diese Auffassung von der Wahrscheinlichkeit führt zu wunderlichen Konsequenzen[8]). Sie beruht ebenfalls auf fiktiven Annahmen. Die genaue Wahrscheinlichkeit kennen wir auf diese Art nie. Wir tun aber so, als ob sie einen definitiven Wert hätte[9]). Eine weitere Fiktion besteht in der Annahme, dass die Ab-

---

[1]) Logik I, 1919, S. 321.

[2]) Wundt, ebenda; ebenso Sigwart, Logik, Tübingen 1911, II, S. 319.

[3]) A. a. O. S. 322: «Der erstere Fall misst die Teile einer Sphäre gegeneinander und gibt deshalb eine mathematische Wahrscheinlichkeit.»

[4]) Daher Poincaré: «on a défini le probable par le probable».

[5]) Sheppard, Probability and Error, Enc. brit. 1928, 18, S. 532 *b*.

[6]) Ebenda, S. 531 *d*.

[7]) Sheppard, Probability and Error, Enc. brit. 1928, 18, S. 532 *c*.

[8]) Abraham Wolf, Probability, Enc. brit. 1928, 18, bekämpft die Häufigkeitsdefinition der Wahrscheinlichkeit.

[9]) Sheppard, a. a. O. S. 532 *a*.

weichungen vom empirisch festgelegten Wert sich wie zufällige Schwankungen bei freiem Spiel von gleichmöglichen Kombinationen verhalten, was zur binomialen Verteilung der Abweichungen und zum Gesetz der grossen Zahlen führt[1]).

Solche Fiktionen sind durchaus berechtigt und nützlich, solange man sich ihrer als Fiktionen bewusst bleibt. Sie werden jedoch gefährlich, wenn man sie für Hypothesen oder gar für Realitäten hält. Wer kann uns daran hindern, die Schwankungen einer empirischen Reihe als zufällige Schwankungen eines einfachen oder zusammengesetzten fingierten Urnenbeispieles zu betrachten oder in einer beliebigen Kurve das Zusammenwirken von mehreren übereinander gelagerten Normalkurven zu sehen, oder endlich die empirische «Streuung» einer Reihe mit einer normalen Streuung zu vergleichen? Dadurch gewinnen wir Massstäbe für die empirisch festgestellten Schwankungen einer Erscheinung. Wir können anhand von ihnen problematische Schlüsse auf den Eintritt kommender Ereignisse ziehen. Aber wir dürfen nicht glauben, dass diese empirischen Ereignisse tatsächlich durch das Zusammenspiel von zufälligen Ursachen bestimmt wurden: mit einem Wort, dass unsere Annahmen etwas anderes gewesen sind als reine Fiktionen.

3. Fiktive Schlüsse aus Einzelurteilen, wie wir die Mittelwerte nennen wollen, erfreuen sich des besten Rufes in der Statistik. Sie nimmt sie so wichtig, dass sie sich selbst als die Wissenschaft von den Mittelwerten bezeichnete[2]). Dem Statistiker wachsen die von ihm aufgetürmten Reihen über den Kopf. Er sucht sie durch eine einzige Zahl zu «charakterisieren»[3]). Es stehen ihm zu diesem Zweck ausserordentlich viele zur Verfügung[4]). Er begnügt sich mit drei: dem arithmetischen Mittel, dem Median- und dem dichtesten Wert, die nur dann, wenn eine Reihe ohnedies charakterisiert ist, zusammenfallen: bei der normalen, gleichästigen Verteilung. Sie sind nicht nur in allen andern Fällen gänzlich voneinander verschieden[5]); ein und derselbe Mittelwert kann auch gänzlich verschiedene Reihen «charakterisieren»[6]). Und nicht nur charakterisiert er Reihen; Reihen charakterisieren ihn. Denn er selbst muss seinerseits charakterisiert werden[7]), was am besten dadurch geschieht, dass man ihn wieder in eine Reihe (in Quartile, Dezile oder perzentile Grade) zurückverwandelt. «Aus Mittelwerten ist nämlich nicht ersichtlich, wie sich die in Mittelwerten zusammengefassten Verschiedenheiten im einzelnen gestalten... Wir sind daher häufig bestrebt, den Mittelwert durch Angaben über die Gestaltung der Reihe der Einzelwerte zu ergänzen» schreibt Zizek in seinem Grundriss (S. 150).

---

[1]) Ebendort, S. 534.

[2]) Bowley, Edgeworth. Joh. Müller (Theorie und Technik der Statistik, Jena 1927, S. 189) findet dies zu weitgehend.

[3]) Zizek, Tischer, Kaufmann, Winkler u. v. a.

[4]) Zizek, Grundriss der Statistik, 1921, S. 144.

[5]) Winkler, Statistik, Leipzig 1925, S. 71: «Dem statistischen Kenner wird es, die Anwendbarkeit aller dreier Werte vorausgesetzt, nahezu gleichgültig sein, durch welche die Kurve vertreten wird.»

[6]) Zizek, Die statistischen Mittelwerte, Leipzig 1908, S. 299.

[7]) Tischer, Grundlegung der Statistik, Jena 1929, S. 140: «Durch eine derartige Vereinfachung werden allerdings die oft für das Wesen der beobachteten Erscheinung sehr wichtigen Abweichungen der Kenntnis des Beobachters einfach entzogen.» Es sei dem Benützer der Mittelwerte erwünscht, die Vereinfachung wieder aufzuheben. Ähnlich Winkler, Statistik S. 67.

Jeder Mittelwert besitzt die Fehler seiner Vorzüge. Dem arithmetischen Mittel wird nachgerühmt, dass es auch die extremen Glieder miteinbeziehe[1]), dem dichtesten Wert, dass er dies nicht tue[2]). Dem arithmetischen Mittel wird von denselben Autoren vorgeworfen, dass es durch Einbezug der extremen Glieder sehr oft wertlos werde (Einkommen des Millionärs im Dorfe) [3]), während dem Zentralwert und dichtesten Wert vorgehalten wird, dass sie durch Nichteinbeziehen der extremen Glieder falsche Vorstellungen vom extremen Verlauf der Kurve erwecken[4]). Dem dichtesten Wert sagt man nach, dass er der einzige sei, der wenigstens ein Glied mit Bestimmtheit repräsentiere[5]). Man teilt die Mittelwerte in typische und nichttypische ein. Die nichttypischen seien ohne jede reale Grundlage eine «rein rechnerische Abstraktion»[6]). Sie seien weitaus die häufigsten[7]).

Welche Mittelwerte soll der Statistiker verwenden? Er soll sie «der Reihe nach durchprobieren»[8]). Welchen er dann nimmt, wird von «seinem Wissen, Können und seinen Erfahrungen abhängen»[8]).

Soweit die statistische Theorie. Nach ihr erheben die Mittelwerte also vor allem den Anspruch, die Reihen zu charakterisieren, manchmal auch, «das Gesetz der grossen Zahlen zum Ausdruck zu bringen»[9]), «die Ursache der Reihe»[10]), zum mindesten aber die Realität der Erscheinungen, widerzuspiegeln. In Wahrheit gibt es aber nur einen einzigen Fall, in dem ein Mittelwert eine Realität repräsentiert[11]), und gerade dieser Fall kommt für die Statistik fatalerweise überhaupt nicht in Betracht. Bei zahlreichen Messungen derselben Strecke ist das arithmetische Mittel der Messungen die wirkliche Strecke[12]) — nach dem Gaus'schen Gesetz, sofern es ein Gesetz und nicht nur eine Hypothese ist, wie Poincaré[13]) behauptet. — Wo jedoch grosse Zahlen verschiedener Elemente einer Masse beobachtet werden, z. B. die Körperlänge von Rekruten, hat das arithmetische

---

[1]) Zizek, 1908, S. 167.

[2]) Zizek, 1908, S. 273.

[3]) Zizek, 1908, S. 242.

[4]) Winkler, Statistik: «Ein Vorteil ist es, dass ganz unregelmässige Aussenfälle ohne Einfluss bleiben, ein Nachteil dagegen, dass unter Umständen auch wesentlichen Kurvenänderungen das Gehör versagt bleibt.» (S. 72.)

[5]) Zizek, 1921, S. 149.

[6]) Zizek, 1921, S. 141.

[7]) Zizek, 1908, S. 273: «Derartige typische Mittelwerte in strengem Sinne kommen jedoch bekanntlich nur sehr selten vor.»

[8]) Joh. Müller, a. a. O. S. 222.

[9]) Zizek, 1926, S. 141.

[10]) Joh. Müller, a. a. O. S. 189.

[11]) Dass ein Mittelwert zufällig die gleiche Grösse haben kann und beim dichtesten Wert auch haben muss wie ein Einzelglied der Reihe, beweist nicht, dass er eine Realität ist; denn dieses Einzelglied kann ebenso wie der Mittelwert gänzlich ungeeignet sein, den Reihenverlauf zu charakterisieren.

[12]) Hugo Forcher, Die statistische Methode, Leipzig 1913, S. 297: «Bei wiederholten Messungen desselben Gegenstandes ist das arithmetische Mittel der wahrscheinlichste Wert des Gegenstandes selbst.»

[13]) Calcul des probabilités, Paris 1912, S. 170 und 216: «Mit einem in Millimeter geteilten Meterstab wird man niemals, sooft man die Messungen wiederholt, eine Länge auf ein Millionstel Millimeter bestimmen können.»

Mittel keine reale Bedeutung; doch steckt hinter der Mittelwertsbestimmung immerhin ein realer Gedanke, schreibt Forcher[1]), weil «die einzelnen Werte einer extensiven Grösse, denen eine derartige Verteilung um den Mittelwert zukommt, als zufällig gestörte Spezialisierungen eines Grundwertes der betreffenden Grösse betrachtet werden. Die vielen Mittelwerte der statistischen Empirie hingegen entbehren in den meisten Fällen jeder realen Basis... Sie können die Massenerscheinung nicht mehr repräsentieren, ihnen kommt keinerlei Realität zu, sie sind aus diesem Grunde wertlos und verwirren nur die Sachlage, statt dieselbe zu klären.»

Dies gilt natürlich auch von den Beziehungszahlen und erst recht von den sogenannten isolierten Mittelwerten. Sie sollen ja nach der Theorie eine Reihe charakterisieren oder extrahieren, diese jedoch besteht hier nur in der Phantasie (z. B. Lohnsumme aller verunfallten Arbeiter, geteilt durch die entsprechende Zahl der Vollarbeiter).

Wie entsteht das arithmetische Mittel? (Seine Berechnung kennt jeder.) Aus den Urteilen: «S_1 ist P_1»; «S_2 ist P_2»; «S_3 und S_4 sind ebenfalls P_2», oder konkreter: 1 Arbeiter hat Fr. 2 Lohn; 5 Arbeiter haben Fr. 4 Lohn; 2 Arbeiter haben Fr. 3 Lohn, wird der Schluss abgeleitet: «alle S sind Pm», «alle 8 Arbeiter haben Fr. $3,_5$ Lohn». Man sieht hier deutlich, dass man bei der Berechnung auf Einzelurteile zurückgehen muss. Das allgemeine Urteil: «5 Arbeiter haben Fr. 4 Lohn» muss aufgelöst werden in: «1 Arbeiter hat Fr. 4, ein zweiter ebenfalls, ein dritter bis fünfter ebenfalls.» Das nicht gewogene arithmetische Mittel ist überhaupt kein Mittel, sondern ein Schnitzer. — Aus den Einzelurteilen wird also ein allgemeines Urteil gewonnen, das aber nur ungefähr richtig ist, das nur dann einen Sinn hat, wenn ohnedies schon die meisten Werte sehr nahe beim Durchschnitt liegen. Das arithmetische Mittel gilt mit Recht als um so typischer, je weniger Prokrustes-Arbeit es zu verrichten hatte[2]).

Der dichteste oder häufigste Wert ersetzt ebenfalls eine Reihe von Einzelurteilen durch ein allgemeines Urteil. Ihm liegt aber eine andere Voraussetzung zugrunde. Sein Element ist die schiefe Verteilungskurve, deren Gipfelhöhe er misst. Daher sein Name «Scheitelwert». Wird der Dromedarhöcker durch einen zweigipfligen Kamelhöcker (Ausdrücke von Bowley) oder gar durch eine «Hochebene» ersetzt, so gibt es entweder mehrere häufigste (!) Werte, die nichts mehr charakterisieren können, oder einen so desorientierten wie die Magnetnadel am Pol, der wahrhaftig kein «Mittelwert der Lage» mehr ist.

Dem Medianwert (dem mittleren Glied einer orgelpfeifenartig geordneten Reihe) liegt die Voraussetzung zugrunde, dass eine grosse Zahl von Einzelwerten ihm ungefähr entsprechen, dass die Abweichungen von ihm sehr wenig zahlreich und um so seltener sind, je grösser sie werden. (Diese Forderung gilt übrigens für

---

[1]) A. a. O. S. 297.

[2]) Schott, Statistik, 1920, hat das arithmetische Mittel mit der Gewohnheit des Prokrustes verglichen, seine Gäste seinem Bett anzupassen, indem er die zu langen amputierte, die zu kurzen streckte; er hat damit treffend das Gewaltsame, das jeder Fiktion anhaftet, zum Ausdruck gebracht.

jeden der drei besprochenen Mittelwerte)[1]. Bei einer J-Kurve (Einkommensverteilung) verliert der Medianwert ebenso wie das arithmetische Mittel allen Sinn. Die vollkommenste Form des Medianwertes findet sich daher bei einer normalen (binomialen) Verteilung, wenn die Einzelwerte in aufsteigender Grösse geordnet werden und die Galtonsche Ogive-Kurve ergeben.

Warum sind Mittelwerte nur auf sogenannte «eigentliche Reihen» anwendbar? Warum fällt es niemanden ein, das arithmetische Mittel aus «uneigentlichen Reihen[2])», z. B. aus den Zahlen der Berufstätigen verschiedener Berufe, zu bilden[3])? Die Antwort ergibt sich aus unseren Ausführungen von selbst: Solche uneigentlichen Reihen sind eben gar keine Häufungen von Einzelurteilen, sondern präzisierte partikuläre Urteile, durch die eine Gesamtmasse in Teilmassen zerlegt wird.

Dass die sogenannten Beziehungszahlen, für die eine Reihe von höchst unglücklich gewählten Bezeichnungen existieren, allesamt ebenfalls nur fiktive Masszahlen[4]) sind, die nichts messen können, sondern höchstens «einen mnemotechnischen Wert» (Forcher) besitzen, weil sie disparate Urteile zueinander in Beziehung setzen, liegt auf der Hand. Selbst sehr real scheinende Berechnungen, z. B. die Berechnung des Schwerpunktes der Bevölkerung eines Landes, sind rein fiktive[5]). Auf einer gewichtslos gedachten Scheibe von der Grösse und Form des Territoriums der Vereinigten Staaten wird von der amtlichen amerikanischen Statistik die wirkliche Verteilung der Bevölkerung festgestellt und der Schwerpunkt der durch sie belasteten Scheibe als «Schwerpunkt der Bevölkerung» bezeichnet. Wendet man diese Vorstellungsweise auf Afrika an, einen Kontinent mit starker Randbesiedelung, so liegt der Schwerpunkt der Bevölkerung in der Sahara.

IV. Das Wesen der Statistik

Unzählige Versuche sind unternommen worden, dem Wesen der Statistik, diesem verschleierten Bild, nahezukommen. Alle sind unzureichend, denn alle umfassen nur einen Teil von dem, was ein berufstätiger, nicht zu sehr spezialisierter Statistiker treibt.

Einer der geistreichsten und unhaltbarsten Versuche ist zweifellos der, jede statistische Zahl als unvollkommenen, zufallsgestörten Ausdruck einer verborgenen Wirklichkeit aufzufassen[6]). Verhältnismässig weit reicht der Blick jener Theoretiker, die alles im «Übertragen» sehen[7]). Bescheidener sind die Forscher, die das

---

[1]) Zizek, Grundriss, 1921, S. 141, verlangt dies von einem typischen Mittelwert. Noch weitgehendere Anforderungen in «Die statistischen Mittelwerte», S. 273.

[2]) Winkler, «Statistik», 1925, S. 40, man könne sie nicht zum Gegenstand ernsthafter Forschung machen.

[3]) Winkler, a. a. O., gibt eine gesuchte Erklärung: die verschiedenen Berufe haben eine verschieden grosse «Berufsfläche», sie haben keinen «Reihungsgrund» (S. 40 und 45).

[4]) Sigwart, a. a. O. II, S. 700: «Es ist eine blosse Fiktion, dass sich diese Verhältnisse im Raume oder der Zeit gleich verteilen.»

[5]) Vaihinger, Philosophie des Als Ob, 5. Aufl., Leipzig 1920, S. 34, hat die Mittelwerte als Fiktionen hingestellt.

[6]) Tschuprow in Nordisk Stat. Tidskrift 1924, S. 434: «Stets gelten die konkreten Einzelzahlen bloss als durch den Zufall mehr oder weniger entstellte Vertreter der „wahren“ Grössen, die ihnen „zugrunde liegen“.»

[7]) Seuttemann, Das Ganze der Statistik und die beherrschende Idee, Hildebr. Jahrb. 1929, S. 519.

Wesen der Statistik im Beschreiben von kollektiven Ganzen[1]), im Charakterisieren oder «Kondensieren» durch Mittelwerte usw.[2]) oder gar lediglich in der «Kunst des Zählens»[3]), also in einer Technik, erblicken. Nichtssagend sind die Bezeichnungen des Wesens der Statistik als Reihenbetrachtung oder als Inbeziehungsetzung statistischer Massen[4]); überholt die Auffassung, dass durch sie Rassen und Arten unterschieden werden können (Galton)[5]). Manche Autoren wollen die Statistik auf einer charakteristischen Einzelheit aufbauen, wie March[6]) auf der Erfahrungstatsache des Ausgleichs der Einzelfälle in grössern Massen.

Die hier vertretene Auffassung sieht das Wesen der Statistik im nähern Bestimmen unbestimmter Urteile. Was ich unter diesen drei Worten verstanden wissen möchte, sei durch folgende Übersicht noch einmal dargelegt:

Alle unsere Erkenntnis nimmt die Form von Urteilen an[7]). Wenn allgemeine Urteile («Alle S sind P») nicht möglich sind, wenn wir nur von einigen Gegenständen des Denkens (Subjektbegriffen) behaupten können, dass ihnen gewisse Merkmale (Prädikatsbegriffe) zukommen, hat uns die Statistik die folgenden Methoden an die Hand gegeben, diese unvollkommene partielle Erkenntnis (von der Form: «Einige S sind P») fruchtbringend zu gestalten:

I. Bestimmen des Subjektsbegriffs (S) des unbestimmten Urteils durch:

a) Präzisieren; aus: «Einige S sind P» wird: «So und so viele S von allen S sind P» (Präzision, Statistisches Urteil);

b) Erweitern; aus: «Einige S sind P» wird: «Alle S sind P» (Induktion; Statistischer Schluss durch Beobachtung der Variationsbreite einer Erscheinung);

c) Abschätzen; Aus dem disjunktiven Urteil «S ist entweder P_1 oder P_2» wird (durch Hinzutreten von präzisierten Urteilen: «Es gibt n Fälle von P_1 und m Fälle von P_2») der Schluss abgeleitet: «Alle S sind wahrscheinlich $\frac{n}{n+m}$» (Probabilität)[8]).

II. Bestimmen des Prädikatsbegriffs (P) des unbestimmten Urteils durch:

a) Präzisieren; aus dem divisiven Urteil: «S ist teils P_1 teils P_2» wird: «S ist zu so und so vielen Teilen P_1 und zu so und so vielen Teilen P_2.» (Partition, Statist. Urteil);

---

1) Tischer, a. a. O.

2) Galton, Inquiries into human faculties, New York 1883, S. 49.

3) Bowley, Statistics, 1919, S. 3.

4) Winkler, Statistik, 1925, S. 84.

5) A. a. O., S. 49.

6) In De la Méthode dans les Sciences, Paris 1924, II, S. 352.

7) «Das Erkennen lässt sich seinem Wesen und der Sache nach als ein Prädizieren auffassen.» Lask, Die Lehre vom Urteil S. 59.

8) Den an logische Symbole nicht gewöhnten Lesern sei hier ein Beispiel gegeben: Blonde Mädchen sind entweder solche mit blauen oder nicht blauen Augen. Durch Bestimmen der n Fälle von Blondinen mit blauen gegenüber den m Fällen von Blondinen mit nicht blauen Augen wird: Alle blonden Mädchen haben die Wahrscheinlichkeit n/n+m, mit blauen Augen auf die Welt zu kommen.

b) Erweitern; aus: «Einige S haben die Merkmale *a*, *b* und *c* und sind daher P» wird: «Andere S haben ebenfalls die Merkmale *a* und *b*, also auch *c*, und sind daher ebenfalls P.» (Analogie; Statistischer Schluss der repräsentativen Methode);

c) Abschätzen; aus den Einzelurteilen: S_1 ist P_1; S_2 ist P_2; S_3 ist ebenfalls P_2 (die nur die Zusammenfassung zum Urteil: Einige S sind P_2 zulassen würden), wird: Alle S sind Pm, wobei Pm ein Mittelwert ist (Fiktiver Schluss).

Die statistischen Begriffe sind tastende Versuche zur Urteilsbildung, wenn sie die Erhebungseinheit definieren; sie sind der Niederschlag von zahlreichen statistischen Urteilen, wenn die Erhebung verarbeitet ist. Die statistischen Urteile stellen Tatsachen fest. Aus ihnen werden die generalisierenden Schlüsse gezogen, die ebenfalls von empirischen Beobachtungen begleitet sein müssen, wenn sie etwas bedeuten sollen. Aber sie, wie die Wahrscheinlichkeitsschlüsse und die fiktiven Schlüsse aus Einzelurteilen — dessen sollte man sich stets bewusst bleiben —, tun bereits einen Schritt aus der wirklichen Welt hinaus ins Imaginäre.