Skripte zur Wissenschaftstheorie, III

Zur Metaphysik der Sozialstatistik

G. Rohwer

Version 1 April 2000

Vorbemerkung. Bei den „Skripten zur Wissenschaftstheorie“ handelt es sich um Texte, die als Leitfäden für Seminare zur Wissenschaftstheorie der Sozialwissenschaften dienen sollen. Der vorliegende Text beschäftigt sich mit der Geschichte der theoretischen Sozialstatistik.

Götz Rohwer
Fakultät für Sozialwissenschaft
Ruhr-Universität Bochum, GB 1
44780 Bochum

Email: `goetz.rohwer @ ruhr-uni-bochum.de`

wtc.tex April 2000
Homepage: `http://www.stat.ruhr-uni-bochum.de [lecture notes]`

Inhaltsverzeichnis

Kapitel 1

Einleitung

Thema dieses Skripts ist die Geschichte der Sozialstatistik, genauer gesagt eine bestimmte Fragestellung: Es soll untersucht werden, wie die Sozialstatistik in ihrer historischen Entwicklung durch Überlegungen geprägt worden ist, die ihren Ursprung in der Wahrscheinlichkeitsrechnung haben.[1] Allerdings geht es mir nicht um eine historische Darstellung, sondern der Rückgriff auf ältere Publikationen soll helfen, ein kritisches Verständnis für die Herausbildung und den Charakter einer Reihe von Auffassungen zu gewinnen, die bis heute wirksam sind. Das Ziel ist, schließlich auch ein etwas besseres Verständnis der statistischen Methoden und Modelle zu finden, wie sie in der heutigen Sozialforschung verwendet werden. In diesem einleitenden Kapitel wird zunächst kurz an die Geschichte der Sozialstatistik erinnert, dann wird die im weiteren zu verfolgende Fragestellung etwas genauer charakterisiert.

1.1 Sozialstatistik

1. Einige Autoren verlegen die Anfänge der Sozialstatistik bereits in die Zeit erster Volkszählungen, die im Altertum durchgeführt wurden.[2] Eine wichtige Quelle für ihre spätere Entwicklung kann in den Ansätzen zu einer „systematischen Staatsbeschreibung“ gesehen werden, die seit der frühen (europäischen) Neuzeit verfolgt wurden. Eine interessante Dokumentensammlung für den Zeitraum von 1456 bis 1813, ergänzt durch einen ausführlichen Kommentar, wurde von Rassem und Stagl (1994) herausgegeben.

2. Eine wesentliche Intention dieser frühen Bestrebungen kann exemplarisch anhand einer Schrift von G. W. Leibniz aus dem Jahr 1685

[1] Es gibt bereits eine Reihe von Untersuchungen, die – wenn auch nicht nur für die Sozialstatistik – von dieser Leitfrage ausgehen. Hinzuweisen ist zunächst auf die beiden Bände zur „Probabilistic Revolution“, die von L. Krüger u.a. (1987) herausgegeben worden sind; weiterhin auf „The Empire of Chance“ von G. Gigerenzer u.a. (1989) und die Monographie zur Geschichte der Ökonometrie von M. S. Morgan, 1990.

[2] John 1884, S. 15ff, v. Mayr 1895, S. 158ff. Eine umfassende Sammlung von Informationen zur Geschichte von Volkszählungen gibt es bei Alterman, 1969. Interessante Informationen über sozialgeschichtliche Hintergründe finden sich auch bei Lindner, Wohak und Zeltwanger, 1984.

verdeutlicht werden:[3]

„Ich nenne *Staats-Tafeln* eine schriftliche kurze Verfassung des Kerns aller zu der Landesregierung gehörigen Nachrichtungen, so ein gewisses Land insonderheit betreffen, mit solchen Vorteil eingerichtet, daß der hohe Landesherr alles darin leicht finden, was er bei jeder Begebenheit zu betrachten, und sich dessen als eines der bequemsten Instrumente zu einer löblichen Selbst-Regierung bedienen könne."

„Durch *Nachrichtungen* verstehe ich nicht allerhand Vernunftschlüsse und Regeln, so verständige Leute bei Gelegenheit und wenn sie darauf zu denken Ursach haben, selbsten leicht finden können, sondern was mehr in facto als Nachsinnen beruhet und daher nicht erfunden, sondern erfahren, erhöret und erlernet werden muß, zum Exempel was in einem Lande für eine Quantität seidener Zeuge oder wüllene Tücher jährlich konsumiert oder vertan werden, das ist eine Nachrichtung und beruhet in facto, es kann es auch keiner erraten, er sei so verständig als er wolle; ob aber ratsam, solche consumation vor sich gehen zu lassen, oder zu verbieten, und enger zu spannen, und ob man solche Manufakturen im Lande selbst einzuführen habe oder nicht, bestehet in ratiocinatio und gehöret nicht zu unserer Staatstafel, sondern kann vielmehr aus denen in der Staatstafel befindlichen Nachrichtungen von verständigen Leuten leicht geschlossen werden."

Einen gewissen Höhepunkt hat diese Idee einer Staatszustandskunde in der „deutschen Universitätsstatistik" erreicht, wie sie sich in Deutschland im 17. und 18. Jahrhundert entwickelt hat.[4] Wichtige Vertreter waren u.a. Hermann Conring (1606–1681), Gottfried Achenwall (1719–1772), und August Ludwig von Schlözer (1735–1809), von dem vor allem der Ausspruch berühmt geworden ist: „Geschichte ist fortlaufende Statistik, Statistik ist stillstehende Geschichte."[5]

3. Während die deutsche Universitätsstatistik darauf aus war, mithilfe historisch-beschreibender Methoden möglichst flächendeckende Kenntnisse über die Beschaffenheit von Staatswesen zu erreichen, hat sich parallel, primär in England, eine „Politische Arithmetik" entwickelt, die darauf abzielte, soziale Sachverhalte möglichst weitgehend „in Zahl-, Gewichts- oder Maßbezeichnungen auszudrücken".[6] Wichtige Vertreter waren John Graunt (1620–1674); William Petty (1623–1687), der – als

[3] Zitiert nach Rassem und Stagl 1994, S. 321–329.

[4] Eine ausführliche Darstellung der Geschichte der älteren Statistik (bis 1835) findet sich bei John, 1884.

[5] Biographische Angaben und Auszüge aus Originalarbeiten finden sich bei Rassem und Stagl, 1994.

[6] William Petty, zitiert nach John 1884, S. 185.

Titel eines seiner Bücher – den Namen „Political Arithmetick" prägte und diese als „the art of reasoning by figures upon things relating to government" definierte; und Edmund Halley (1656–1742), der das Konzept der Sterbetafeln entwickelt hat. Für Deutschland gilt Johann Peter Süßmilch (1707–1767) als einer der ersten wichtigeren Vertreter der politischen Arithmetik.[7]

4. Die weitere Entwicklung der Sozialstatistik ist im wesentlichen den Ideen der Politischen Arithmetik gefolgt. In seiner Programmschrift „Political Arithmetick" definierte Petty die von ihm verfolgte Methode so:

„Die dazu von mir angewandte Methode ist noch nicht sehr gebräuchlich. Denn anstatt nur komparative und superlative Worte oder intellektuelle Argumente zu gebrauchen, habe ich den Weg eingeschlagen (als ein Muster der Politischen Arithmetik, das ich schon seit langem im Sinn hatte), mich in den Termini *Zahl*, *Gewicht* und *Maß* auszudrücken, mich nur aus sinnlicher Erfahrung abgeleiteter Argumente zu bedienen und nur solche Ursachen zu betrachten, die sichtbare Grundlagen in der Natur haben. Dabei überlasse ich solche Argumente den Erörterungen anderer, Argumente, die von wankelmütigen Neigungen, Ansichten, Gelüsten und Leidenschaften einzelner abhängen, da ich in der Tat erklären muß, daß ich mich zu solchen Gründen (wenn man sie überhaupt Gründe nennen darf) nicht äußern kann [...]. Nun sind die durch *Zahl*, *Gewicht* und *Maß* ausgedrückten Beobachtungen oder Thesen, worauf ich die folgenden Darlegungen gründe, entweder wahr oder nicht offensichtlich falsch. Und wenn sie nicht bereits wahr, bestimmt und offensichtlich sind, so kann sie aber die souveräne Macht zu Beobachtungen und Thesen von dieser Art machen, denn *nam id certum est quod certum reddi potest* [dasjenige ist glaubwürdig, was als glaubwürdig dargestellt werden kann]. Und wenn sie falsch sind, dann sind sie nicht so falsch, daß sie das Argument zunichte machen, für das sie vorgebracht werden. Doch schlimmstenfalls taugen sie als Annahmen, um den Weg zu den Kenntnissen zu weisen, nach denen ich strebe."[8]

Wie schon Leibniz in dem oben angeführten Zitat betont auch Petty, daß Wissen über gesellschaftliche Verhältnisse empirisch – durch „aus sinnlicher Erfahrung abgeleitete Argumente" – gewonnen werden muß.

5. Ein zweiter Gesichtspunkt liegt in der Betonung von „Zahl, Gewicht und Maß". Vornehmlich durch diese quantitative Orientierung kam es auch zur Entwicklung eines Gegensatzes zur älteren „Universitätsstatistik". Der Gegensatz, wie er sich etwa bis Mitte des 19. Jahrhunderts

[7] Nähere Informationen wiederum bei John 1884, sowie Rassem und Stagl 1994, Porter 1986, S. 18ff. Eine Zusammenstellung deutscher Übersetzungen seiner Schriften findet sich bei Petty, 1986. Auszüge aus den Schriften von Graunt und Halley finden sich bei Newman 1956, Bd. 3.

[8] Zitiert nach der dt. Übersetzung der Ausgabe von 1690, in Petty 1986, S. 221f.

herausgebildet hatte, wurde von Knies (1850, S. 81ff) so dargestellt:

„Man muß unterscheiden: eine historisch schildernde, beschreibende Darstellung, welche in der Vorführung des Stoffes, zur Erreichung der erstrebten Resultate die Methode der geschichtlichen Wissenschaften wählt; und eine mathematisch rechnende, mit Zahlenbestimmungen wirthschaftende, welche die Basis, die Evolutionen, den Calcül der exacten Wissenschaften verlangt und anwendet. [...]

1) Die ganze Achenwall-Schlözersche Schule, von dem Vater der Statistik an bis auf seine spätesten Enkel; die Anhänger der deutschen Universitäts- wie der Büchingschen Statistik, mögen sie Alles und Alles, eine Encyklopädie des Himmels und der Erde oder nur so einzelne Dinge wie die Verfassung oder Verwaltung der Staaten vorführen, haben sich für die beschreibende, schildernde Darstellung der historischen Wissenschaften entschieden. [...]

2) Eine wesentlich von dieser eben geschilderten verschiedene Methode haben andere Statistiker gewählt. Ihre Zahl ist geringer, sie unterscheiden sich untereinander außerdem in mancher anderen Beziehung, aber darin stimmen sie überein, daß nur das in Zahlen ausgedrückte und ausdrückbare Datum Gegenstand der Statistik sei, daß die bloße Schilderung, die historische Beschreibung aus der statistischen Wissenschaft zu verbannen sei. [...]

Es steht sich gegenüber die mit Worten zu beschreibende Thatsache und das von der Zahl begleitete Datum, die Schilderung der Phrase und das exacte Maß der Ziffer, das Bild, das Gemälde mit seinem allgemeinen Umriß und der exacte Calcul."

6. In der zweiten Hälfte des 19. Jahrhunderts setzte sich mehr und mehr die zweite der beiden von Knies unterschiedenen Konzeptionen der Sozialstatistik durch. Gegen Ende des 19. Jahrhunderts entwickelte insbesondere Georg v. Mayr (1841–1925) eine in Deutschland für längere Zeit dominierende Auffassung der Sozialstatistik, deren Grundgedanken er folgendermaßen darstellte:

„Die Wissenschaft von den sozialen Massen. Aus der geordneten Beobachtung der einzelnen Elemente der vorbezeichneten sozialen Massen erwächst der wissenschaftliche Gewinn des Einblicks in den Bestand und die Gliederung der Massen und damit der Erkenntniß aller in diesen Massen zum Ausdruck kommenden gesellschaftlichen Zustände und Erscheinungen. Am vollständigsten ist diese Erkenntniß einer sozialen Masse dann, wenn erschöpfende Beobachtung sämmtlicher Elemente, aus denen sie zusammengesetzt ist, vorliegt, und wenn diese Beobachtung zugleich in exakter Weise, d.h. mittelst Zählens und Messens, durchgeführt wird. Der Grundgedanke der so gearteten wissenschaftlichen Arbeit ist, aus der exakten Beobachtung der Elemente, aus deren angemessener Gruppierung, Gliederung und Vergleichung die Erkenntniß des Ganzen, d.i. der sozialen Masse zu gewinnen. [...]

Am vollkommensten wird diese Erforschung der sozialen Masse durch die erschöpfende Massenbeobachtung ihrer Elemente in Zahl und Maß bewerkstelligt. Die so geartete wissenschaftliche Erforschung der sozialen Masse nennen wir *Statistik*. Die Statistik ist hiernach recht eigentlich die Wissenschaft von den sozialen Massen." (v. Mayr 1895, S. 5)

Als Definition des Wortes 'Statistik' sind diese Ausführungen sicherlich nicht mehr gültig, denn die Statistik hat sich – schon zu Lebzeiten v. Mayrs – weitgehend zu einer Methodenwissenschaft entwickelt.[9] Sie verdeutlichen jedoch ein Verständnis von *Sozialstatistik*, das bis heute fortlebt. Das wesentliche Ziel besteht darin, mithilfe von systematisch erhobenen Daten und sich daran anschließenden statistischen Begriffsbildungen gesellschaftliche Verhältnisse zu beschreiben.

1.2 Theoretische Ansprüche

1. Nicht nur im Hinblick auf das heutige Verständnis von Statistik als Methodenlehre sind die Auführungen v. Mayrs nicht mehr zutreffend. Auch im Hinblick auf das heutige Selbstverständnis empirischer Sozialforschung – soweit sie statistische Methoden verwendet – gibt es nur noch

[9]Hinweise finden sich bei Porter 1986, S. 235f. Porter erwähnt u.a. J. J. Fox: „who argued in 1860 in a paper read before the Statistical Society of London: Statistics cannot really be called a science, since it 'has no facts of its own; in so far as it is a science, it belongs to the domain of Mathematics. Its great and inestimable value is, that it is a 'method' for the prosecution of other sciences. It is a 'method of investigation' founded upon the laws of abstract science; founded on the mathematical theory of probabilities; founded upon that which had been happily termed the logic of large numbers.' " v. Mayrs Auffassung, daß die Statistik – jedenfalls als Sozialstatistik – als eine selbständige Wissenschaft konzipiert werden sollte (vgl. v. Mayr, 1907), ist auch dadurch obsolet gemacht worden, daß ihr Gegenstandsbereich durch die sich differenzierenden Sozialwissenschaften okkupiert worden ist. W. Wundt sagte dazu in seiner „Logik" (1908, S. 494) : „Ursprünglich bedeutet bekanntlich das Wort 'Statistik' nichts anderes als 'Staatskunde', fällt also vielmehr mit der heutigen Staatswissenschaft als mit der Bevölkerungslehre zusammen. Dann wurde es speziell auf die im politischen Interesse angewandten Abzählungsmethoden, auf die sogenannte 'politische Arithmetik', und von da aus auf die Bevölkerungslehre und alle die Gebiete, die auf die Anwendung ähnlicher quantitativer Methoden angewiesen sind, übertragen. (Vgl. John, Geschichte der Statistik. 1884.) Nach dieser Verallgemeinerung, nach der von 'statistischen Methoden' auch auf den Gebieten der Naturwissenschaft, z.B. in der Metereologie, und der Psychologie z.B. in der Psychophysik, geredet werden kann, ist es offenbar das einzig richtige, unter Statistik nur noch eine Methode, aber keine besondere Wissenschaft zu verstehen. Auch ist der Ausdruck im letzteren Sinne umso überflüssiger, als es kein Gebiet gibt, in welchem die statistische Methode angewandt wird, das nicht nach anderen sachlicheren Merkmalen bereits zureichend definiert und benannt wäre."

wenige Berührungspunkte. Orientiert man sich an der gegenwärtigen Methodenliteratur, geht es in der empirischen Sozialforschung nicht in erster Linie darum, gesellschaftliche Verhältnisse zu beschreiben (dies erscheint bestenfalls als Aufgabe einer separat konzipierten „Wirtschafts- und Sozialstatistik"), sondern es wird angestrebt, soziale Sachverhalte und ihre Entwicklung zu *erklären*. Verbreitet hat sich insbesondere eine Rhetorik, wonach das Ziel empirischer Sozialforschung in der Formulierung und Prüfung „hypothetischer Gesetzmäßigkeiten" besteht, von denen angenommen wird, daß sie das gesellschaftliche Leben beherrschen.[10]

2. Welche Vorstellungen sich hinter dem gegenwärtigen Selbstverständnis empirischer Sozialforschung verbergen, ist schwer durchschaubar. In diesem Skript soll deshalb versucht werden, einem Verständnis auf indirekte Weise näher zu kommen. Es soll untersucht werden, wie in der Geschichte der Sozialstatistik theoretische Fragestellungen konzipiert und verfolgt worden sind. Wie sich zeigen wird, haben viele der dafür grundlegenden Begriffsbildungen und Gedankengänge ihren Ursprung in der Entwicklung der Wahrscheinlichkeitsrechnung. Ich möchte also hauptsächlich untersuchen, wie sich durch die Übernahme wahrscheinlichkeitstheoretischer Begriffsbildungen und Vorstellungen eine theoretische Sozialstatistik herausgebildet hat, etwa seit der 2. Hälfte des 19. Jahrhunderts.

3. Die Einflüsse der Wahrscheinlichkeitsrechnung auf die Sozialstatistik kommen noch heute darin zum Ausdruck, daß in der statistischen Methodenlehre, wie sie im Bereich der Sozialwissenschaften vermittelt wird, oft zwischen einer „deskriptiven" und einer „schließenden" Statistik unterschieden wird. Zum Beispiel heißt es in einem neueren Lehrbuch von D. Hochstädter (1991):

„Man kann zwei Erkenntnisziele der Statistik unterscheiden: *deskriptive Statistik* $\longleftrightarrow$ *schließende Statistik*. Die deskriptive Statistik beschränkt sich auf die reine *Beschreibung* der erhobenen Massenerscheinung. Ihr Gegenstand sind Umfang, Gliederung, graphische Darstellung oder zeitliche Entwicklung des erhobenen Datenmaterials sowie die Verhältnisse verschiedener Massen zueinander.

[10] Exemplarisch sei auf einen programmatischen Aufsatz von E.K. Scheuch und D. Rüschemeyer über „Soziologie und Statistik" (1956, S. 348) hingewiesen: „Abgesehen von einer dauernden Bindung an die empirische Kontrolle stellt sich die von der modernen Wissenschaftslehre ausgehende Soziologie zudem die Aufgabe, über die Beschreibung hinaus allgemeine Gesetze (das heißt Hypothesen über konstante Beziehungen zwischen Variablen) aufzustellen." Um zu verstehen, was hier mit „moderne Wissenschaftslehre" gemeint ist, vgl. man exemplarisch Reichenbach, 1951.

Die *schließende Statistik* versucht aus den Beobachtungen allgemeingültige Aussagen zu erhalten. Dazu verwendet sie vor allem den Schluß von einer *bekannten* Teilmasse (Stichprobe) auf die *unbekannte* Grundgesamtheit. Ihr Gegenstand ist die Schätzung von unbekannten (Verteilungs-) Parametern und das Prüfen von (Grundgesamtheits-) Hypothesen. Grundlage dazu bildet die *Wahrscheinlichkeitsrechnung*." (Hochstädter 1991, S. 6)

Die Erläuterungen zur „schließenden Statistik" sind allerdings ambivalent. Einerseits beziehen sie sich auf die Frage, wie man Wahrscheinlichkeitsaussagen über Hypothesen, die sich auf Merkmalsverteilungen in jeweils bestimmten Gesamtheiten beziehen, auf der Grundlage von unvollständigen Daten einschätzbar machen kann. Bei *diesen* Hypothesen würde es sich um die von einer deskriptiven Sozialstatistik intendierten Aussagen handeln, die nur infolge mangelhafter Daten als mehr oder weniger wahrscheinliche Hypothesen formuliert werden müssen. Andererseits soll jedoch die „schließende Statistik" dazu dienen, „allgemeingültige Aussagen" zu begründen, also Aussagen oder Hypothesen, die sich nicht auf Merkmalsverteilungen in einem abgrenzbaren Raum-Zeit-Gebiet beziehen, sondern „allgemeine Regelmäßigkeiten" zum Ausdruck bringen.[11]

4. Diese Ambivalenz durchzieht fast regelmäßig die Erläuterungen zur Inferenzstatistik in der Methodenliteratur. Es ist deshalb zweckmäßig, hier gleich zu Beginn auf eine Unterscheidung hinzuweisen. Sozialstatistik im Sinne eines empirischen Unternehmens hat es stets mit real existierenden und infolgedessen endlichen Gesamtheiten von Dingen, Menschen oder Situationen zu tun. Der grundlegende begriffliche Ansatz besteht in der Konzeption statistischer Variablen der Form

$$X : \Omega \longrightarrow \tilde{X}$$

Hierbei verweist Ω auf die Gesamtheit der Objekte oder Situationen, auf die man sich beziehen möchte, und die statistische Variable X ordnet jedem Element $\omega \in \Omega$ einen Merkmalswert $X(\omega)$ in einem Merkmalsraum $\tilde{X}$ zu. Ziel ist es, Aussagen über die statistische Verteilung von X, d.h. über die Häufigkeitsverteilung der Merkmalswerte aus $\tilde{X}$ in der Gesamtheit Ω zu machen. Alle weiteren statistischen Begriffsbildungen

[11] In einer Erläuterung, die sich an das oben angeführte Zitat anschließt, sagt Hochstädter: „Will man dagegen die Erhebungsdaten verwenden, um allgemeine Regelmäßigkeiten zu erkennen, die nicht nur für die speziellen Erhebungsfälle gelten, sondern allgemeine Gültigkeit besitzen, dann spricht man von *statistischer Inferenz*." Der hier verwendete Begriff 'statistische Inferenz' ist indessen gleichermaßen ambivalent.

knüpfen an einer solchen Konzeption statistischer Variablen an, wobei natürlich auch mehrdimensionale Variablen betrachtet werden können, wodurch Fragen nach gemeinsamen und bedingten Häufigkeitsverteilungen entstehen.[12]

5. Dieser konzeptionelle Ansatz zeigt nun auch, wo ein statistisches Inferenzproblem auftreten kann. Es tritt auf, wenn man die Werte einer statistischen Variablen nicht (oder nicht hinreichend genau) für alle Mitglieder der Gesamtheit Ω kennt, für die die Variable definiert ist. Das ist in der Sozialstatistik typischerweise dann der Fall, wenn sich die verfügbaren Daten nur auf eine Teilmenge (Stichprobe) der Mitglieder von Ω beziehen. Man kann dann über die statistischen Verteilungen der Variablen in der Gesamtheit Ω keine definitiven Aussagen machen, sondern nur Hypothesen formulieren; und es stellt sich die Frage, wie die Wahrscheinlichkeit solcher Hypothesen mit den verfügbaren Daten eingeschätzt werden kann. — In diesem Kontext kann man von einem *sozialstatistischen Inferenzproblem* sprechen. Wesentlich ist, daß man es in diesem Fall mit *definiten* Hypothesen zu tun hat, die sich auf Merkmalsverteilungen in real existierenden Gesamtheiten beziehen. Oder anders formuliert: die Hypothesen beziehen sich in diesem Fall auf *Sachverhalte*, die der Hypothesenbildung vorausgesetzt werden können.

6. Dieses sozialstatistische Inferenzproblem setzt als Kontext eine deskriptiv konzipierte Sozialstatistik voraus und wird uns deshalb im Rahmen dieses Skripts nicht näher beschäftigen. Soweit gelegentlich darauf eingegangen wird (insbesondere in Abschnitt 4.4), dient das nur zu einer Abgrenzung von Fragestellungen der theoretischen Sozialstatistik, die

[12] Der hier skizzierte Begriff statistischer Variablen kann generell als Ausgangspunkt aller statistischen Begriffsbildungen, nicht nur für die Sozialstatistik im engeren Sinne, angesehen werden. Unterschiede ergeben sich durch die Art der Objektmengen, auf die man sich beziehen möchte; daß man sich aber überhaupt auf Objektmengen bezieht, gilt für alle Anwendungen der Statistik. M. Kendall und A. Stuart beginnen ihre „Advanced Theory of Statistics“ (1977, I, S. 1) mit folgenden Worten: „The fundamental notion in statistical theory is that of the group or aggregate, a concept for which statisticians use a special word – ‘population’. This term will be generally employed to denote any collection of objects under consideration, whether animate or inanimate; for example, we shall consider populations of men, of plants, of mistakes in reading a scale, of barometric heights on different days, and even populations of ideas, such as that of the possible ways in which a hand of cards might be dealt. The notion common to all these things is that of aggregation. The science of Statistics deals with the properties of populations.“ In diesem Zitat deutet sich auch bereits an, daß der Begriff einer statistischen Gesamtheit nicht nur im Sinne einer real existierenden definiten Gesamtheit verwendet wird. Einige der daraus resultierenden Unklarheiten und Probleme werden in Kap. 4 genauer besprochen.

von grundsätzlich anderer Art sind. In der deskriptiven Sozialstatistik und bei dem aus ihrer Perspektive konzipierten Inferenzproblem geht es darum, wie Sachverhalte beschaffen sind (wobei es sich bei den Sachverhalten typischerweise um statistische Merkmalsverteilungen oder daraus ableitbare Größen handelt). Wie diese Sachverhalte zustandegekommen sind, spielt dabei keine Rolle. Theoretische Sozialstatistik beginnt dagegen mit *Fragestellungen, die sich auf das Zustandekommen und die Entwicklung sozialer Sachverhalte richten.* Man möchte Einsichten in die Prozesse gewinnen, durch die sich soziale Sachverhalte bilden und verändern. Darauf beziehen sich auch, so läßt sich jedenfalls vermuten, die „hypothetischen Gesetzmäßigkeiten“, die im Selbstverständnis der gegenwärtig betriebenen empirischen Sozialforschung eine so große Rolle spielen.

7. Um den Unterschied in der Art der Fragestellungen zu verdeutlichen, kann der Begriff eines „datenerzeugenden Prozesses“ dienen, der in der statistischen Literatur oft verwendet wird. Bemerkenswert ist die Zweideutigkeit dieses Begriffs. Einerseits kann er sich auf Prozesse beziehen, durch die man Informationen über in unserer Erfahrungswelt gegebene Sachverhalte gewinnen kann. Der Begriff bezieht sich dann darauf, wie die *Daten* – nicht jedoch die Sachverhalte, auf die sich die Daten beziehen – zustande kommen. Andererseits kann man sich mit dem Begriff darauf beziehen, wie die *Sachverhalte* (Situationen, Ereignisse, ...) entstehen, über die wir in Form von Daten mehr oder weniger viele Informationen haben. Das ist offenbar ein wichtiger Unterschied, der auch terminologisch fixiert werden sollte. Ich schlage deshalb vor, das Wort ‘datenerzeugender Prozeß’ für die erste der beiden Bedeutungen zu reservieren. Das Wort verweist dann auf Prozesse, durch die *Daten* zustande kommen, und setzt mithin voraus, daß Sachverhalte, auf die sich die Daten beziehen, in unserer Erfahrungswelt gegeben sind. In dieser Bedeutung sind datenerzeugende Prozesse für ein Verständnis des sozialstatistischen Inferenzproblems wichtig. Man muß wissen, wie Daten zustandegekommen sind, um Aussagen über die Sachverhalte machen zu können, auf die sich die Daten beziehen.

8. Eine nützliche Einteilung *daten*erzeugender Prozesse, auf die wir uns später gelegentlich beziehen werden, wurde von D. R. Cox und E. J. Snell (1981) vorgeschlagen:

> “While the same technique of analysis, e.g. regression analysis, may be used in numerically identical form for any one of the main types of study, the limitations on scientific interpretation are quite different in the different types. It is helpful to distinguish between the following:

(i) Experiments, in which the system under study is set up and controlled by the investigator. Typically, one of a number of alternative treatments is applied to each individual, or experimental unit, and responses are measured. [...]
(ii) Pure observational studies, in which data have been collected on individuals in some system, the investigator having had no control over the collection of the data, other than perhaps some role in checking the quality of the data.[13]
(iii) Sample surveys, in which a sample is drawn from a well-defined population by methods, usually involving randomization, under the investigator's control. Conclusions can be drawn with confidence about the descriptive properties of the population in question, but the interpretation of, for example, relationships between variables raises problems similar to (ii). [...]
(iv) Controlled prospective studies, in which a group of individuals, chosen by the investigator, have various explanatory variables measured and are then followed through time, often to see whether some particular event of significance (e.g. death) occurs. To the extent that all important explanatory variables can be measured, and of course this is never totally possible, these studies have some of the virtues of an experiment.
(v) Controlled retrospective studies, such as the epidemiologial investigation summarized in Example V, in which a characteristic response variable has been observed on individuals, and the history of those individuals is examined to isolate relevant explanatory variables." (Cox und Snell 1981, S. 10f)

9. Die Einteilung von Cox und Snell bezieht sich auf unterschiedliche *daten*erzeugende Prozesse. Theoretische Sozialstatistik beginnt dagegen mit einer gedanklichen Bezugnahme auf Prozesse, durch die soziale Sachverhalte, Situationen oder Ereignisse zustandekommen. Solche Prozesse erzeugen keine Daten, sondern die Sachverhalte, auf die sich Daten beziehen können. Man sollte deshalb nicht von *daten*erzeugenden Prozessen sprechen. Es handelt sich vielmehr um „soziale Prozesse“, durch die die sozialen Sachverhalte entstehen, auf die sich die Erkenntnisinteressen der Sozialstatistik richten.[14] Das Wort 'soziale Prozesse' hat allerdings keine klare Bedeutung. Es macht zwar darauf aufmerksam, daß es sich

[13]Diese Formulierung ist ersichtlich nicht sehr präzise. Es kann natürlich bestimmt werden, wie und welche Aspekte beobachtet werden sollen. Der wesentliche Unterschied zu Experimenten besteht darin, daß bei Beobachtungen kein Einfluß *auf das Zustandekommen* der zu beobachtenden Sachverhalte genommen wird.

[14]Auch wenn in der Literatur von *daten*erzeugenden Prozessen gesprochen wird, ist oft eigentlich gemeint, daß man sich auf die Prozesse beziehen möchte, durch die die Sachverhalte entstehen, über die dann mehr oder weniger viel Information in Form von Daten verfügbar ist. Als Beispiel kann auf eine Abhandlung von J. Marschak (1950) hingewiesen werden, in der erläutert wird, wie in der Ökonometrie von statistischen Modellen gesprochen wird: „Hypotheses about economic structure are also

nicht um „datenerzeugende Prozesse“ handelt; aber es verweist nicht auf Sachverhalte, die empirisch festgestellt werden könnten. Bestenfalls liefert das Wort einen Hinweis auf ein *theoretisches* Problem: Wie kann man sich explizierbare Vorstellungen darüber machen, wie soziale Sachverhalte zustande kommen?

10. Orientiert man sich an dieser Fragestellung, wird deutlich, daß eine einfache Antwort nicht gegeben werden kann. Vielmehr führt die Fragestellung unmittelbar zu konzeptionellen Fragen der Gesellschaftstheorie. Um so bemerkenswerter ist es jedoch, daß in der theoretischen Sozialstatistik Fragen der Gesellschaftstheorie fast immer ausgespart bleiben. *Stattdessen* wurden und werden Begriffsbildungen und Gedankengänge der Wahrscheinlichkeitsrechnung verwendet, um Fragen zu formulieren, die sich auf das Zustandekommen sozialer Sachverhalte beziehen. Der Grundgedanke der theoretischen Sozialstatistik, wie sie sich bisher entwickelt hat, besteht darin, soziale Sachverhalte so zu betrachten, *als ob* sie durch hypothetische Zufallsgeneratoren erzeugt würden. Soweit man von Theoriebildung sprechen kann, besteht sie darin, solche hypothetischen Zufallsgeneratoren zu konstruieren, um sie als ein Deutungsschema für soziale Prozesse zu verwenden, durch die soziale Sachverhalte entstehen.[15]

known as economic theories. They try to state relations that describe the behavior and environment of men and determine the values taken at any time by economic variables such as prices, output, and consumption of various goods and services, and the prices and amounts of various assets. [...] Also, economic variables as well as those of experimental science are, in principle, random (stochastic) variables: that is, their properties are described by probability distributions. In particular, the stochastic character of the observed data can often be ascribed to their dependence on stochastic nonobservable variables: such nonobservable variables are random 'errors' in the observation of single variables or random 'shocks' sufferend by the relations connecting them. [...] Thus, economic data are generated by systems of relations that are, in general, stochastic, dynamic, and simultaneous.“ (Marschak 1950, S. 2f) Allerdings bleibt eine Ambivalenz, auf die D. Freedman in einer Arbeit, die sich kritisch mit Anwendungen der Regressionsrechnung in den Sozialwissenschaften beschäftigt, so hingewiesen hat: „In social-science regression analysis [...] usually the idea is to fit a curve to the data, rather than figuring out the process which generated the data. As a matter of fact, investigators often talk about 'modeling the data'. This is almost perverse: surely the object is to model the phenomenon, and the data are interesting only because they contain information about that phenomenon.“ (Freedman 1985, S. 348)

[15]Einigen Statistikern ist dieses „als ob“ durchaus bewußt. Z.B. schrieb O. Anderson (1961, S. 954f): „Es unterliegt in der Tat keinem Zweifel, dass die moderne statistische Theorie sich auf einer Reihe von verschiedenen Fiktionen, verschiedenen 'Als Ob' aufbaut. [...] Solange die anzuwendenden statistischen Methoden 'stochastisch', d.h. wahrscheinlichkeits-theoretisch ausgerichtet sind, sind sie von gewissen 'Als Ob'

11. Noch eine Bemerkung zur Verwendung des Ausdrucks 'theoretische Sozialstatistik'. Wie ich diesen Ausdruck in diesem Skript verwenden möchte, wurde eben angedeutet. In den weiteren Ausführungen wird sich zeigen, daß diese Verwendung zweckmäßig ist, um zusammenfassend auf einen Komplex theoretischer Gedankengänge in der Sozialstatistik hinzuweisen, der seinen konzeptionellen Ausgangspunkt in der Annahme hat, daß man sinnvoll von Wahrscheinlichkeiten für soziale Sachverhalte und Ereignisse sprechen kann. Vielleicht genauer, aber auch umständlicher, könnte man von einer „an Vorstellungen der Wahrscheinlichkeitstheorie orientierten" Sozialstatistik sprechen. Stattdessen spreche ich kurz von theoretischer Sozialstatistik. Diese Bezeichnung erscheint jedenfalls insofern berechtigt, als ein erklärtes Ziel der Gedankengänge, auf die mit dem Ausdruck verwiesen werden soll, darin besteht, soziale Sachverhalte und Ereignisse nicht nur zu beschreiben, sondern sich gedanklich auf Prozesse ihres Zustandekommens zu beziehen. Es sollte aber darauf hingewiesen werden, daß die Bezeichnung 'theoretische Sozialstatistik' im Laufe der Zeit in unterschiedlichen Bedeutungen verwendet worden ist. Zum Beispiel unterschied G. v. Mayr zwischen einem theoretischen und einem praktischen Teil der Statistik:

> „Den *theoretischen* Theil bilden in der Hauptsache die Untersuchungen, welche das Wissensgebiet der Statistik abgrenzen, dessen allgemeine Grundlagen bestimmen und die Methode und Technik darlegen. [...] Der *praktische* Theil der statistischen Wissenschaft umfaßt die Gesammtheit der *materiellen* wissenschaftlichen Errungenschaften auf dem Gebiete der Massenbeobachtung des Gesellschaftslebens." (v. Mayr 1895, S. 123)

Dementsprechend nannte v. Mayr den ersten Band seines Hauptwerks zur Statistik und Gesellschaftslehre „Theoretische Statistik". Gerade G. v. Mayr ist jedoch nicht nur kein Anhänger, sondern eher ein Kritiker der Verwendung wahrscheinlichkeitstheoretischer Begriffsbildungen und Gedankengänge in der Sozialstatistik gewesen.[16] Ebenso gab ein anderer Kritiker der Verwendung wahrscheinlichkeitstheoretischer Überlegungen, H. Wolff, seinem Lehrbuch den Titel „Theoretische Statistik" (1932). Dagegen bezieht sich das Reden von theoretischer Statistik in diesem Skript in erster Linie auf Gedankengänge, die L. v. Bortkiewicz als 'Stochastik' bezeichnet hat:

> „Die an der Wahrscheinlichkeitstheorie orientierte, somit auf 'das Gesetz der großen Zahlen' sich gründende Betrachtung empirischer Vielheiten möge als *Stochastik* [...] bezeichnet werden. Die Stochastik ist nicht sowohl Wahrscheinlichkeitstheorie schlechthin, als vielmehr Wahrscheinlichkeitstheorie in ihrer Anwendung, sei es auf empirische Vielheiten überhaupt, sei es auf empirische Vielheiten einer bestimmten Art [wie in der Sozialstatistik]. (v. Bortkiewciz 1917, S. 3)

Dadurch wird auch die Hauptfrage kenntlich, die in diesem Skript diskutiert werden soll: welchen potentiellen Sinn wahrscheinlichkeitstheoretische Begriffsbildungen und Vorstellungen in der Sozialstatistik haben können; oder anders formuliert: ob v. Bortkiewicz Recht hat, wenn er im Anschluß an das eben angeführte Zitat behauptet:

---

kaum zu trennen, insbesondere dann, wenn sie einfach die Normalverteilung jener häufig gar nicht genau erfassbaren Ausgangsgesamtheiten ('populations') *postulieren*, aus denen die gegebenen Teilgesamtheiten 'zufällig' entstanden sein sollen. Eine der Hauptbeschäftigungen des modernen Statistikers besteht darin, einzelne Zahlen oder ganze Zahlenreihen, die tatsächlich vorliegen, mit anderen Zahlen oder Zahlenreihen zu vergleichen, die sich gemäss gewissen vorgegebenen Theorien, Hypothesen oder Modellen (wie man sich jetzt in der Ökonometrie auszudrücken pflegt) ergeben *sollten*. Das Ziel besteht darin, festzustellen, ob die Abweichungen der beobachteten von den erwarteten Zahlen noch allein *'dem Zufall'* zugeschrieben werden könnten oder ob der Unterschied bereits als hierfür zu gross erscheint. Man braucht dabei gar nicht an die Neyman-Pearson-Wald'sche Theorie der Hypothesenprüfung zu denken (die z.B. eine so grosse Autorität wie R. A. Fisher lange nicht in allen ihren Konsequenzen anerkennen will), denn jeder Vergleich von relativen Häufigkeiten, Durchschnitten und anderen Masszahlen, jede Anwendung der χ^2 Methode, jeder Versuch der Feststellung des Verteilungsgesetzes einer statistischen Variablen usw. usw. stellt uns vor dasselbe Problem." Wenn Anderson in diesen Ausführungen von Fiktionen spricht, sind Hypothesen gemeint, die sich auf Prozesse des Zustandekommens von Sachverhalten oder Ereignissen beziehen und bei denen es sich insofern um Fiktionen handelt, als sie selbst nicht empirisch geprüft werden können. Solche Fiktionen sind zwar metaphysisch, aber sie sollten nicht von vornherein als sinnlos abgetan werden. Jedenfalls sollten sie deutlich von einem gelegentlich geäußerten, aber belanglosen Vorwurf unterschieden werden, der sich darauf bezieht, daß statistische Aussagen auf Abstraktionen beruhen. Z.B. hat H. Wolff (1932, S. 129) dies so formuliert: „Alle statistischen Methoden der Beobachtung sind demnach abstraktiv; sie sind als abstraktive Methoden Fiktionen im Sinne der Vaihingerschen Fiktionentheorie." Das sollte hier erwähnt werden, weil sich auch Anderson auf Vaihingers „Philosophie des Als Ob" beruft.

[16] Hier eine seiner Stellungnahmen: „Die Wahrscheinlichkeitsrechnung hat eine innerliche Berechtigung nur bei solchen Vorgängen des sozialen Lebens, welche einige Analogie mit der Wiederholung einer Beobachtung an demselben Objekt oder der Herausnahme einer Kugel aus einer selbstbegrenzten konstanten Menge verschiedenfarbiger Kugeln in gleichbleibender Zusammensetzung haben. Solche Analogien finden sich z.B. auf dem Gebiete der Geschlechtsvertheilung der Geborenen. Darüber hinaus gibt es aber zahlreiche und gerade die wichtigsten sozialen Vorgänge und Erscheinungen, bei welchen von einem derartigen unveränderlichen Beobachtungsgebiet keine Rede ist, bei denen vielmehr gerade der stetige Wandel der Massen das Charakteristische ist. Hier vermag die Wahrscheinlichkeitsrechnung überhaupt nicht einzusetzen." (v. Mayr 1895, S. 28f)

„... daß erst die Durchdringung der Statistik mit der stochastischen Auffassungsweise ihr nicht nur einen höheren theoretischen Wert, sondern auch eine größere praktische Bedeutung verleiht.“ (ebda.)

1.3 Wahrscheinlichkeitsrechnung

1. Einen wichtigen Hintergrund für die eben angedeutete Konzeption einer theoretischen Sozialstatistik bildet die Wahrscheinlichkeitsrechnung, die sich seit dem 17. Jahrhundert, vornehmlich anhand einer Beschäftigung mit Glücksspielen entwickelt hat. Das erste größere Werk zur Wahrscheinlichkeitsrechnung stammt von dem Mathematiker Jakob Bernoulli (1654–1705).[17] Es trägt den Titel *Ars conjectandi* („Kunst der Vermutungen“) und wurde zuerst posthum im Jahr 1713 veröffentlicht. Das Buch beschäftigt sich fast ausschließlich mit Glücksspielen. Zwar ging es Bernoulli wie den anderen Autoren, die an der Entwicklung der Wahrscheinlichkeitsrechnung mitgewirkt haben, nicht in erster Linie um Glücksspiele, sondern sie waren an einer generell verwendbaren „Kunst der Vermutungen“ interessiert.[18] Ihre primäre Beschäftigung mit Glücksspielen verweist jedoch auf einen spezifischen konzeptionellen Ansatz: Gegenstand der Wahrscheinlichkeitsrechnung sind Zufallsgeneratoren, d.h. Verfahren, durch die in wiederholbarer Weise Ereignissse erzeugt werden können. Und zwar, um als Gegenstand einer mathematischen Theoriebildung dienen zu können, *artifizielle* Verfahren, die man sich selbst konstruiert hat und deren Konstruktionsprinzipien infolgedessen bekannt sind. Glücksspiele können als prototypische Beispiele für solche Zufallsgeneratoren aufgefaßt werden und sie werden bis heute zur Motivation von Begriffsbildungen und Gedankengängen der Wahrscheinlichkeitsrechnung verwendet.

2. Eine genaue Diskussion der Grundbegriffe einer aleatorisch konzipierten Wahrscheinlichkeitsrechnung soll hier nicht vorgenommen werden.[19] Nur die Definition des Begriffs ‘Zufallsgenerator’ sei kurz angedeutet. Ich verstehe darunter ein unter gleichen Bedingungen wiederholbares *Verfahren*, mit dem im Prinzip beliebig viele Sachverhalte oder Ereignisse erzeugt werden könnten, die sich durch Elemente eines vorgegebenen

[17]Einige biographische Angaben finden sich bei Johnson und Kotz 1997, S. 15ff.

[18]Vgl. Daston, 1988.

[19]Man vgl. dazu das Skript „Epistemische und aleatorische Wahrscheinlichkeit“, das in unserer Homepage verfügbar ist: `http://www.stat.ruhr-uni-bochum.de [lecture notes]`.

Merkmalsraums charakterisieren lassen. Zur formalen Notation wird die Schreibweise

$$\mathcal{G} = (\mathcal{V}, (\mathcal{S}, \tilde{X}))$$

verwendet. Der Zufallsgenerator $\mathcal{G}$ besteht aus einem Verfahren $\mathcal{V}$, mit dem beliebig viele Situationen des Typs $(\mathcal{S}, \tilde{X})$ erzeugt werden können, die ihrerseits durch jeweils genau ein Element einer Kennzeichnungsmenge, eines Merkmalsraums $\tilde{X}$, charakterisiert werden können. Außerdem soll gelten, daß vor einer Anwendung des Verfahrens mit dem gegenwärtig verfügbaren Wissen keine sicheren Voraussagen darüber gemacht werden können, welches Element des Merkmalsraums $\tilde{X}$ realisiert wird. Die Schreibweise $(\mathcal{S}, \tilde{X})$ wird verwendet, um auf den Typ der Situationen zu verweisen, die mit dem Zufallsgenerator erzeugt werden können. Die Schreibweise soll es erlauben, daß man sich beim Reden über einen Zufallsgenerator auf eine oder mehrere Situationen dieses Typs beziehen kann.[20] Hat man verstanden, wie $(\mathcal{S}, \tilde{X})$ als ein Situationstyp zur Charakterisierung jeweils bestimmter Situationen verwendet werden kann, kann man in analoger Weise auch von Situationen des Typs $(\mathcal{S}, \tilde{x})$ sprechen, wobei $\tilde{x}$ ein Element aus dem Merkmalsraum $\tilde{X}$ ist. Zu sagen, daß s eine Situation vom Typ $(\mathcal{S}, \tilde{x})$ ist, soll bedeuten, daß s eine Situation des Typs $(\mathcal{S}, \tilde{X})$ ist, die durch $\tilde{x} \in \tilde{X}$ charakterisiert werden kann. Es sei darauf hingewiesen, daß es sich bei $\tilde{x}$ auch um eine beliebige Teilmenge von $\tilde{X}$ handeln kann. Eine Situation des Typs $(\mathcal{S}, \tilde{x})$ ist dann eine Situation des Typs $(\mathcal{S}, \tilde{X})$, die durch irgendein Element aus $\tilde{x}$ charakterisiert werden kann. Somit kann man auch von einer Algebra für Situationstypen sprechen.

3. Exemplarisch kann man an eine Urne denken, die mit vollständig symmetrischen Objekten (meistens Kugeln) gefüllt ist, die sich nur durch

[20]Es sollte beachtet werden, daß $(\mathcal{S}, \tilde{X})$ selbst weder eine Situation ist noch auf eine Situation verweist. Situationen, für die wir die Symbole $s, s_1, s_2, \ldots$ verwenden, sind stets bestimmte (reale oder vorgestellte) Situationen. Der Ausdruck $(\mathcal{S}, \tilde{X})$ soll vielmehr dafür dienen, um über solche jeweils bestimmten Situationen sprechen zu können. Ich verwende die Schreibweise

$$s \models (\mathcal{S}, \tilde{X})$$

um zum Ausdruck zu bringen, daß s eine bestimmte (reale oder vorgestellte) Situations des Typs $(\mathcal{S}, \tilde{X})$ ist. Das Symbol ‘$\models$’ wird anstelle der mengentheoretischen Notation ‘$\in$’ verwendet, um zu betonen, daß ein Situationstyp nicht sinnvoll als eine Menge möglicher Situationen konzipiert werden kann. Mit einem Zufallsgenerator $(\mathcal{V}, (\mathcal{S}, \tilde{X}))$ können zwar beliebig viele Situationen des Typs $(\mathcal{S}, \tilde{X})$ hergestellt werden; aber es gibt kein Verfahren, mit dem sinnvoll eine „Menge aller möglichen Situationen dieses Typs“ definiert werden könnte.

ihre Farbe oder eine Nummer unterscheiden. Eine solche Urne kann folgendermaßen als ein Zufallsgenerator verwendet werden: zunächst werden die Kugeln in der Urne gemischt, dann wird blind eine der Kugeln aus der Urne herausgezogen. Die Farbe oder Nummer der herausgezogenen Kugel gibt an, welches der möglichen Ereignisse eingetreten ist. Schließlich wird die Kugel wieder in die Urne hineingelegt, um gleiche Ausgangsbedingungen für eine erneute Verwendung des Zufallsgenerators herzustellen. An diesem Beispiel kann auch der aleatorische Wahrscheinlichkeitsbegriff erläutert werden. Enthält die Urne insgesamt n Kugeln und haben m dieser Kugeln die Farbe F, dann ist die aleatorische Wahrscheinlichkeit für das Ziehen einer Kugel mit der Farbe F gleich m/n. Entsprechend kann allen Arten von Ereignissen, die mit einem Zufallsgenerator erzeugt werden können, eine aleatorische Wahrscheinlichkeit zugeordnet werden. Dies führt zum Begriff einer Wahrscheinlichkeitsverteilung: eine Funktion, die jedem Typ von Ereignissen, die mit einem Zufallsgenerator erzeugt werden können, eine bestimmte aleatorische Wahrscheinlichkeit zuordnet. Somit kann ein Zufallsgenerator aus der Perspektive der Wahrscheinlichkeitsrechnung vollständig und hinreichend durch eine Wahrscheinlichkeitsverteilung charakterisiert werden.

4. Man kann sicherlich sagen, daß sich die Wahrscheinlichkeitsrechnung zunächst als eine Theorie der Zufallsgeneratoren entwickelt hat. Damit stellte sich jedoch ein Problem. Die Theorie war verständlich und einsehbar, soweit sie sich mit *artifiziellen* Zufallsgeneratoren beschäftigte, deren Wahrscheinlichkeitsverteilung aus einer Kenntnis ihrer Konstruktion begründbar ist. Damit war jedoch zugleich der mögliche Anwendungsbereich der Theorie im wesentlichen auf solche artifiziellen Zufallsgeneratoren beschränkt. Wie aber konnte begründet werden, daß es für die Theorie auch außerhalb dieses Bereichs sinnvolle Anwendungsmöglichkeiten gibt?

5. Jedenfalls einige der Autoren, die an der Entwicklung der Wahrscheinlichkeitsrechnung mitgewirkt haben, waren sich dieses Problems durchaus bewußt. Dies gilt insbesondere für J. Bernoulli. In seiner *Ars conjectandi* schrieb er:

> „Wir sind also dahin gelangt, dass zur richtigen Bildung von Vermuthungen über irgend eine Sache nichts anderes zu thun erforderlich ist, als dass wir zuerst die Zahl dieser Fälle genau ermitteln und dann bestimmen, um wieviel die einen Fälle leichter als die anderen eintreten können. Und hier scheint uns gerade die Schwierigkeit zu liegen, da nur für die wenigsten Erscheinungen und fast nirgends anders als in Glücksspielen dies möglich ist; die Gücksspiele wurden aber von den ursprünglichen Erfindern, damit die Spieltheilnehmer gleiche Gewinnaussichten haben sollten, so eingerichtet, dass die Zahlen der Fälle, in welchen sich Gewinn oder Verlust ergeben muss, im voraus bestimmt und bekannt sind, und dass alle Fälle mit gleicher Leichtigkeit eintreten können. Bei den weitaus meisten andern Erscheinungen aber, welche von dem Walten der Natur oder von der Willkür der Menschen abhängen, ist dies keineswegs der Fall. So sind z.B. bei Würfeln die Zahlen der Fälle bekannt, denn es giebt für jeden einzelnen Würfel ebensoviele Fälle als er Flächen hat; alle diese Fälle sind auch gleich leicht möglich, da wegen der gleichen Gestalt aller Flächen und wegen des gleichmässig vertheilten Gewichtes des Würfels kein Grund dafür vorhanden ist, dass eine Würfelfläche leichter als eine andere fallen sollte, was der Fall sein würde, wenn die Würfelflächen verschiedene Gestalt besässen und ein Theil des Würfels aus schwererem Materiale angefertigt wäre als der andere Theil. So sind auch die Zahlen der Fälle für das Ziehen eines weissen oder eines schwarzen Steinchens aus einer Urne bekannt und können alle Steinchen auch gleich leicht gezogen werden, weil bekannt ist, wieviele Steinchen von jeder Art in der Urne vorhanden sind, und weil sich kein Grund angeben lässt, warum dieses oder jenes Steinchen leichter als irgend ein anderes gezogen werden sollte. Welcher Sterbliche könnte aber je die Anzahl der Krankheiten (d.i. ebensovieler Fälle), welche den menschlichen Körper an allen seinen Theilen und in jedem Alter befallen und den Tod herbeiführen können, ermitteln und angeben, um wieviel leichter diese als jene Krankheit, die Pest als die Wassersucht, die Wassersucht als Fieber den Menschen zugrunde richtet, um daraus eine Vermuthung über das Verhältniss von Leben und Sterben künftiger Geschlechter abzuleiten? Oder wer könnte die unzähligen Fälle von Veränderungen aufzählen, welchen die Luft täglich unterworfen ist, um daraus schon heute vermuthen zu wollen, welche Beschaffenheit sie nach einem Monate oder gar nach einem Jahre hat? [...] Da diese und ähnliche Dinge von ganz verborgenen Ursachen abhängen, welche überdies noch durch die unendliche Mannigfaltigkeit ihres Zusammenwirkens unsere Erkenntnis beständig täuschen, so würde es völlig sinnlos sein, auf diese Weise etwas erforschen zu wollen.
>
> Aber ein anderer Weg steht uns hier offen, um das Gesuchte zu finden und das, was wir *a priori* nicht bestimmen können, wenigstens *a posteriori*, d.h. aus dem Erfolge, welcher bei ähnlichen Beispielen in zahlreichen Fällen beobachtet wurde, zu ermitteln.“ (Bernoulli 1713, S. 146f)

6. Die Unterscheidung ist deutlich. Wenn man sich mit Glücksspielen oder, allgemeiner, mit Zufallsgeneratoren beschäftigt, gewinnt man Kenntnisse über Wahrscheinlichkeiten aus Einsichten in ihre Konstruktion. Es handelt sich um ein Wissen *a priori* in dem Sinne, daß man es aus einer Untersuchung von Konstruktionsverfahren gewinnen kann und nicht darauf angewiesen ist, es aus Beobachtungen realisierter Spiele zu bilden.[21] Wenn es sich jedoch andererseits um Vorkommnisse handelt, „welche von dem Walten der Natur oder von der Willkür der Menschen abhängen“, kann unser Wissen nur aus einer Beobachtung der Vorkommnisse selbst und – wenn möglich – der Prozesse, durch die sie entstehen, gewonnen werden. Zur Begründung von Erwartungen, die sich auf *mögliche* Vorkommnisse beziehen, kann man sich dann bestenfalls auf in der Vergangenheit gemachte Beobachtungen berufen und ggf. auf eine Kenntnis von Regeln für das Verhalten von Dingen und Akteuren. Und somit stellt sich die Frage, ob dafür eine Wahrscheinlichkeitstheorie nützlich sein kann, die ihre grundlegenden Begriffe und Vorstellungen aus einer Bezugnahme auf Zufallsgeneratoren gewinnt; insbesondere dann, wenn es sich um Erwartungen für mögliche Vorkommnisse handelt, die nicht durch Zufallsgeneratoren, sondern durch Interaktionsprozesse von Akteuren hervorgebracht werden.

7. Wenn es sich um Vorkommnisse handelt, die nicht durch artifizielle Zufallsgeneratoren entstehen, sondern die „von dem Walten der Natur oder von der Willkür der Menschen abhängen“, läuft der Vorschlag Bernoullis darauf hinaus, ihre Wahrscheinlichkeiten durch empirisch ermittelbare Häufigkeiten zu schätzen. Die Idee ist: man kann eine gewisse Anzahl von Situationen beobachten, in denen irgendein Merkmal auftreten oder nicht auftreten kann, dann die relativen Häufigkeiten bestimmen und sie als Schätzwerte für die unbekannten Wahrscheinlichkeiten verwenden. Zum Beispiel:

> „... wenn Jemand schon seit langen Jahren das Wetter beobachtet und sich angemerkt hat, wie oft es heiter oder regnerisch war, oder wenn Jemand zwei Spielern sehr oft zugeschaut und gesehen hat, wie oft dieser oder jener gewinnt, so kann er gerade dadurch das Verhältniss bestimmen, welches die Zahlen der Fälle, in denen dieselben Ereignisse unter den vorangegangenen gleichen Umständen auch nachher eintreten oder nicht eintreten können, wahrscheinlicher Weise zueinander haben.“ (Bernoulli 1713, S. 248)

8. Bernoulli war sich durchaus bewußt, daß die Verwendung relativer Häufigkeiten, die aus einer gewissen Anzahl von Beobachtungen ermittelt worden sind, zur Schätzung unbekannter Wahrscheinlichkeiten einer Rechtfertigung bedarf. Nachdem er das Problem anhand einiger Beispiele erläutert hat, gehen seine Überlegungen so weiter:

[21] Man sollte darauf achten, daß mit der Charakterisierung ‘a priori’ hier nicht gemeint ist, daß es sich um ein Wissen „vor aller Erfahrung“ handelt; vielmehr soll der Ausdruck in diesem Kontext darauf hinweisen, daß man dieses Wissen aus Einsichten in die Beschaffenheit bzw. Konstruktion eines Verfahrens gewinnt.

> „Man muss vielmehr noch Weiteres in Betracht ziehen, woran vielleicht Niemand bisher auch nur gedacht hat. Es bleibt nämlich noch zu untersuchen, ob durch Vermehrung der Beobachtungen beständig auch die Wahrscheinlichkeit dafür wächst, dass die Zahl der günstigen zu der Zahl der ungünstigen Beobachtungen das wahre Verhältnis erreicht, und zwar in dem Maasse, dass diese Wahrscheinlichkeit schliesslich jeden beliebigen Grad der Gewissheit übertrifft, oder ob das Problem vielmehr, so zu sagen, seine Asymptote hat, d.h. ob ein bestimmter Grad der Gewissheit, das wahre Verhältnis der Fälle gefunden zu haben, vorhanden ist, welcher auch bei beliebiger Vermehrung der Beobachtungen niemals überschritten werden kann.“ (Bernoulli 1713, S. 248f)

Bernoulli behauptet dann, daß die hier geforderte Begründung gerade durch sein „Gesetz der großen Zahlen“ geliefert werde.

9. Die Überlegung erscheint sinnvoll, wenn es sich um wiederholbare Ereignisse handelt, die man sich als durch einen Zufallsgenerator erzeugt vorstellen kann. Auf dieser Voraussetzung beruht auch die Formulierung und Begründung des Bernoullischen „Gesetzes“. Insbesondere drei Aspekte sind wichtig.

a) Nur durch die Voraussetzung eines Zufallsgenerators erhält man eine Explikation für einen quantitativ bestimmten (aleatorischen) Wahrscheinlichkeitsbegriff.

b) Die Voraussetzung ist insbesondere erforderlich, um annehmen zu können, daß es überhaupt eine numerisch bestimmte Wahrscheinlichkeit gibt. Denn ohne diese Annahme wäre ganz unverständlich, was mithilfe empirisch ermittelbarer relativer Häufigkeiten *geschätzt* werden soll.

c) Schließlich ist die Voraussetzung erforderlich, um die Situationen, auf deren Beobachtung sich die Schätzung gründen soll, als voneinander unabhängig annehmen zu können; denn auch diese Annahme ist für die Begründung des Bernoullischen „Gesetzes“ wesentlich.

Wenn man sich auf einen Zufallsgenerator beziehen kann, sind diese Voraussetzungen erfüllt, auch dann, wenn man über die genaue Beschaffenheit des Zufallsgenerators nur mangelhafte Kenntnisse hat. Man denke zum Beispiel an eine Urne, die in einem uns unbekannten Verhältnis weiße und schwarze Kugeln enthält. Infolgedessen kann eine Wahrscheinlichkeitsverteilung, die diese Urne, wenn sie als ein Zufallsgenerator verwendet wird, charakterisiert, nicht berechnet werden. Dennoch erscheint die Annahme sinnvoll, daß es ein ganz bestimmtes – wenn auch unbekanntes

– Verhältnis der beiden Sorten von Kugeln in der Urne gibt. Aber wie ist es, „wenn jemand schon seit langen Jahren das Wetter beobachtet und sich angemerkt hat, wie oft es heiter oder regnerisch war"? Kann man sich das Zustandekommen von heiteren und regnerischen Tagen so vorstellen, wie man sich das Ziehen von weißen und schwarzen Kugeln aus einer Urne vorstellen kann?

10. Auf dieses Problem hat zuerst Leibniz in einem Brief an Bernoulli hingewiesen.[22] Bernoulli hat das Problem in seiner *Ars conjectandi* angesprochen, wo er sich mit einigen Einwänden gegen seinen Gedankengang auseinandersetzt:

„... will ich kurz die Einwände widerlegen, welche einige Gelehrte dagegen erhoben haben.

1. Zuerst machen sie den Einwurf, dass das Verhältnis zwischen den Steinchen [der Kugeln in einer Urne] von anderer Beschaffenheit sei als dasjenige zwischen den Krankheiten und den Luftveränderungen; die Zahl jener sei bestimmt, die Zahl dieser aber unbestimmt und unsicher.

Darauf antworte ich: Beide sind hinsichtlich unserer Erkenntniss gleich ungewiss und unbestimmt. Dass aber irgend ein Ding an sich und seiner Natur nach ungewiss und unbestimmt beschaffen sei, kann von uns ebenso wenig verstanden werden, als wir verstehen können, dass Gott etwas zugleich erschaffen und nicht erschaffen hat; denn alles was Gott geschaffen hat, hat er gerade dadurch, dass er es geschaffen hat, auch bestimmt." (Bernoulli 1713, S. 250f)

Wenn ich es richtig verstehe, soll das Argument die Annahme begründen, daß alles, was geschieht, gesetzmäßig bestimmten Erzeugungsprozessen unterliegt. Das ist zwar eine metaphysische Annahme, aber sie läßt es sinnvoll erscheinen, durch Beobachtungen der erzeugten Sachverhalte und Ereignisse Einsichten in Beschaffenheiten der sie erzeugenden Prozesse zu gewinnen.

11. Allerdings stellt sich die Frage, wie ein solcher Glaube mit Erfahrungen historischer Kontingenz vereinbart werden kann, insbesondere bei sozialen Sachverhalten, die – wie Bernoulli sich ausgedrückt hat – auch „von der Willkür der Menschen abhängen" (vgl.o.). Es ist ja nicht nur erforderlich, daß man hinreichend viele Situationen beobachten kann, um relative Häufigkeiten zur Schätzung von Wahrscheinlichkeiten zu ermitteln; erforderlich ist außerdem die Annahme, daß sich alle Prozesse, durch die die Situationen entstehen, *durch den gleichen* Zufallsgenerator deuten lassen. Zu diesem Einwand heißt es:

[22] Ein Teil dieses Briefes findet sich bei Schneider 1988, S. 59f.

„Drittens machen sie den Einwand, dass die Zahl der Krankheiten nicht beständig dieselbe sei, sondern dass täglich neue entstehen. Darauf entgegne ich: Dass sich im Laufe der Zeiten die Krankheiten vermehren können, leugne ich nicht, und sicherlich würde derjenige, welcher aus heutigen Beobachtungen auf antediluvianische Zeiten zurückschliessen wollte, gewaltig von der Wahrheit abirren. Daraus folgt aber nichts weiter als dass bisweilen neue Beobachtungen aufgestellt werden müssen; auch bei den Steinchen würden neue Beobachtungen nothwendig werden, wenn man annehmen müsste, dass ihre Anzahl in der Urne sich geändert hätte." (Bernoulli 1713, S. 251)

Mit dieser Bemerkung wird gewissermaßen eine zweite Argumentationslinie aufgebaut. Sollte zunächst durch eine metaphysische Überlegung gerechtfertigt werden, daß sich die Erscheinungen unserer Erfahrungswelt gesetzmäßig bestimmten Erzeugungsprozessen verdanken, wird jetzt eingeräumt, daß diese Erzeugungsprozesse selbst einem historischen Wandel unterliegen. Bernoullis Vorschlag, sich gelegentlich neue Daten zu verschaffen, um Eigenschaften der hypothetisch unterstellten Zufallsgeneratoren neu zu schätzen, liefert gewissermaßen eine pragmatische Ergänzung zum vorab betonten Glauben an Gesetzmäßigkeiten.[23]

1.4 Orientierung und Spekulation

1. Bernoullis Überlegungen sind vor allem deshalb interessant, weil in ihnen ein Spannungsverhältnis von zwei Gedankengängen zum Ausdruck kommt, das die Geschichte der theoretischen Sozialstatistik bis heute durchzieht. Einerseits ein Glaube an „ewige Gesetze", denen die Entwicklung allen Geschehens unterliegt (zu Bernoullis Zeiten prototypisch exemplifiziert an den Bewegungsvorgängen des Planetensystems und der Newtonschen Physik);[24] andererseits eine Orientierung an den kontingenten Entwicklungen einer von Akteuren hervorgebrachten Geschichte. Dementsprechend vermischen sich in der theoretischen Sozialstatistik zwei unterschiedliche Problemstellungen. Einerseits die Frage,

[23] So hat wohl auch Leibniz die Angelegenheit betrachtet, als er in einem Brief an Bernoulli (1703) schreibt: „Wenn sich aber auch empirisch keine vollkommen richtige Schätzung erzielen läßt, wird die empirische Schätzung nichtsdestotrotz in der Praxis brauchbar und ausreichend sein." (Zitiert bei Schneider 1988, S. 60.)

[24] P. S. de Laplace (1749–1827), der für die Entwicklung der Wahrscheinlichkeitsrechnung und ihrer Anwendungen besonders einflußreich geworden ist, schrieb in seinem *Essai philosophique sur les probabilités* (1814, S. 3): „Die Regelmäßigkeit, welche uns die Astronomie in der Bewegung der Kometen zeigt, ist ohne Zweifel bei allen Erscheinungen vorhanden."

wie man sich theoretische Vorstellungen über die Prozesse machen kann, durch die sich gesellschaftliche Verhältnisse entwickeln, andererseits Fragen, die sich darauf richten, wie man Beobachtungen gesellschaftlicher Erscheinungen für pragmatisch brauchbare Prognosen nutzen kann.

2. Vergleichsweise leicht verständlich ist die pragmatische Orientierung, die auf Voraussagen und Erwartungsbildungen zielt. Man kann sie sich an beliebig vielen Beispielen aus dem praktischen Leben verdeutlichen. Es ist auch verständlich, daß statistische Daten für Erwartungsbildungen hilfreich sein können. Zum Beispiel kann eine statistische Erhebung Daten liefern, um einschätzbar zu machen, wie sich die in den Schulen zu erwartenden Schülerzahlen in einem gewissen Zeitraum entwickeln werden. Das in der Geschichte der Sozialstatistik zunächst wichtigste Beispiel betrifft die Konstruktion von Sterbetafeln, die von Lebensversicherungsgesellschaften für Erwartungsbildungen über Lebensdauern und Sterbefälle verwendet wurden. Erwartungsbildungen dieser Art erfordern keine Gesellschaftstheorie, sondern verdanken sich praktischer Kenntnisse sozialer Prozesse. Zum Beispiel benötigt ein Kaufmann keine Gesellschaftstheorie, um sich Erwartungen über die Anzahl der Kunden zu bilden, die vielleicht in sein Geschäft kommen werden. Die Ermittlung statistischer Daten darüber, wie es bisher gewesen ist, kann gleichwohl dazu dienen, Informationen, die für Erwartungsbildungen eine Rolle spielen können, zugänglich, explizit überschaubar und einschätzbar zu machen.

3. Würde sich die Sozialstatistik darauf beschränken, Daten verfügbar zu machen, die für pragmatische Orientierungen nützlich sein können, könnte man sich wissenschaftstheoretische Überlegungen sparen. Sozialstatistik wäre dann ein Hilfsmittel für praktische Orientierungen und hätte im übrigen keinerlei gesellschaftstheoretische Implikationen oder Ansprüche. Tatsache ist jedoch, daß die Sozialstatistik in ihrer historischen Entwicklung – bis hin zur gegenwärtigen empirischen Sozialforschung – immer wieder auch mit sozialwissenschaftlichen Ansprüchen aufgetreten und zu einem Ausgangspunkt für gesellschaftstheoretische Spekulationen geworden ist. Es sind diese Ansprüche und Spekulationen, durch die wissenschaftstheoretische und ideologiekritische Überlegungen herausgefordert werden.

4. Ein Verständnis der Wahrscheinlichkeitsrechnung ist hierfür wichtig, weil sie das grundlegende Denkmuster geliefert hat, das in der theoretischen Sozialstatistik zur Konzeption sozialer Sachverhalte und Ereignisse verwendet wird: die grundlegende Annahme besteht darin, daß

es für das Zustandekommen sozialer Sachverhalte und Ereignisse Wahrscheinlichkeiten gibt. Das aber impliziert, wenn man dem Selbstverständnis der Wahrscheinlichkeitsrechnung folgt, daß soziale Sachverhalte und Ereignisse so betrachtet werden, als ob sie durch Zufallsgeneratoren hervorgebracht werden.[25] Im vorangegangenen Abschnitt wurde gezeigt, wie diese Betrachtungsweise von Bernoulli durch eine metaphysische Überlegung zu rechtfertigen versucht wurde: „alles was Gott geschaffen hat, hat er gerade dadurch, dass er es geschaffen hat, auch bestimmt.“ Gott, oder die Natur, hat gewissermaßen die Zufallsgeneratoren eingerichtet, durch die die Sachverhalte und Ereignisse in unserer Erfahrungswelt entstehen. Ganz ähnlich hat sich auch Abraham de Moivre ausgedrückt, der wie Bernoulli eine maßgebliche Rolle in der Entwicklung der Wahrscheinlichkeitsrechnung gespielt hat:

> “As, upon the Supposition of a certain determinate Law according to which any Event is to happen, we demonstrate that the Ratio of Happenings will continually approach to that Law, as the Experiments or Observations are multiplied: so, *conversely*, if from numberless Observations we find the Ratio of the Events to converge to a determinate quantity, as to the Ratio of P to Q; then we conclude that this Ratio expresses the determinate Law according to which the Event is to happen. For let that Law be expressed not by the Ratio $P : Q$, but by some other, as $R : S$; then would the Ratio of the Events converge to this last, not to the former: which contradicts our *Hypothesis*. And the like, or greater, Absurdity follows, if we should suppose the Event not to happen according to any Law, but in a manner altogether desultory and uncertain; for then the Events would converge to no fixt Ratio at all.
>
> Again, as it is thus demonstrable that there are, in the constitution of things, certain Laws according to which Events happen, it is no less evident from Observation, that those Laws serve to wise, useful and beneficent purposes; to preserve the stedfast Order of the Universe, to propagate the several Species and Beings, and furnish to the sentient Kind such degrees of happiness as are suited to their State.
>
> But such Laws, as well as the original Design and Purpose of their Establishment, must all be *from without*; the *Inertia* of matter, and the nature of all created Beings, rendering it impossible that any thing should modify its own essence, or give to itself, or to any thing else, an original determination

[25] Dies entspricht dem traditionellen Verständnis der Denkvoraussetzungen für Anwendungen des Wahrscheinlichkeitsbegriffs in der Sozialstatistik. Mit Auffassungen, die hypothetische Zufallsgeneratoren durch „degrees of belief“ explizieren, werden wir uns in diesem Skript nicht näher auseinandersetzen. Zahlreiche historische Bemerkungen darüber, wie sich das traditionelle Verständnis von Wahrscheinlichkeit mit statistischen Fragen verbunden hat, finden sich bei John, 1895.

or propensity. And hence, if we blind no ourselves with metaphysical dust, we shall be led, by a short and obvious way, to the acknowledgment of the great *Maker* and *Governour* of all; *Himself all-wise, all-powerful* and *good*." (de Moivre 1756, S. 251f)

5. Erst im 19. Jahrhundert sind indessen Begriffsbildungen und Gedankengänge der Wahrscheinlichkeitsrechnung verwendet worden, um eine bis dahin weitgehend deskriptiv orientierte Sozialstatistik explizit mit gesellschaftstheoretischen Spekulationen zu verknüpfen. Eine besonders wichtige Rolle spielte Adolph Quetelet (1796–1874), insbesondere sein auch in andere Sprachen übersetztes Buch „Über den Menschen und die Entwicklung seiner Fähigkeiten, oder Versuch einer Physik der Gesellschaft" (1835, dt. Übersetzung 1838).[26] A. A. Tschuprow, ein wichtiger Mitbegründer der theoretischen Sozialstatistik,[27] hat folgende Einordnung vorgenommen:

„Die uns als eine beschreibende Wissenschaft von den Massenerscheinungen des menschlichen Lebens entgegentretende Statistik hat bekanntlich eine dreifache Wurzel in der Geschichte. Die Staatenbeschreibung der Sansovino, Botero u.a. hat ihr durch die Vermittlung der deutschen 'Universitätsstatistik' den Charakter einer beschreibenden Wissenschaft aufgeprägt und in dem Begriffe der Staatsmerkwürdigkeiten als des Gegenstandes ihrer beschreibenden Tätigkeit ihre formale Grundlage zugewiesen. Die politische Arithmetik hat den Begriff der Merkwürdigkeit nach längerer Opposition seitens der zum Geographischen neigenden Vertreter der Wissenschaft mit dem gegenwärtig allgemein anerkannten Inhalte ausgefüllt, indem sie das Zahlenmäßige in den Massenerscheinungen in dessen Gesetzmäßigkeit erkannte und in den Vordergrund der statistischen Betrachtung geschoben hat. Endlich hat die Wahrscheinlichkeitslehre der Vorstellung der in den Massenerscheinungen des menschlichen Lebens sich offenbarenden Gesetzmäßigkeit das theoretische Rückgrat verliehen und dem synthetischen Genie Quetelets gestattet, die Statistik als eine Physique sociale zu einem abgeschlossenen wissenschaftlichen System zu gestalten." (Tschuprow 1906, S. 647f)

Quetelets zentraler Gedankengang war: Die Entwicklung der Individuen und ihres gesellschaftlichen Lebens erfolgt aufgrund von Gesetzmäßigkeiten; und diese Gesetzmäßigkeiten lassen sich erkennen, indem mithilfe statistischer Methoden ein sog. „mittlerer Mensch" konstruiert und untersucht wird.[28] Allerdings sind Quetelets Ausführungen sehr unklar und bestehen weitgehend nur in metaphorischen Bezugnahmen auf Gedankengänge der Wahrscheinlichkeitsrechnung.[29] Es erscheint deshalb ausreichend, wenn im weiteren besprochen wird, wie spätere Autoren an ihn angeknüpft haben.

[26]Einige biographische Hinweise finden sich bei Johnson und Kotz 1997, S. 64ff, sowie bei Lippert, 1910.

[27]Tschuprow lebte 1874 – 1926. Hinweise auf seine Biographie und Arbeiten finden sich bei Ondar, 1981.

[28]Im „Handwörterbuch der Staatswissenschaften" (Lippert, 1910) wird Quetelet so charakterisiert: „Gründer der modernen Sozialstatistik durch seine Definition der Statistik als der 'Wissenschaft von den Gesetzen, die das menschliche Leben ordnen und beherrschen'."

[29]Vgl. Porter, 1985 u. 1986, S. 41ff.

Kapitel 2

Hypothetische Zufallsgeneratoren

Theoretische Sozialstatistik, so wie dies Wort in der Einleitung erläutert wurde, beginnt mit der Annahme, daß die Prozesse, durch die soziale Sachverhalte und Ereignisse entstehen, als Realisationen von Zufallsgeneratoren gedeutet werden können. In diesem Kapitel soll diese Idee etwas genauer besprochen werden.

2.1 Verteilungsgesetze

1. Als Ausgangspunkt kann das Wort 'Verteilungsgesetz' dienen, das in der statistischen Literatur sehr oft angetroffen werden kann. Soweit mit dem Wort auf den Begriff einer statistischen Verteilung verwiesen werden soll, ist es leicht verständlich. Denken wir an eine statistische Variable

$$X : \Omega \longrightarrow \tilde{X}$$

Ω ist eine Menge von Objekten oder Situationen; $\tilde{X}$ ist eine Menge von Merkmalswerten (Attributen), durch die die Mitglieder von Ω charakterisiert werden können; die statistische Variable X ist eine Abbildung, die jedem Element aus Ω seine Charakterisierung in $\tilde{X}$ zuordnet. Eine solche statistische Variable kann folglich durch eine Häufigkeitsverteilung

$$h : \mathcal{P}(\tilde{X}) \longrightarrow \mathbf{R}$$

charakterisiert werden, die jeder Teilmenge $\tilde{x} \subseteq \tilde{X}$ (jedem Element der Potenzmenge $\mathcal{P}(\tilde{X})$) die relative Häufigkeit $h(\tilde{x})$ zuordnet, mit der Merkmalswerte aus $\tilde{x}$ in der Gesamtheit Ω vorkommen.[1] Die Funktion h zeigt die *Verteilung* der durch $\tilde{X}$ gegebenen Merkmale in der Gesamtheit Ω und wird dementsprechend eine *Häufigkeits-* oder *Verteilungsfunktion* genannt.

[1] In diesem Text verwende ich das Symbol $\mathcal{P}$ für Potenzmengen, das Symbol $\mathbf{R}$ soll auf die Menge der reellen Zahlen verweisen. Natürlich sind relative Häufigkeiten auf die Zahlen zwischen 0 und 1 beschränkt.

2. Bei einer Verteilungsfunktion handelt es sich allerdings nicht um ein Gesetz. Eine Verteilungsfunktion zeigt, mit welchen Häufigkeiten eine Reihe vorgegebener Merkmale in einer Gesamtheit von Dingen, Menschen oder Situationen auftreten. Sie liefert eine spezifische (statistische) Beschreibung einer Menge von Dingen oder Situationen. Das Wort ‘Gesetz’ verweist hingegen auf Unterstellungen, die mit dem Begriff einer Verteilungsfunktion nicht unmittelbar verknüpft werden können. Es vermittelt die Konnotation, daß irgendetwas der Fall sein muß. Aber das ist bei Häufigkeitsverteilungen, wie sie in der Sozialstatistik ermittelt werden, typischerweise nicht der Fall. Solche Häufigkeitsverteilungen können sich nicht nur verändern; wir wissen auch, daß sie sich im Zeitablauf ändern. — Warum sprechen Statistiker gleichwohl oft von einem Verteilungs*gesetz*?

3. Der Sprachgebrauch hat unterschiedliche geistesgeschichtliche Hintergründe. Eine der Traditionslinien verweist auf die Wahrscheinlichkeitstheorie. Denkt man an ihre historische Entwicklung, hat sie sich zunächst als eine Theorie der Zufallsgeneratoren entwickelt (vgl. Abschnitt 1.3). Wie statistische Variablen können indessen auch Zufallsgeneratoren durch Verteilungsfunktionen charakterisiert werden. Zwar beziehen sich Verteilungsfunktionen dann nicht unmittelbar auf relative Häufigkeiten in statistischen Kollektiven, sondern auf Realisierungsmöglichkeiten von Sachverhalten bzw. auf Wahrscheinlichkeiten; man denke z.B. an Zufallsgeneratoren in Gestalt von Würfeln oder Urnen, aus denen Kugeln gezogen werden. Es macht aber auch in diesem Kontext Sinn, von einer Verteilung, genauer von einer *Wahrscheinlichkeitsverteilung*, zu sprechen. Jeder Zufallsgenerator kann durch eine bestimmte Wahrscheinlichkeitsverteilung charakterisiert werden. Dadurch wird in einer abstrakten Weise zum Ausdruck gebracht, wie der Zufallsgenerator funktioniert. Somit wird auch verständlich, in welcher Weise in der Wahrscheinlichkeitstheorie von einem Verteilungsgesetz gesprochen wird: das Wort bezieht sich in diesem Kontext auf die Funktionsweise eines Zufallsgenerators; diese Funktionsweise wird durch die Angabe einer Wahrscheinlichkeitsverteilung charakterisiert. Insofern gibt in diesem Zusammenhang eine Wahrscheinlichkeitsverteilung das „Gesetz“ an, nach dem ein Zufallsgenerator funktioniert.

4. Aber liefert diese Verwendung des Wortes ‘Verteilungsgesetz’ zur Charakterisierung von Zufallsgeneratoren auch einen Zugang zum Verständnis entsprechender Redeweisen in der Statistik? Man kann versuchen, einen Zusammenhang herzustellen. Angenommen, daß man sich auf einen Zufallsgenerator beziehen kann. Dann kann man ihn verwenden, um eine Menge von Situationen zu erzeugen. Zum Beispiel kann man einen Würfel nehmen und n-mal würfeln. Als Ergebnis erhält man ein statistisches Kollektiv, eine Menge Ω_n, die aus n Situationen besteht, die mithilfe des Zufallsgenerators erzeugt worden sind. Dazu korrespondierend kann man eine statistische Variable X_n definieren, um die in den einzelnen Situationen realisierten Merkmalswerte zu repräsentieren. Schließlich kann man die Häufigkeitsverteilung von X_n berechnen und erhält eine *statistische* Verteilung. Sie entspricht zwar nicht genau der Wahrscheinlichkeitsverteilung, die den Zufallsgenerator charakterisiert. Aber wie in der Wahrscheinlichkeitstheorie gezeigt wird, kann man gleichwohl Wahrscheinlichkeitsaussagen machen, die sich auf den Zusammenhang zwischen der Wahrscheinlichkeitsverteilung eines Zufallsgenerators und statistischen Verteilungen in Gesamtheiten von Situationen, die mit dem Zufallsgenerator erzeugt werden, beziehen.

5. Stellt man auf diese Weise einen Zusammenhang her, wird dadurch zwar die in Ω_n ermittelte statistische Verteilung nicht zu einem Verteilungsgesetz. Aber man hat auf eine nachvollziehbare Weise einen Zusammenhang zwischen einer statistischen Verteilung und einem Verteilungsgesetz hergestellt. Das Verteilungsgesetz charakterisiert den Zufallsgenerator. Die für die Herstellung des Zusammenhangs entscheidende Voraussetzung besteht darin, daß die Mitglieder eines statistischen Kollektivs, in unserem Beispiel Ω_n, *durch einen Zufallsgenerator erzeugt* worden sind. Dies macht es in gewisser Weise möglich, die in Ω_n ermittelte statistische Verteilung durch ein Verteilungsgesetz zu erklären. Es handelt sich um eine Erklärung, weil bzw. insoweit man gewisse Einsichten in einen Prozeß gewinnt, durch den der zu erklärende Sachverhalt (in diesem Fall eine statistische Verteilung) zustandegekommen ist.

6. Es handelt sich allerdings um eine sehr spezifische Art der Erklärung. Man kann von einer *probabilistischen* Erklärung sprechen. Eine solche Erklärung liefert kein Wissen darüber, wie in jedem Einzelfall der Prozeß abläuft, der zu einem bestimmten Ergebnis führt; und insofern vermittelt eine probabilistische Erklärung strenggenommen überhaupt keine Einsichten in die Beschaffenheit der Prozesse, die die Eigenschaften der Mitglieder eines statistischen Kollektivs hervorbringen. Dennoch kann man sagen, daß man eine Art von Erklärung für eine statistische Verteilung bekommt.

7. Eine andere Frage ist, ob solche Erklärungen wissenswert sind. Man kann vielleicht sagen, daß sie ausreichend sind, wenn man nur an statistischen Verteilungen interessiert ist, so wie dies z.B. von J. K. Lindsey (1973, S. 1) in folgender Weise zum Ausdruck gebracht worden ist:

"The sociologist is not concerned with determining how an individual will react in a given set of circumstances. Instead, more general relationships are developed to show that, under given circumstances, a certain proportion of individuals will do this, another proportion something else and so on, until all of the common responses or attributes have been included."

Es ist allerdings keineswegs selbstverständlich, soziologische Erkenntnisinteressen von vornherein auf statistische Verteilungen einzuschränken. Schon die eben bemerkte Tatsache, daß probabilistische Erklärungen keinerlei Einsichten in Prozesse vermitteln, durch die die Eigenschaften der Mitglieder eines statistischen Kollektivs zustande kommen, liefert ein Motiv, sich auch um ein Verständnis solcher Prozesse zu bemühen.

8. Aber selbst wenn man das von Lindsey formulierte Erkenntnisinteresse akzeptiert, bleibt eine grundlegende Frage noch offen. Probabilistische Erklärungen setzen voraus, daß man sich auf Zufallsgeneratoren beziehen kann, durch die statistische Kollektive zustande kommen. *Wenn* dies möglich ist, kann man sich auf ein Verteilungsgesetz berufen, um in abstrakter Weise den Prozeß zu charakterisieren, durch den die Elemente eines statistischen Kollektivs zustande kommen bzw. jeweils bestimmte Eigenschaften bekommen; und der Hinweis auf ein solches Verteilungsgesetz liefert dann eine Art von Erklärung für eine statistische Verteilung. Aber unter welchen Umständen und mit welcher Berechtigung kann man voraussetzen, daß eine Menge von Dingen oder Situationen durch einen Zufallsgenerator zustandegekommen ist? In der Wahrscheinlichkeitstheorie braucht diese Frage nicht gestellt zu werden, da man sich dort mit *artifiziellen* Zufallsgeneratoren beschäftigt.[2] Die Frage bekommt aber eine wichtige und kritische Bedeutung, wenn man Gedankengänge und Begriffsbildungen der Wahrscheinlichkeitstheorie für eine theoretische Sozialstatistik verwenden möchte. Denn es ist offensichtlich, daß die Kollektive der Sozialstatistik und die Eigenschaften ihrer Mitglieder nicht in irgendeinem realistischen Sinne des Wortes durch Zufallsgeneratoren entstehen. Welchen Sinn könnte es haben, gleichwohl *hypothetische* Zufallsgeneratoren anzunehmen, durch die Mitglieder sozialstatistischer Kollektive und ihre Eigenschaften entstanden sein könnten?

[2]Dabei kann es sich einerseits um praktisch realisierbare Verfahren handelt, z.B. Glücksspiele, oder um rein begriffliche Konstruktionen, wie sie in der mathematischen Wahrscheinlichkeitstheorie ausgedacht werden.

2.2 Die Kollektivmaßlehre

1. Die Idee, statistische Kollektive durch hypothetische Zufallsgeneratoren zu deuten, kann, wie schon in Abschnitt 1.2 angedeutet worden ist, als Grundidee der theoretischen Sozialstatistik betrachtet werden. In Andeutungen und Beispielen kann man sie bereits bei J. Bernoulli, de Moivre, Laplace, Quetelet und anderen Autoren finden, die die Entwicklung einer theoretischen Sozialstatistik vorbereitet haben. Als ein Beispiel dafür, wie diese Grundidee in der Frühphase der Entwicklung einer theoretischen Sozialstatistik verfolgt worden ist, beschäftigen wir uns in diesem Abschnitt mit Fechners Kollektivmaßlehre. Gustav Theodor Fechner lebte 1801 – 1887 und ist hauptsächlich als Begründer der Psychophysik bekannt geworden. Sein statistisches Hauptwerkt, die „Kollektivmaßlehre", wurde posthum 1897 von G. F. Lipps herausgegeben und hatte in den folgenden Jahrzehnten einen wichtigen Einfluß auf die Entwicklung der Statistik in Deutschland.[3]

2. Um das Thema seiner Kollektivmaßlehre anzugeben, beginnt Fechner mit folgender Definition:

„Unter einem Kollektivgegenstand (kurz K.-G.) verstehe ich einen Gegenstand, der aus unbestimmt vielen, nach Zufall variierenden, Exemplaren besteht, die durch einen Art- oder Gattungsbegriff zusammengehalten werden.

So bildet der Mensch einen Kollektivgegenstand im weiteren Sinne, der Mensch von bestimmtem Geschlechte, bestimmtem Alter und bestimmter Rasse einen solchen im engeren Sinne, wie überhaupt das, was man den Umfang eines K.-G. nennen kann, sich nach der Ausdehnung des Gattungs- oder Artbegriffs, unter den er tritt, ändert.

Die Exemplare eines K.-G. können räumlich oder zeitlich verschieden sein und hiernach einen räumlichen oder zeitlichen K.-G. bilden. So können die Rekruten eines Landes oder Ähren eines Kornfeldes als Exemplare eines räumlichen K.-G. gelten. Soe giebt die (mittlere) Temperatur des 1. Januar, an einem gegebenen Orte durch eine Anzahl von Jahren verfolgt, ebenso viele Exemplare eines zeitlichen K.-G. Statt des 1. Januar kann man jeden anderen Jahrestag, statt eines bestimmten Tages einen bestimmten Monat, statt der Temperatur den Barometerstand setzen u.s.w. und wird damit Exemplare eben so vieler zeitlicher K.-G. erhalten." (Fechner 1897, S. 3)

Warum interessiert sich Fechner für solche Kollektivgegenstände? Er gibt zunächst folgende Antwort:

[3]Hinweise auf Fechners Einfluß finden sich bei Heidelberger, 1987.

„Anthropologie, Zoologie, Botanik haben es überhaupt wesentlich mit K.-G. zu thun, da es sich darin nicht um eine Charakterisierung einzelner Exemplare, sondern nur um das handeln kann, was einer Gesamtheit derselben zukommt, die aus dem oder jenem Gesichtspunkte als Gattung oder Art in größerer oder geringerer Weite zusammengefasst wird. Die Meteorologie bietet nach eben angeführten Beispielen in ihren nicht periodischen Witterungsphänomenen zahlreiche Beispiele davon dar; und selbst in der Artistik lässt sich von solchen sprechen, sofern Bücher, Visitenkarten darunter gehören." (Fechner 1897, S. 3f)

Es geht also zunächst darum, deskriptives Wissen über Gesamtheiten von Dingen zu gewinnen. Das kommt deutlich auch in folgender Bemerkung zum Ausdruck:

„Gelte es z.B. Rekruten eines gegebenen Landes, so fragt es sich: wie groß sind die Rekruten im Mittel, wie stark schwanken die einzelnen Maße um ihr Mittel, wie groß sind die größten und kleinsten, wie verhalten sich die Rekrutenmaße nach diesen Bestimmungen in den einzelnen Jahrgängen, wie in verschiedenen Ländern unter einander. Solche und damit zusammenhängende, später zu betrachtende Fragen lassen sich bei jedem K.-G. aufwerfen." (Fechner 1897, S. 4)

Soweit handelt es sich bei Fechners Kollektivmaßlehre um deskriptive Statistik. Es gibt jedoch zwei Gesichtspunkte, die darüber hinausweisen.

3. Der erste Gesichtspunkt kommt bereits in der oben zuerst zitierten Definition zum Ausdruck: ein Kollektivgegenstand im Sinne Fechners besteht aus *unbestimmt vielen* Elementen. Dem korrespondiert die Formulierung, daß ein K.-G. „durch einen Art- oder Gattungsbegriff zusammengehalten" wird. Ein Kollektivgegenstand in diesem Sinne unterscheidet sich offensichtlich von einer statistischen Gesamtheit, die stets aus einer bestimmten und mithin endlichen Gesamtheit von Dingen oder Situationen in unserer Erfahrungswelt besteht. Es handelt sich in diesem Fall um *definite* Gesamtheiten; das ist die entscheidende Voraussetzung, um sich empirisch auf eine Gesamtheit beziehen und statistische Variablen definieren zu können. Dagegen sind Kollektivgegenstände im Sinne Fechners *indefinite* Gesamtheiten, die aus einer *unbestimmten* Anzahl von Elementen bestehen. Zugrunde liegt eine Idee, die sich bis heute in der statistischen Literatur findet: daß man sich ausgehend von Art- oder Gattungsbegriffen Mengen bilden könne, die alles umfassen, was unter den jeweiligen Begriff fallen *kann*. Zum Beispiel eine „Menge aller Menschen" oder eine „Menge aller Rekruten" oder eine „Menge aller Visitenkarten". Fechner weiß natürlich, daß man nicht unmittelbar empirisch über solche indefiniten Mengen sprechen kann. Empirisches Wissen kann man stets nur über definite Gesamtheiten gewinnen, deren Elemente sich in unserer Erfahrungswelt fixieren lassen. Aber sein Erkenntnisinteresse zielt dennoch auf die indefiniten Mengen, die den Gattungsbegriffen korrespondieren. Indem Fechner einige Menschen untersucht, möchte er etwas über die „Menge aller Menschen" herausfinden, Einsichten in die „Gattung Mensch" gewinnen.

4. Man kann natürlich bezweifeln, daß das möglich ist. Angenommen man hat Informationen über die Lebensdauern bei einer bestimmten Menge von Menschen, die in der Vergangenheit gelebt haben. Mit welcher Berechtigung kann man solche Informationen verwenden, um Aussagen über die Lebensdauern der Mitglieder einer indefiniten „Menge aller Menschen" zu machen? Zu einer solchen Menge (wenn es sich überhaupt um eine nachvollziehbare Begriffsbildung handeln würde) würden ja auch alle Menschen gehören, die erst in der Zukunft geboren werden; aber was können wir heute schon über die Lebensdauern von Menschen wissen, die erst in einigen 100 Jahren geboren werden?

5. Fechner erörtert zwar Fragen dieser Art nicht, er führt jedoch einen zweiten Gesichtspunkt ein, der eine scheinbare Lösung des Problems vorstellbar machen soll.

„Über alle diese Einzelfragen [wie in den oben angeführten Zitaten angedeutet worden ist] erhebt sich eine allgemeinere, die wichtigste, um die es sich überhaupt in dieser Lehre handeln kann und demgemäß im folgenden handeln wird, die Frage nach dem Gesetze, wie sich die Exemplare eines K.-G. nach Maß und Zahl verteilen. Unter dem Ausdruck *Verteilung* aber ist die Bestimmung zu verstehen, wie sich die Zahl der Exemplare eines gegebenen K.-G. mit ihrer Größe ändert. Bei jedem, in einer größeren Zahl von Exemplaren vorhandenen K.-G. kommen die kleinsten und größten Exemplare, kurz Extreme, am seltensten vor, am häufigsten solche von einer gewissen mittleren Größe. Aber giebt es nicht ein allgemeines, auf alle oder wenigstens die meisten K.-G. anwendbares Gesetz der Abhängigkeit der Zahl von der Größe der Exemplare? In der That wird sich ein solches aufstellen lassen, und eine Hauptaufgabe des folgenden auf seine Feststellung gehen." (Fechner 1897, S. 4f)

Die Idee besteht also darin, daß man bei (jedenfalls einigen) Kollektivgegenständen die Existenz eines Verteilungsgesetzes annehmen kann, das die Häufigkeitsverteilung der Merkmalsausprägungen „beherrscht".[4]

6. Zum besseren Verständnis sollte man darauf achten, daß Fechner das Wort 'Verteilung' in zwei unterschiedlichen Kontexten verwendet.

[4] Diese Formulierung, daß Verteilungsgesetze Merkmalsausprägungen „beherrschen", durchzieht die statistische Literatur; Fechner verwendet sie z.B. auf S. 5 seines Buches.

Zunächst in einem empirischen Kontext, in dem man sich auf eine definite Gesamtheit beziehen kann. Oder anders gesagt: man kann sich dann auf eine bestimmte Menge von Daten beziehen, z.B. auf die Körpergrößen von 1000 Rekruten oder auf 100 Zeitpunkte, zu denen man eine Temperatur festgestellt hat. Für diesen Kontext kann die in Abschnitt 2.1 eingeführte Notation für statistische Variablen verwendet werden:

$$X : \Omega \longrightarrow \tilde{X}$$

wobei Ω auf eine definite Gesamtheit verweist und die statistische Variable X jedem ihrer Mitglieder einen bestimmten Merkmalswert in $\tilde{X}$ zuordnet. Fechners Begriff einer Verteilung entspricht in diesem Kontext genau unserem Begriff einer statistischen Häufigkeitsverteilung. Eine solche Häufigkeitsverteilung ist aber natürlich kein Verteilungsgesetz; sie bringt nur einen empirisch ermittelten Sachverhalt zum Ausdruck. Dem entspricht, daß in diesem Kontext die Menge Ω kein Kollektivgegenstand im Sinne Fechners ist. Ω besteht zum Beispiel aus 1000 Rekruten, aber der entsprechende Kollektivgegenstand im Sinne Fechners würde aus „allen Rekruten“ bestehen, aus der Gesamtheit aller ..., die unter den Gattungsbegriff ‘Rekrut’ fallen.

7. Fechner verwendet indessen das Wort ‘Verteilung’ in begrifflich analoger Weise nicht nur für definite Gesamtheiten, sondern auch für seine Kollektivgegenstände. Darin besteht ja die Idee: daß man sich indefinite Mengen als durch Gattungsbegriffe definierbare Mengen in Analogie zu definiten Mengen vorstellen könne. Würde man dieser Idee folgen, könnte man also unsere Notation übernehmen und Ω als symbolischen Verweis auf eine indefinite Menge verwenden, die durch einen Gattungsbegriff definiert wird. Eine sich hieran anschließende Konzeption statistischer Variablen würde schließlich den Eindruck erwecken, daß es für jedes Mitglied eines Kollektivgegenstandes einen bestimmten Merkmalswert gibt (also auch z.B. für Rekruten, die heute noch nicht geboren sind), und würde es erlauben, von einer Merkmalsverteilung innerhalb des Kollektivgegenstands zu reden. Dies ist der zweite Kontext, in dem Fechner von Verteilungen spricht, und auf diesen Kontext bezieht sich auch seine Vorstellung von Verteilungs*gesetzen*.

8. Die Begriffsbildungen, die den zweiten Kontext erzeugen, in dem Fechner von Verteilungsgesetzen spricht, sind ersichtlich obskur. Hier setzt indessen der Versuch ein, diese begrifflichen Schwierigkeiten durch die Annahme zu vermeiden, daß die Elemente eines Kollektivgegenstands „zufällig“ – durch einen Zufallsgenerator – zustande kommen. An die Stelle der Annahme, daß es korrespondierend zu einem Gattungsbegriff auch bereits eine indefinite Menge von Dingen gibt, die unter den Gattungsbegriff fallen, tritt dann eine andere Vorstellung: *daß man sinnvoll von Prozessen sprechen kann*, die die Mitglieder eines Kollektivgegenstands bzw. ihre jeweils bestimmten Eigenschaften erzeugen. Dann ist es nicht erforderlich, sich auf indefinite Mengen zu beziehen (die Gesamtheit der Dinge oder Situationen, die die Prozesse erzeugen *könnten*), sondern man kann versuchen, sich auf die Prozesse und ihre Eigenschaften zu beziehen. Schließlich erlaubt diese Betrachtungsweise auch eine Explikation des Wortes ‘Verteilungsgesetz’, die nicht mehr auf eine Analogie zu statistischen Verteilungen angewiesen ist. Es ist ja diese Analogie, die zunächst dazu geführt hat, sich Kollektivgegenstände als indefinite Mengen vorzustellen. Definiert man dagegen einen Kollektivgegenstand durch Prozesse, die seine Elemente erzeugen, kann man versuchen, auch die weiteren Begriffsbildungen *auf die Prozesse* zu beziehen. Insbesondere kann der Begriff eines Verteilungsgesetzes auf die Prozesse bezogen werden, die die Mitglieder eines Kollektivgegenstands erzeugen, so wie dies in Abschnitt 2.1 bereits angedeutet worden ist.

2.3 Deutung durch Zufallsgeneratoren

1. Dann stellt sich allerdings die Frage, wie man geeignete Vorstellungen über die Prozesse bilden kann, durch die die Mitglieder eines Kollektivs entstehen und jeweils bestimmte Eigenschaften bekommen. Das kann man sich in jedem Einzelfall überlegen. Wie und wodurch bekommen Rekruten ihre Körpergrößen? Wie und wodurch bekommen Erwerbstätige ihre Einkommen? Wie und wodurch kommt es an bestimmten Orten zu bestimmten Temperaturen? In jedem Fall müßte man zunächst die Prozesse untersuchen, durch die in unserer Erfahrungswelt Dinge und Situationen entstehen und jeweils bestimmte Eigenschaften bekommen. In der theoretischen (Sozial-) Statistik wird indessen ein vollständig anderer Weg verfolgt. Die Prozesse werden nicht untersucht, sondern durch Zufallsgeneratoren gedeutet.

2. Wie die zu Beginn des vorangegangenen Abschnitts angeführten Zitate zeigen, bezieht sich Fechner auf eine große Vielzahl unterschiedlicher Kollektivgegenstände. Aber in keinem Fall untersucht er die Prozesse, die deren Mitglieder hervorbringen bzw. durch die sie jeweils bestimmte Eigenschaften bekommen. *Stattdessen* ist sein Grundgedanke, daß die Mitglieder eines Kollektivgegenstands „zufällig variieren“. Auf diese Weise sucht er den Anschluß an Begriffsbildungen und Gedankengänge der Wahrscheinlichkeitsrechnung:

„... da nach dem Begriffe der K.-G. ein solcher aus *nach Zufall variierenden* Exemplaren besteht, finden jedenfalls auch die allgemeinen Wahrscheinlichkeitsgesetze des Zufalls – und jeder Mathematiker weiß, dass es solche giebt – darauf Anwendung. In der That werden die Verteilungsverhältnisse der K.-G. allgemein von solchen *beherrscht*, indes dieselben Wahrscheinlichkeitsgesetze bei physikalischen und astronomischen Maßbestimmungen nur *nebensächlich* für die Sicherheitsbestimmung der erlangten Mittelmaße in Betracht kommen, hiermit hier eine ganz andere und viel unwesentlichere Rolle spielen als in der Maßlehre der K.-G. Insofern aber der Zufall unter bestimmten, für die verschiedenen K.-G. verschiedenen äußeren und inneren Bedingungen spielt, lassen sich, durch alle Zufälligkeiten durch, die verschiedenen K.-G. durch charakteristische, aus ihren Verteilungsverhältnissen ableitbare Konstanten unterscheiden. Diese sind es, worin die Bestimmtheit derselben gegeneinander ruht; und diese gilt es mit Berücksichtigung der allgemeinen Wahrscheinlichkeitsgesetze aufzusuchen." (Fechner 1897, S. 5)

Wichtig ist die Annahme, daß die Mitglieder eines Kollektivgegenstands (bzw. ihre Eigenschaften) durch einen „zufälligen Prozeß" zustande kommen, so daß Begriffsbildungen der Wahrscheinlichkeitsrechnung – insbesondere die Idee der Verteilungsgesetze – anwendbar werden. Bemerkenswert ist auch, daß sich Fechner von Fragestellungen der Theorie der Beobachtungsfehler abgrenzt, bei der Begriffsbildungen der Wahrscheinlichkeitstheorie dienen, um Vorstellungen über Meßfehler zu bilden. Dadurch unterscheidet sich Fechner von Quetelet, der Abweichungen von statistischen Mittelwerten in Analogie zu Meßfehlern gedeutet hat.[5] Die letzten beiden Sätze in dem angeführten Zitat deuten an, daß Fechner

[5] Anläßlich seiner Diskussion der von Quetelet verwendeten Normalverteilung, des „Gaußschen Gesetzes", schreibt Fechner (1897, S. 64): „Von Gauss selbst ist das Gesetz nicht für Kollektivabweichungen, als Abweichungen der einzelnen Exemplargrößen a von ihrem arithmetischen Mittel, sondern bemerkter- und bekanntermaßen für Beobachtungsfehler, als Abweichungen der einzelnen Beobachtungswerte eines Gegenstandes von ihrem arithmetischen Mittel aufgestellt; und an sich ist nichts weniger als selbstverständlich, dass eine Übertragung des Gesetzes von letzteren auf erstere stattfinde. In der That ist es doch von vornherein etwas sehr Anderes, Abweichungen vor sich zu haben, die wegen mangelnder Schärfe der Messinstrumente oder Sinne und zufälliger äußerer Störungen bei wiederholter Messung eines *einzelnen* Gegenstandes vom arithmetischen Mittel der Maße erhalten werden, und Abweichungen, welche die *vielen* Exemplare eines K.-G. von ihrem arithmetischen Mittel aus Gründen darbieten, welche in der Natur der Gegenstände selbst und der sie beeinflussenden äußeren Umstände gelegen sind. Es ließ sich also auch durchaus nicht a priori voraussagen, dass die Natur in diesen Abweichungen vom Mittel das Gesetz der Beobachtungsfehler befolgt, sondern galt erst, eine direkte Prüfung desselben an K.-G. selbt vorzunehmen." Schließlich schlägt Fechner auch eine Modifikation der Normalverteilung vor, die in der Lage ist, auch nichtsymmetrische Verteilungen darzustellen. Es sei indessen erwähnt, daß Fechner keineswegs der erste war, der – kritisch gegen Quetelet – allgemeinere Verteilungsfunktionen für statistische Gesamtheiten vorgeschlagen hat. Hinweise zur Geschichte finden sich u.a. bei Arne Fisher 1928, S. 181ff.

eine besonders wichtige Aufgabe auch darin sehen möchte, Verteilungen durch gewisse Maßzahlen charakterisierbar zu machen; etwa durch unterschiedliche Arten von Mittelwerten und Streuungsmaßen. Auch neuere Statistik-Lehrbücher sehen darin oft eine besonders wichtige Aufgabe. Für die Frage, ob und ggf. in welcher Weise man in der Sozialstatistik von Verteilungsgesetzen sprechen kann, ist das jedoch ohne Belang.[6]

3. Während bei Fechner die begriffliche Bezugnahme auf Gedankengänge der Wahrscheinlichkeitsrechnung oft nur in vagen Andeutungen besteht, haben dies nachfolgende Autoren sehr explizit gemacht. Exemplarisch sei auf Emanuel Czuber hingewiesen, der sich in einem Teil seiner „Wahrscheinlichkeitsrechnung" (1908) mit der Kollektivmaßlehre beschäftigt. Ihr Ziel sieht er, wie Fechner, darin, zu Charakterisierungen von „Arten" oder „Gattungen" zu gelangen:

„Um eine Art oder Gattung von Objekten zu beschreiben, hat man einen solchen Komplex von Merkmalen zusammenzustellen, mit dessen Hilfe es möglich

[6] Eine ganz ähnlich konzipierte Suche nach „Konstanten" findet sich bei K. Pearson (1896, S. 256): „*Variation.* – If a curve be constructed, of which the ordinate y is such that $y\,dx$ measures the frequency with which an organ lying in size between x and $x + dx$, occurs in a considerable population (500 to 1000 or more), the constants which, for any particular organ for any particular animal determine the form of this curve, are termed the *constants of variation*, or more briefly, the variation of the given organ." Eine sehr allgemeine Formulierung findet sich auch bei F. Y. Edgeworth (1908, S. 499): „... the general problem which may be enunciated as follows: There are given numerous observations relating to the attributes or 'organs' of particular cases or individual specimens; each observation being of type $(x_t, y_t, z_t, \ldots)$, where $x_t, y_t, z_t, \ldots$ are concurrent values of the variables $x, y, z, \ldots$, which represent the attributes or organs that are under measurement. It is given also that if the set, or 'series' of observations were prolonged indefinitely, under unaltered conditions, the group of attributes thus constituted would (tend to) conform to a frequency-function, or 'surface' of any number of dimensions, of which the form is given: say, $w = f(x, y, z, \ldots; c_1, c_2, \ldots)$; $c_1, c_2, \ldots$ being constants, in general not given in magnitude. Such being the *data*, the *quæsita* are as follows. It is required to determine the most probable, or best available, values of the *primary* constants, the averages (mode, arithmetic mean, &c.) of the organs. It is required also to determine the constants $c_1, c_2, \ldots$ which constitute the *secondary* frequency-constants." Wie Fechner greift auch Edgeworth auf die Vorstellung zurück, daß sich die Beobachtungen unter gleichen Bedingungen beliebig vermehren lassen. Das erschließt den Zusammenhang zum Begriff eines Zufallsgenerators, wie er in der Wahrscheinlichkeitsrechnung verwendet wird.

ist, über die Zugehörigkeit eines Objekts zur Gattung zu entscheiden; eine solche Beschreibung, eine *Gattungsbeschreibung*, erfordert die vergleichende Betrachtung einer Vielheit von Objekten; sie ist das Produkt einer Abstraktion, weil von manchen Eigenschaften, die bei der Individualbeschreibung notwendig wären, abgesehen wird. Die Gattungsbeschreibung bestimmt den Begriff der Gattung." (Czuber 1908, S. 344)

Czuber betont dann, daß die Elemente einer Gesamtheit „gleichartig" sein müssen, um eine „Gattung" zu repräsentieren. Mit dieser merkwürdigen Vorstellung einer „Gleichartigkeit" werden wir uns jedoch erst später genauer beschäftigen. Hier interessiert vielmehr, wie Czuber den Zusammenhang zur Wahrscheinlichkeitsrechnung herstellt.

4. Czuber erinnert zunächst daran, daß man sich auch in der Wahrscheinlichkeitsrechnung auf Kollektivgegenstände bezieht, nämlich durch eine Bezugnahme auf Zufallsgeneratoren, durch die Mitglieder eines Kollektivgegenstands erzeugt werden können:

„Die Wahrscheinlichkeitsrechnung führt in den Ergebnissen wiederholter Versuchsreihen zu theoretisch konstruierten Kollektivgegenständen, für welche sich die Verteilung a priori angeben läßt, mit der dann die faktische Verteilung verglichen werden kann [...]." (Czuber 1908, S. 350)

Dann folgt eine Bemerkung zu Quetelet:

„Die Fehlertheorie operiert gleichfalls mit Kollektivreihen, als deren einfachster Vertreter eine Reihe wiederholter Messungen einer und derselben Größe anzusehen ist, und sie bedient sich einer Funktion, die die Verteilung der Fehler nach ihrer Größe beschreibt [...].

An diese Fälle knüpfen denn auch die ersten Ansätze zu einer Kollektivmaßlehre an, die von A. Quetelet stammen; dieser nahm ohne weiteres volle Analogie zwischen der Verteilung von Körperlängen, Brustumfängen u. dgl. einer großen Anzahl erwachsener Personen und der Wiederholungszahlen weißer (oder schwarzer) Kugeln in einer großen Anzahl gleich ausgedehnter Ziehungsreihen aus einer Urne an, die weiße und schwarze Kugeln in gleicher Menge enthält; die Grenzform dieser Verteilung ist das Gaußsche Exponentialgesetz $\phi(x)$." (Czuber 1908, S. 350)

Schließlich wird der Ansatz von Fechner mit folgenden Bemerkungen erläutert und gewürdigt:

„Auch der eigentliche Begründer der Kollektivmaßlehre, G. Th. Fechner, schloß sich an dieses Gesetz an. Er nahm in seine Definition des Kollektivgegenstandes ausdrücklich das Moment der Zufälligkeit auf, indem er einen solchen als eine Gesamtheit von unbestimmt vielen *nach Zufall* variierenden Exemplaren erklärte, die durch einen Art- oder Gattungsbegriff zusammengehalten werden. Seine vielfachen Untersuchungen an bereits vorhandenen und an eigens geschaffenen Kollektivgegenständen führten ihn aber weiter. Während nämlich die zu $\phi(x)$ gehörige Verteilungskurve in bezug auf ihre größte Ordinate Symmetrie aufweist, gelangte Fechner zu Kollektivreihen, die eine von der Symmetrie so stark abweichende Verteilung aufwiesen, daß die Asymmetrie als wesentlich, nicht als eine bloße Störung einer in der Natur der Sache gelegenen Symmetrie erkannt werden mußte. Er glaubte, dieser Tatsache mathematisch dadurch Rechnung tragen zu können, daß er das gewöhnliche (einfache) Gaußsche Gesetz zu einem *zweiseitigen* erweiterte, indem er die Verteilungskurve aus zwei Zweigen zusammensetzte, deren jeder einem Gesetz der Form $\phi(x)$, aber von ungleichem Präzisionsmaß, folgt und deren Gipfelpunkte zusammenfallen. Nicht in dieser anfechtbaren Erweiterung liegt Fechners Verdienst, sondern in der Schaffung einer Systematik für die Untersuchung von Kollektivgegenständen, in der Einführung einer Teminologie für diesen Zweig der angewandten Mathematik. Wiewohl er ihn durch die Betonung des Zufälligen mit der Wahrscheinlichkeitsrechnung in Beziehung gebracht hatte, sah er ihn doch als selbständig an. In Wirklichkeit aber ist die Kollektivmaßlehre eine Fortbildung der Methoden der Wahrscheinlichkeitsrechnung, deren Ziel ja in der Bestimmung relativer Häufigkeiten besteht, und anwendbar ohne Rücksicht darauf, nach welchem Prinzip der Kollektivgegenstand entstanden ist." (Czuber 1908, S. 350f)

Hier wird also die Kollektivmaßlehre bereits ganz selbstverständlich als ein spezieller Zweig der angewandten Wahrscheinlichkeitsrechnung aufgefaßt. Die Frage, mit welcher Berechtigung für die Entstehungsprozesse von Kollektivgegenständen hypothetische Zufallsgeneratoren unterstellt werden dürfen, wird nicht gestellt. Czuber gehört auch bereits einer Generation von wahrscheinlichkeitstheoretisch gebildeten Statistikern an, die wußten, daß man für jede beliebige empirische Verteilung ein „theoretisches Verteilungsgesetz" (einen Zufallsgenerator) *konstruieren* kann. So erklärt sich auch, daß es für ihn ganz gleichgültig ist, „nach welchem Prinzip der Kollektivgegenstand entstanden ist".[7]

[7]In der neueren statistischen Literatur kommt eine durchaus vergleichbare Haltung darin zum Ausdruck, daß Statistik als eine dem Anspruch nach *allgemeine* Methodenwissenschaft aufgefaßt wird. Z.B. sagen Cox und Snell (1981, S. 3): „Statistical analysis deals with those aspects of the analysis of data that are not highly specific to particular fields of study. That is, the object is to provide concepts and methods that will, with suitable modification, be applicable in many different fields of application; indeed one of the attractions of the subject is precisely this breadth of potential application." Noch deutlicher heißt es bei Schaich (1990, S. 2): „Ihrer Eigenart nach sind die statistischen Methoden in ihrer theoretischen Grundlegung genau so wie in ihrer konkreten Anwendung vom *Erfahrungsbereich*, in welchem sie angewendet werden, *unabhängig*." Allerdings gibt es auch Gegenstimmen. Z.B. hat A.P. Dempster

5. Ganz ähnliche Auffassungen wurden auch von L. v. Bortkiewicz vertreten.[8] In seinem Buch „Die Iterationen“ (1917) beginnt Bortkiewicz mit folgender allgemeinen Definition einer *empirischen* Statistik:

„Faßt man beliebige Erscheinungen, die unter einen gemeinsamen Begriff fallen und sei es nur in bezug auf die räumlichen Grenzen, zwischen denen sie eingeschlossen sind, und inbezug auf den Zeitpunkt oder den Zeitabschnitt, in dem sie beobachtet werden, sei es außerdem noch in irgendwelcher Hinsicht miteinander übereinstimmen, gedanklich zusammen, so kommt ein Ganzes zustande, das man als eine *empirische Vielheit* bezeichnen kann. Dadurch, daß man die Elemente (Einzelfälle), aus denen sich eine empirische Vielheit zusammensetzt, als solche der Beobachtung unterwirft und die so gewonnenen Beobachtungsresultate summiert, macht man die betreffende empirische Vielheit zum Gegenstand der *Statistik*. Denn Statistik ist nichts anderes als eine auf 'Massenbeobachtung' und Summierung ihrer Ergebnisse beruhende Erkenntnis empirischer Vielheiten.“ (v. Bortkiewicz 1917, S. 1)

Dann wird jedoch der Ansatz einer *theoretischen* Sozialstatistik auf folgende Weise eingeführt:

„Empirische Vielheiten, sofern sie Gegenstand der Statistik [...] sind, brauchen nicht unbedingt aus großen Zahlen von Elementen zu bestehen oder selbst in größerer Zahl auftreten. Wohl aber ist beides oder mindestens eines von beidem erforderlich, wenn es gilt, über das Verhalten irgendwelcher empirischer Vielheiten vom Standpunkte der *Wahrscheinlichkeitstheorie* aus etwas auszusagen. Diesem Standpunkte zufolge werden nämlich die Elemente, aus denen sich die betreffenden empirischen Vielheiten zusammensetzen, unter die Herrschaft des 'Zufalls' gestellt und wird die dadurch bedingte Unbestimmtheit und Unberechenbarkeit des Verhaltens der empirischen Vielheiten auf die Weise überwunden, daß man sie aus hinreichend zahlreichen Elementen bestehen oder in hinreichend großer Zahl auftreten läßt.“ (v. Bortkiewicz 1917, S. 2f)

(1987, S. 2) betont: „I believe that there needs to be much more integration of the logic of statistical inference with the reasoning processes used routinely by working scientists.“ Eine noch radikalere Position hat N. W. Henry vorgetragen. Bezugnehmend auf einen Aufsatz von D. Freedman (1997), in dem sich dieser mit Problemen und Ungereimtheiten statistischer Modellbildung beschäftigt, schreibt Henry (1997, S. 64): „I propose that one way to eliminate these problems is to give up the notion that statistics is an independent discipline. The problems themselves will remain, of course, but they will no longer underly something called the logic of statistical inquiry. Instead they will be part of the foundations of the disciplines in which probability and statistical models are used. Solutions ... if that is the right word ... will need to be phrased in terms of the language and concepts of the disciplines, and they may vary from one to another.“

[8] Bortkiewicz lebte 1868 bis 1931. Einige biographische Angaben findet man bei Johnson und Kotz 1997, S. 181f, und, etwas ausführlicher, bei Gumbel, 1978.

Die theoretische, an Vorstellungen der Wahrscheinlichkeitsrechnung orientierte Statistik beginnt also mit der Annahme, daß die Elemente einer empirischen Vielheit (einer statistischen Gesamtheit) bzw. ihre Eigenschaften durch einen hypothetisch unterstellten Zufallsgenerator entstehen, der gewissermaßen die empirischen Sachverhalte beherrscht.

2.4 Kann der Ansatz begründet werden?

1. Während hypothetische Zufallsgeneratoren in der heutigen statistischen Literatur weitestgehend zu einer nicht mehr zu hinterfragenden Selbstverständlichkeit geworden sind, gab es in der Frühphase der Entwicklung einer theoretischen (Sozial-) Statistik durchaus Diskussionen über die Frage, mit welcher Berechtigung statistische Kollektive als durch hypothetische Zufallsgeneratoren erzeugt aufgefaßt werden können. Z.B. schrieb H. Bruns in seiner „Wahrscheinlichkeitsrechnung und Kollektivmasslehre“ (1906, S. 87):

„Bei den Glücksspielen nämlich oder, allgemeiner gesprochen, bei den Vorgängen, die mit Sicherheit auf das Urnenschema reduziert werden können, hatten – das wußte man – die Begriffe Zufall und Wahrscheinlichkeit einen guten Sinn und führten bei weiterer Entwickelung auf Sätze, die die Probe der Erfahrung aushielten. Daraufhin vollzog sich nun, und zwar recht früh, die Unterschiebung, daß die Sätze der W.-R. schlechthin überall da anwendbar seien, wo überhaupt von Zufall und Wahrscheinlichkeit die Rede ist. Die Prüfung, ob denn die für die Zufälle und Wahrscheinlichkeiten eines Glücksspiels charakteristischen Merkmale jedesmal vorhanden seien, wurde einfach unterlassen.“

Aber kann man prüfen, ob sich empirisch beobachtbare Sachverhalte und Ereignisse hypothetisch unterstellbaren Zufallsgeneratoren verdanken? Ich glaube, daß die Frage verneint werden muß, daß es sich vielmehr um ein Deutungsschema handelt, das grundsätzlich immer verwendet werden kann; so daß also schließlich die Fragestellung verändert werden muß: Für welche Erkenntnisinteressen könnte es sinnvoll sein, ein solches Deutungsschema zu verwenden?

2. Um zu einer Diskussion der Fragestellung zu gelangen, können einige Überlegungen aus den „Principien der Wahrscheinlichkeitsrechnung“ von J. v. Kries (1886) einen Leitfaden liefern. Im Vergleich zu anderen Autoren setzen die Überlegungen v. Kries' sehr grundsätzlich ein. Er beginnt mit folgender Problemstellung:

„Wenn wir aus einem mit schwarzen und weissen Kugeln gefüllten Gefässe viele Male hintereinander je eine Kugel herausziehen, die gezogene immer wieder zurücklegen und mit den übrigen vermischen, und auf diese Weise in einer grossen Zahl von Ziehungen n Mal eine schwarze und m Mal eine weisse Kugel erhalten, so scheint der Schluss gerechtfertigt, dass das Zahlen-Verhältniss der in dem Gefässe enthaltenen schwarzen und weissen Kugeln sich nicht sehr erheblich von dem Werte n/m unterscheide. Demgemäss nehmen wir an, dass die Wahrscheinlichkeit, in einer folgenden Ziehung eine schwarze resp. weisse Kugel zu erhalten, sehr annähernd $= n/(n+m)$ resp. $m/(n+m)$ gesetzt werden dürfe. Man bezeichnet diese Verfahrungsweise als eine *aposteriorische Wahrscheinlichkeits-Bestimmung*. Legen wir uns die Frage vor, inwieweit dieselbe als eine allgemein anwendbare angesehen werden darf, so müssen wir darauf achten, dass für die in dem obigen Beispiele unfragliche Berechtigung derselben gewisse Kenntnisse in Betracht kommen, welche deutlich zum Bewusstsein zu bringen die gewöhnliche Darstellung unterlässt. So wissen wir schon sicher, dass bei Ziehungen aus einem Gefässe die Wahrscheinlichkeit, eine schwarze oder weisse Kugel zu greifen, lediglich bestimmt wird durch das Zahlen-Verhältniss, in welchem solche vorhanden sind, dass jeder Ziehung, unabhängig von dem Erfolge aller anderen, eine bestimmte Chance des einen und des anderen Ergebnisses zuzuschreiben ist, dass die Zahl der Kugeln zwischen den einzelnen Ziehungen sich nicht verändert, mit einem Worte, dass die ganze zu beurteilende Erscheinung ein Zufalls-Spiel mit constanten Chancen und Unabhängigkeit der Einzelfälle darstellt. Was wir empirisch bestimmen, ist demnach nur der Zahlen-Wert dieser Chance. Hieraus geht hervor, dass die Anwendbarkeit jenes Verfahrens ganz und gar keine *allgemeine*, sondern eine an sehr wesentliche Voraussetzungen gebundene ist. Wenn, der üblichen Vorschrift entsprechend, bei jeder beliebigen Reihe gleichartiger Fälle, welche teils so, teils anders verlaufen, schlechtweg nach 'der Wahrscheinlichkeit eines bestimmten Verlaufs in einem derartigen Falle' gefragt wird, so ist, wie schon früher hervorgehoben wurde, zunächst fraglich, ob ein Wert, welchem diese Bedeutung zuzuschreiben wäre, überhaupt existirt." (v. Kries 1886, S. 132f)

3. Hieran anschließend bestimmt v. Kries die Aufgabe, mit der er sich im weiteren beschäftigen möchte, folgendermaßen.

„Demgemäss kann nun auch die Aufgabe, mit der wir uns hier beschäftigen, durchaus nicht darin bestehen, eine Methode zur numerischen Bestimmung irgend welcher ganz beliebiger Wahrscheinlichkeiten anzugeben; es wird vielmehr, entsprechend den wichtigen objectiven Voraussetzungen, welche in den Wahrscheinlichkeits-Sätzen zum Ausdruck gelangen, eine Begründung derselben zu fordern sein, welche weit mehr und ganz Anderes, als die blosse Bestimmung eines Zahlen-Wertes einschliesst. Wir überzeugen uns nun leicht, dass eine solche zum grossen Teile in Vorstellungen ganz allgemeiner Natur zu suchen ist, welche sich auf die Art und Weise, wie die betreffenden Erscheinungen zu Stande kommen, und auf das Verhalten der sie bedingenden Umstände beziehen." (v. Kries 1886, S. 133f)

Wichtig erscheint mir vor allem der zuletzt ausgesprochene Gedanke: daß man sich auf die Prozesse beziehen muß, durch die Sachverhalte oder Ereignisse hervorgebracht werden. Das ist jedenfalls dann der Fall, wenn man in der Wahrscheinlichkeitsrechnung Ereignisse als *durch Zufallsgeneratoren erzeugt* konzipiert und mithin nicht unmittelbar über Ereignisse spricht, sondern über Eigenschaften der Prozesse, durch die sie hervorgebracht werden.

4. v. Kries bezieht sich im weiteren zunächst auf explizit konstruierte Zufallsgeneratoren, deren Wahrscheinlichkeitsverteilung durch ihre Konstruktion begründbar erscheint. Dann geht es so weiter:

„Im Gegensatz zu den bisher betrachteten deductiven Ableitungen wollen wir nunmehr sehen, wie sich eine im engeren Sinne empirisch zu nennende, d.h. eine auf die Beobachtung der betreffenden Fälle selbst zu gründende Gewinnung von Wahrscheinlichkeits-Sätzen zu gestalten hat. Die erste Bedingung dafür, dass man überhaupt hoffen kann, eine zahlenmässige Wahrscheinlichkeit zu erhalten, ist offenbar die, dass eine Anzahl von Fällen, die eine gewisse Gleichartigkeit zeigen, zum Teil einen, zum Teil einen anderen Verlauf nehmen, dass bei den einen dieser, bei den anderen jener Erfolg auftritt. In einem derartigen Verhalten liegt schon eine gewisse Uebereinstimmung mit der Reihe einzelner Fälle eines Zufalls-Spiels. Diese Uebereinstimmung wird eine noch vollkommenere, wenn unsere Kenntnisse bezüglich jedes einzelnen Falles unvollständig sind, so dass der Erfolg jedesmal, wie beim Zufalls-Spiel, ungewiss erscheint. So beobachten wir z.B. im Laufe längerer Zeit eine grosse Zahl von Typhus-Fällen; ein Teil der Kranken stirbt, ein anderer Teil wird wieder gesund; und in jedem einzelnen Falle ist es nicht mit Sicherheit im Voraus anzugeben, ob das Eine oder das Andere Statt finden werde. Unter solchen Umständen nun scheint es wenigstens denkbar, dass die ganze Erscheinungs-Reihe sich in der That einem gewöhnlichen Zufalls-Spiele analog verhalte: dass unsere Kenntnis des einzelnen Falles, welcher zufolge wir ihn als einen Typhus bezeichnen, einen gewissen Bereich des Verhaltens offen lasse, und dass innerhalb dieses vergleichbare und indifferente ursprüngliche Spielräume von bestimmtem Grössen-Verhältnis dem tötlichen Ausgange und der Genesung entsprechen. Es bestände alsdann für jeden Typhus-Fall eine gewisse Chance des einen und des anderen Verlaufes, und es erschiene als Sache des Zufalls, ob dieser oder jener eintritt." (v. Kries 1886, S. 140f)

Es ist bemerkenswert, daß v. Kries hier von einer „Analogie" spricht. Um mehr geht es zunächst auch nicht. Die Frage ist, ob man sich bei einigen der in unserer Erfahrungswelt wiederholt auftretenden Sachverhalte oder Ereignisse die Prozesse, durch die sie hervorgebracht werden, *in Analogie* zur Erzeugung von Sachverhalten oder Ereignissen durch Zufallsgeneratoren vorstellen kann. (Auch auf die Frage, welchen Erkenntnisinteressen eine solche Analogie dienen könnte, soll hier zunächst nicht näher eingegangen werden.) — v. Kries' Gedankengang wird dann allerdings mit folgender Frage fortgesetzt: „Es fragt sich nun: wie kann die Ueberzeugung gewonnen werden, dass die Sache sich in Wirklichkeit so verhält?" Die Formulierung ist durchaus merkwürdig, denn „in Wirklichkeit" werden Typhus-Fälle natürlich nicht durch Zufallsgeneratoren erzeugt, sondern jeder einzelne Fall verdankt sich einem spezifischen Entstehungs- und Entwicklungsprozeß, den man unter Umständen auch genauer untersuchen könnte.

5. Man kann jedoch versuchen, die Überlegungen v. Kries' so zu verstehen, daß sie sich auf die Frage beziehen, welche Sinnvoraussetzungen erforderlich sind, um sich Entstehungsprozesse für Sachverhalte oder Ereignisse, z.B. für den tödlichen Ausgang von Typhus-Fällen, in Analogie zu Zufallsgeneratoren vorstellbar zu machen. v. Kries bespricht auch als nächstes eine solche Sinnvoraussetzung:

„Die erste und wesentlichste Frage ist offenbar die, ob für die betreffenden Verhältnisse irgendwelche allgemeine Bedingungen, deren Kenntnis zu der Angabe bestimmter Chancen berechtigte, wirklich constant bestehen. Wir erwähnten vorhin schon die Sicherheit analoger Voraussetzungen bei den Zufalls-Spielen; dass der Würfel, mit dem wir eine grössere Zahl von Würfen ausführen, dauernd die gleichen Bezeichnungen auf seinen Seiten trägt, dass das Gefäss, aus welchem wir Kugeln herausgreifen, vor jeder Ziehung dieselbe Zahl schwarzer und weisser Kugeln enthält u.s.w.: das unterliegt keinem Zweifel. Einer ganz beliebigen Reihe gleichartiger Erscheinungen gegenüber würde das entsprechende Verhalten nicht von vorn herein als sicher angenommen werden dürfen. Es wäre, wie man kurz sagen kann, erst zu prüfen, ob dieselbe sich überhaupt wie ein nach denselben Regeln dauernd fortgesetztes Zufalls-Spiel verhält." (v. Kries 1886, S. 141f)

Damit wird offenbar auf eine wichtige Frage hingewiesen: die Analogie mit einem Zufallsgenerator erfordert als eine Sinnvoraussetzung, daß die Prozesse, die die Sachverhalte oder Ereignisse hervorbringen, stets nach den gleichen Regeln ablaufen. Bei artifiziellen oder gedanklich konstruierten Zufallsgeneratoren wird dies durch ihre Definition erreicht oder vorausgesetzt. Sie sind *als Verfahren definiert*, die in einer stets gleichbleibenden Weise (wobei der Sinn dieser Formulierung durch das Verfahren bestimmt wird) wiederholt werden können. Sollte sich z.B. nach dem Werfen herausstellen, daß ein Würfel zu einer Kugel geworden ist oder seine Beschriftung verloren hat, gilt der Wurf nicht mehr als ein Beispiel für das anfänglich definierte Würfelspiel; ebenso z.B. dann, wenn man den Würfel absichtsvoll so auf den Tisch legt, daß eine bestimmte Seite

nach oben zeigt.

6. Aber wie verhält es sich mit dieser Sinnvoraussetzung, wenn man sich nicht auf artifizielle Zufallsgeneratoren bezieht oder, wie in der mathematischen Wahrscheinlichkeitsrechnung, sich mit gedanklich konstruierten Zufallsgeneratoren beschäftigt, bei denen die Wiederholbarkeit unter gleichen Bedingungen definitorisch vorausgesetzt werden kann? Denn darum handelt es sich ja in der theoretischen Statistik: daß man Zufallsgeneratoren verwenden möchte, um Modelle für den Ablauf von Prozessen zu bilden, die in unserer Erfahrungswelt Sachverhalte und Ereignisse hervorbringen. v. Kries bemerkt, daß dann diese Sinnvoraussetzung „geprüft" werden sollte, und er deutet auch an, wie eine solche Prüfung vorgenommen werden kann:

„Die Ueberzeugung nun, dass gewisse allgemeine, die Chancen des Verlaufs bestimmende Verhältnisse für eine Reihe gleichartiger Fälle in dauernder Constanz bestehn, kann offenbar auf die Thatsache gegründet werden, dass in einer sehr ausgedehnten Beobachtung die relative Häufigkeit der verschiedenen Verlaufsweisen dauernd sehr annähernd dieselbe geblieben ist. Man wird diesen Nachweis mit dem, was gewöhnlich aposteriorische Wahrscheinlichkeits-Bestimmung genannt wird, nicht verwechseln; es handelt sich nicht um die Ermittelung eines Zahlen-Wertes, welche im Allgemeinen aus einer Reihe von einigen hundert Fällen schon mit überschüssiger Sicherheit und Genauigkeit würde geschehen können, sondern es handelt sich um die Untersuchung, ob ein solcher Zahlen-Wert in der That in dauernder Constanz sich erhält. Dazu ist jedesmal die Beobachtung einer grösseren Anzahl derartiger Reihen erforderlich." (v. Kries 1886, S. 142f)

Mit diesen Ausführungen nehmen die Überlegungen v. Kries' gewissermaßen eine pragmatische Wendung. Es bleibt auf bemerkenswerte Weise unbestimmt, was damit gemeint sein soll, „dass in einer sehr ausgedehnten Beobachtung die relative Häufigkeit der verschiedenen Verlaufsweisen dauernd sehr annähernd dieselbe geblieben ist". Wie verhält es sich z.B. bei den Typhus-Fällen, die v. Kries oben als Beispiel verwendet hat? Denkt man daran, wie sich die Behandlungsmöglichkeiten von Typhus seit den Tagen v. Kries' verändert haben, verläuft diese Krankheit heute sicherlich nicht mehr unter den gleichen allgemeinen Bedingungen wie damals. Versagt infolgedessen in diesem Fall die Analogie mit einem Zufallsgenerator?

7. Offenbar steht man vor dem gleichen Problem, das bereits im Zusammenhang der Überlegungen J. Bernoullis in Abschnitt 1.3 angesprochen worden ist. Man kann zwei unterschiedliche Standpunkte einnehmen. Einerseits kann man grundsätzlich argumentieren, daß Zufallsgeneratoren keine angemessene Repräsentation von Prozessen liefern können, deren Regeln sich im Zeitablauf verändern. Man müßte dann konsequenterweise den Ansatz einer theoretischen Sozialstatistik, wie er bisher charakterisiert worden ist, kritisieren. Denn nicht nur unterliegen die Prozesse, durch die soziale Sachverhalte und Ereignisse entstehen, einem fortwährenden Wandel und verlaufen unter fortwährend sich verändernden historischen Bedingungen; man kann außerdem darauf hinweisen, daß auch die Akteure dieser Prozesse lernfähig sind und mithin auch selbst eine Quelle immer neuer Spielregeln sind. Tatsächlich ist auch innerhalb der Sozialstatistik diese Konsequenz von einigen Autoren – mehr oder weniger radikal – gezogen worden. Auf eine Variante, die sich an einer Gegenüberstellung von Sozial- und Naturwissenschaften orientiert, wird in Abschnitt 4.3.4 etwas genauer eingegangen.

8. Andererseits kann man auch einen Standpunkt einnehmen, wie er bereits von J. Bernoulli vertreten worden ist (vgl. Abschnitt 1.3): wenn die Bedingungen, unter denen Sachverhalte oder Ereignisse entstehen, einem historischen Wandel unterworfen sind, muß man gelegentlich neue Daten erheben, um die Eigenschaften der hypothetisch unterstellten Zufallsgeneratoren mit ihrer Hilfe neu zu berechnen. Für die Annahme, daß sich Erscheinungen einem unterstellten Zufallsgenerator verdanken, wird dadurch gewissermaßen ein empirischer Vorbehalt formuliert. Die Annahme eines jeweils bestimmten Zufallsgenerators – z.B. einer bestimmten Sterbetafel – erscheint solange sinnvoll, bis neue Beobachtungen eine Revision erforderlich machen.[9]

9. Der Entwicklung der theoretischen Sozialstatistik, bis hin zu den statistisch orientierten Varianten der gegenwärtigen empirischen Sozialforschung, liegt dieser pragmatische Standpunkt zugrunde.[10] Auch andere Befürworter einer theoretischen Sozialstatistik, soweit sie sich mit der v. Kries'schen Frage nach der Berechtigung hypothetischer Zufallsgeneratoren beschäftigt haben, enden regelmäßig mit einem pragmatischen

[9]So kann man folgende Bemerkung von R.A. Fisher (1966, S. 25) verstehen: „In 'The Improvement of Natural Knowledge', that is, in learning by experience, or by planned chains of experimentation, conclusions are always provisional and in the nature of progress reports, interpreting and embodying the evidence so far accrued."

[10]Aber die Formulierung ist nicht vollständig richtig, denn der „pragmatische Standpunkt", wie er *im praktischen Handeln* tatsächlich eingenommen wird, impliziert gerade keine einfachen Gleichförmigkeitsannahmen. Es handelt sich bereits um ein fragwürdiges *Postulat* der Theoriebildung, wenn etwa gesagt wird: „The basic assumption is that what has been consistently true or true on the average, in the past, will continue to occur, with measurable limits of variation, in the future." (Cohen 1938, S. 671f)

Argument. Als Beispiel soll auf A. Tschuprow hingewiesen werden:

„Im Gesetz der großen Zahlen haben wir das Werkzeug gefunden, mit deren Hülfe das roh-empirische Material der durch die statistische Beobachtung feststellbaren Häufigkeiten veredelt und der methodologisch wichtige Kern von der Schlacke des Zufälligen gereinigt werden kann. Ehe wir jedoch den folgenden Schritt im Ausbau des statistischen Verfahrens machen und zeigen, in welcher Weise die auf die allgemeinen Ursachen sich beziehenden Folgerungen, zu denen wir gelangen, verwertet werden können, muß der Nachweis geliefert werden können, daß unsere deduktiv begründete Auffassung vom Zusammenhange zwischen den empirischen Häufigkeiten und den apriorischen objektiven Wahrscheinlichkeiten keine aus der Luft gegriffene Konstruktion ist, der kein praktischer Wert zukommt, sondern daß sie vielmehr durch die Erfahrung bestätigt wird. Gibt es überhaupt Verhältnisse, unter denen ein Zusammenhang zwischen den Häufigkeiten und den Wahrscheinlichkeiten tatsächlich nachweisbar ist? Sind namentlich diejenigen Verhältnisse, unter denen die Arbeit des Statistikers vor sich geht, derart, daß er das Recht hat, die von ihm beobachteten Häufigkeiten der Ereignisse als Realisationen von Wahrscheinlichkeiten zu betrachten, die nur durch Zufälligkeiten in dem theoretisch berechenbaren Maße verschleiert werden?" (Tschuprow 1905, S. 38f)

Es ist klar, wie die erste der beiden Fragen beantwortet werden kann: durch einen Hinweis auf *artifizielle* Zufallsgeneratoren.

„Die erste Frage läßt sich kurz durch den Hinweis auf die Experimente auf dem Gebiete der Zufallsspiele erledigen. In allen Fällen, wo das Experiment sorgfältig ausgeführt war, so daß der betrachteten Ereignisreihe tatsächlich dasjenige Ursachensystem zugrunde lag, für welches die Wahrscheinlichkeiten berechnet worden waren, hat die Erfahrung mit den theoretischen Vorausberechnungen sehr gut übereingestimmt." (Tschuprow 1905, S. 39)

Die Schwierigkeit liegt aber natürlich in der zweiten Frage. Dazu sagt Tschuprow:

„... muß gezeigt werden, daß das Wahrscheinlichkeitsschema, das wir bis jetzt nur mit den Erfahrungen auf dem Gebiete der Zufallsspiele verglichen haben, auch auf die Massenerscheinungen anwendbar ist, welche der Statistiker zu erforschen hat. Es darf ja nicht jede aus einer Reihe von Beobachtungen abgeleitete Verhältniszahl als empirischer Ausdruck einer bestimmten objektiven Wahrscheinlichkeit aufgefaßt werden; damit dies der Fall sei, müssen gewisse Voraussetzungen erfüllt sein. Ob nun diese Voraussetzungen mit der Arbeitsweise des Statistikers übereinstimmen? A priori läßt sich diese Frage nicht entscheiden, da wir die statistisch erfaßbaren Erscheinungen und die sie bedingenden Ursachen nicht genau genug kennen. Die Lösung muß durch Betrachtung des empirischen Materials gewonnen werden, und das große Verdienst Lexis' ist eben, die Frage nach der Anwendbarkeit der Wahrscheinlichkeitsrechnung auf das Material der statistischen Forschung theoretisch scharf gestellt und genaue Methoden der Lösung derselben geschaffen zu haben.

Es ist ohne weiteres klar, daß die Betrachtung vereinzelter statistischer Zahlen zur Entscheidung der Frage nichts beitragen kann. Ob eine Sterblichkeit von so und so viel pro Mille als empirischer Ausdruck einer objektiven Wahrscheinlichkeit aufgefaßt werden darf, hierüber zu entscheiden, gibt es keine Mittel. Hat man dagegen eine Reihe von Häufigkeiten vor sich, die sich auf eine Anzahl verschiedener statistischer Massen beziehen, so läßt sich, wie Lexis gezeigt hat, entscheiden, ob sie alle als empirische Ausdrücke einer und derselben Wahrscheinlichkeit gelten dürfen.“ (Tschuprow 1905, S. 40)

Auch ohne Details des von Lexis vorgeschlagenen Verfahrens zu kennen,[11] kann man erkennen, daß Tschuprows Argumentation unzureichend ist. Mit der von Lexis vorgeschlagenen Varianzanalyse kann man bestenfalls Hinweise darauf gewinnen, ob für zwei oder mehr statistische Gesamtheiten ein gemeinsamer Zufallsgenerator unterstellt werden kann. Aber das setzt schon voraus, daß man die Elemente jeder Gesamtheit als durch einen Zufallsgenerator erzeugt auffassen kann. Für die von Tschuprow eingangs gestellte Frage nach der Berechtigung dieser Annahme liefert das Lexis'sche Verfahren keine Antwort. Diese Frage ist auch unabhängig davon, ob man es mit nur einer oder mit mehreren statistischen Gesamtheiten zu tun hat; und insofern gilt auch Tschuprows Bemerkung, daß *diese* Frage überhaupt nicht empirisch entschieden werden kann.

10. Es ist allerdings klar, daß sich dadurch auch der Charakter der den Ansatz konstituierenden Annahme verändert: daß man soziale Sachverhalte und Ereignisse sinnvoll als durch hypothetische Zufallsgeneratoren erzeugt auffassen könne. Es handelt sich nicht länger um eine Annahme, die durch eine empirische Untersuchung geprüft werden könnte, so wie dies v. Kries in den oben zitierten Ausführungen nahelegt. Denn ginge es darum, sie einer solchen Prüfung zu unterziehen, könnte sie jederzeit als unzutreffend zurückgewiesen werden. Vielmehr bekommt die Annahme den Charakter einer gedanklichen Fiktion, die es erlauben soll, über soziale Sachverhalte und Ereignisse so zu reden, *als ob* sie durch einen Zufallsgenerator zustande kommen. Die Frage ist dann nicht, ob man auf diese Weise eine sinnvolle und einsichtsvolle begriffliche Repräsentation sozialer Prozesse erreichen kann. Diese Frage, die

[11]Es handelt sich um die sog. Lexis'sche Dispersionstheorie, die erste Ausarbeitung eines Verfahrens der Varianzanalyse. Wir werden uns mit dem theoretischen Ansatz von Lexis genauer in Abschnitt 3.2 beschäftigen.

sich einem soziologischen Erkenntnisinteresse verdankt, ist der theoretischen Sozialstatistik ganz fremd. Stattdessen tritt eine praktische Frage in den Vordergrund: ob statistische Modelle, die auf einer Annahme hypothetischer Zufallsgeneratoren beruhen, zur Begründung konditionaler Erwartungen nützlich sein können.[12]

11. Es wäre sicherlich müßig, diesen pragmatischen Standpunkt kritisieren zu wollen.[13] Auch die Feststellung, daß durch ihn einem Verständnis sozialer Prozesse nicht gedient werden kann, liefert nicht schon eine Kritik. Es folgt nur, daß man die Gedankengänge der theoretischen Sozialstatistik ignorieren kann, wenn man *soziologische* Erkenntnisinteressen verfolgen möchte. Zu einem Gegenstand kritischer Diskussionen sollten jedoch Gedankengänge der theoretischen Sozialstatistik dann werden, wenn sie sich nicht auf Modellbildungen für praktische Orientierungszwecke beschränken, sondern mit dem Anspruch auftreten, auch zu einem Verständnis gesellschaftlicher Verhältnisse beitragen zu können.

2.5 Welche Erkenntnisinteressen?

Über die Frage, welchen Erkenntnisinteressen Sozialstatistik dienen soll und ob bzw. wie ihnen durch aus der Wahrscheinlichkeitstheorie übernommene Begriffsbildungen und Vorstellungen gedient werden kann, gab und gibt es unterschiedliche Auffassungen. Einige Gesichtspunkte sollen im folgenden erwähnt werden.

[12]Hier setzen sog. „subjektive“ Wahrscheinlichkeitsdeutungen ein, die Wahrscheinlichkeitsaussagen als Aussagen über Erwartungen – oder allgemeiner, aber auch unklarer, über „degrees of belief“ – konzipieren. Z.B. schrieb H. Jeffreys (1948, S. 56): „In fact I should say that the notion of chance [durch Zufallsgeneratoren definierbare Wahrscheinlichkeiten] is never more than a considered hypothesis that we are at full liberty to reject. Its usefulness is not that chances ever exist, but that it is sufficiently precisely stated to lead to inferences definite enough to be tested, and when it is found wrong we shall in the process find out how much it is wrong.“ Es ist bemerkenswert, daß Jeffreys hier nicht sagt, daß die hypothetische Unterstellung von Zufallsgeneratoren geprüft werden kann, sondern daß die Unterstellung dienen soll, Hypothesen zu formulieren.

[13]Auch deshalb, weil in der neueren statistischen Literatur oft ein Standpunkt vertreten wird, der eine Kritik gegenstandslos macht: „In applying mathematics to subjects such as physics or statistics we make tentative assumptions about the real world which we know are false but which we believe may be useful nonetheless.“ (G. E. P. Box 1976, S. 792)

2.5.1 Suche nach Typischem

1. Ein in der frühen Literatur zur Sozialstatistik oft betontes Erkenntnisinteresse richtet sich darauf, daß man *typische* Erscheinungen erkennen möchte. Der Gedanke ist zunächst plausibel: sozialstatistische Daten können verwendet werden, um Einblicke in tatsächlich realisierte Häufigkeitsverteilungen zu gewinnen. Man ist dann nicht gezwungen, „das Typische" in Gestalt einer einzigen (mittleren) Merkmalsausprägung anzugeben, sondern kann es in Gestalt einer Merkmalsverteilung sichtbar machen. Anstatt zum Beispiel zu sagen, daß ein Studium der Sozialwissenschaften typischerweise ... Jahre dauert, kann man darstellen, wie sich die tatsächlichen Studiendauern verteilen; und man kann vielleicht auch zeigen, wie sich die Verteilungen der Studiendauern an verschiedenen Universitäten unterscheiden und wie sie sich im Zeitablauf verändert haben.

2. Aber es ist sicherlich fragwürdig, ob es zur Verfolgung dieses Erkenntnisinteresses hilfreich sein kann, sich soziale Sachverhalte und Ereignisse als durch Zufallsgeneratoren zustandegekommen vorzustellen. Wo in der Geschichte der Sozialstatistik die Suche nach Typischem von einer solchen Vorstellung ausgegangen ist, sind regelmäßig nur merkwürdige und verfehlte Spekulationen entstanden. Ein berühmtes Beispiel liefert bereits der „mittlere Mensch" von A. Quetelet, auf den bereits in Abschnitt 1.4 hingewiesen worden ist.[14]

3. Merkwürdig verfehlt erscheint insbesondere Quetelets Gedanke, sich Abweichungen von einem „mittleren Menschen" (oder allgemein von Mittelwerten) als „zufällige Fehler" vorzustellen. Aber selbst ein vergleichsweise reflektierter Sozialstatistiker wie W. Lexis knüpft noch an diesen Gedanken an. Zum Beispiel schrieb er in seinen „Abhandlungen zur Theorie der Bevölkerungs- und Moralstatistik" (1903, S. 101):

> „Als typische Grössen bezeichnen wir diejenigen, deren beobachtete Einzelwerte sich nach dem Gesetz der zufälligen Abweichungen [gemeint ist das „Gaußsche Fehlergesetz", die Normalverteilung] um ihren Mittelwert gruppieren. Dieser Mittelwert stellt also möglichst genau den Typus dar, der in jedem einzelnen Falle gewissermassen erstrebt, aber infolge von zufälligen Störungen, die ebenso leicht in positiver wie in negativer Richtung wirken können, fast niemals genau erreicht wird. [...] Quetelet hat das grosse Verdienst, die in der Astronomie ausgebildete Fehlertheorie auf die menschlichen Körperdimensionen angewandt und nachgewiesen zu haben, dass diese als typische Grössen im obigen Sinne zu betrachten sind und ähnliche Untersuchungen lassen sich auch auf andere anthropologische und demographische Gegenstände ausdehnen."

[14] Adolf Wagner, zunächst ein Anhänger Quetelets, formulierte den Grundgedanken so: „Unsere Untersuchungen über den Menschen führen uns vermittelst der Beobachtung zahlreicher Individuen nothwendig zur Aufstellung eines *mittleren* oder *Durchschnittsmenschen*, welcher in körperlicher, geistiger, sittlicher Hinsicht der *Typus* z.B. der beobachteten Nation ist." (Wagner 1864, S. 8)

Lexis macht dann allerdings eine Einschränkung, aus der scheinbar eine Forschungsaufgabe für eine theoretische Sozialstatistik resultiert:

> „Selbstverständlich ist nicht jeder Mittelwert aus gleichartigen Beobachtungsgrössen ein Typus, sondern er hat nur dann die Bedeutung eines solchen, wenn die obige Bedingung in der Verteilung der Einzelwerte erfüllt ist, während er in allen anderen Fällen zweckmässiger als blosser *Durchschnittswert* zu bezeichnen ist." (ebda.)

Aber der Gedanke ist merkwürdig: Warum sollte ein Mittelwert dann „das Typische" zum Ausdruck bringen, wenn er sich auf eine Verteilung bezieht, die einer Normalverteilung ähnlich ist?

4. Die Betrachtungsweise verdankt sich einem Gedankengang, der damals in der Sozialstatistik sehr verbreitet war und, wie wir später sehen werden, in der theoretischen Statistik immer noch eine Rolle spielt. Zur Verdeutlichung beziehe ich mich auf die „Anthropometrie", in der die Idee „typischer Mittelwerte" besonders intensiv verfolgt wurde. Ein Beitrag von Lexis mit dem Titel „Anthropologie und Anthropometrie" (1909) gibt darüber einen guten Überblick. Worum es geht, wird von Lexis so erläutert:

> „Einerseits nämlich hat die Methode der anthropologischen Messungen, die ‘Anthropometrie’ eine allgemeine statistische Bedeutung, indem sie die geeignetsten Beispiele für die mathematisch-statistische Untersuchung *typischer* Größen liefert. [...] Auch lassen sich die anthropometrischen Methoden unmittelbar zur Untersuchung gewisser wirtschaftlicher und sozialer Fragen verwenden, wie z.B. folgender: Bilden die menschlichen Rassen und Nationalitäten verschiedene Typen, die körperlich mehr oder weniger zur Arbeit geeignet sind? Wie verhält sich die Leistungsfähigkeit des männlichen Geschlechts zu der des weiblichen? Welchen Einfluß hat die Art der Arbeit auf die anthropologische Beschaffenheit des Arbeiters überhaupt und namentlich auf die physische Entwicklung der zur Arbeit herangezogenen Kinder? Entstehen vielleicht als Auswüchse unserer Natur degenerierte menschliche Typen?" (Lexis 1909, S. 524)

Die „Bedeutung der typischen Größen" wird so erläutert:

„Typische Größen sind solche, die sich in der Wirklichkeit, wenn auch nicht in völliger Gleichheit, so doch nur mit *zufälligen Abweichungen* von einem Normalwert wiederholen. In der Anthropologie treten dieselben auf als gewisse Maße oder Maßverhältnisse des menschlichen Körpers, wie z.B. die Körperlänge oder der Brustumfang. Mißt man die Körperlänge einer großen Anzahl erwachsener Männer eines bestimmten Völkerstammes, so wird man zwischen dem größten und dem kleinsten einen ganz erheblichen Unterschied finden, selbst wenn man die pathologischen Fälle von Riesen und Zwergen ganz ausschließt. Stellt man ähnliche Messungen bei einem anderen Volksstamme an, so erhält man ebenfalls Resultate mit beträchtlichen Schwankungsgrenzen und zwar werden auf einer gewissen Längenstrecke Fälle von beiden Stämmen vorkommen. Hat man z.B. Schotten und Franzosen gemessen, so werden die kleinsten Personen der ersteren Nationalität kleiner sein als die größten der letzteren, also unterhalb des Maximalmaßes der letzteren Schotten und Franzosen zusammen gefunden werden. Dennoch ist der schottische Typus größer als der französische. Man erkennt dies schon daran, daß die letzten Ausläufer der schottischen Messungen nicht so tief nach unten reichen wie die der französischen. Doch würde sich aus der Vergleichung der mehr zufälligen extremen Fälle kein bestimmtes Urteil über die wirkliche Größendifferenz der beiden Typen ableiten lassen, sondern zu diesem Zweck ist es erforderlich, die eigentlichen typischen Körpergrößen beider Nationalitäten zu bestimmen und miteinander zu vergleichen.“ (Lexis 1909, S. 524)

Das Problem ist deutlich: zwei Häufigkeitsverteilungen sollen durch „typische Mittelwerte“ charakterisiert werden, um sie vergleichbar zu machen.

5. Dafür sind zwar keine wahrscheinlichkeitstheoretischen Überlegungen erforderlich; insbesondere kann man ganz ohne die Annahme auskommen, daß die jeweils vorliegenden Beobachtungen durch einen hypothetischen Zufallsgenerator erzeugt worden sind. Aber genau diese Vorstellung liefert den theoretischen Ansatz:

„Wie wird nun dieser typische Wert irgendeiner anthropologischen Maßgröße gefunden? Es ist das große Verdienst Quetelets, diese Frage wissenschaftlich beantwortet zu haben. Der typische Wert wird immer möglichst nahe durch das arithmetische Mittel aus den Ergebnissen der einzelnen Messungen dargestellt, aber es ist nicht umgekehrt auch jedes arithmetische Mittel als eine typische Größe anzusehen, sondern bei der letzteren müssen noch andere Bedingungen zutreffen. Es kommt hier die Theorie der zufälligen Fehler zur Anwendung, wie sie von den Astronomen zur Bestimmung des wahrscheinlichsten Wertes und des wahrscheinlichen Fehlers von gleichartigen, mehrfach wiederholten Beobachtungen angewandt wird. G. Hagen hat in seiner Wahrscheinlichkeitsrechnung, deren erste Auflage schon 1837 erschien, eine anschauliche Vorstellung von der Entstehung der regelmäßigen Verteilung der rein zufälligen Fehler gegeben, die dann von Quetelet in eine noch populärere Form gebracht und auf die zufälligen Abweichungen anthropologischer Maßgrößen angewandt worden ist. Er nimmt an, daß die Natur bei jedem einzelnen Menschen einer gleichartigen Gruppe eine gleiche Körperlänge, einen gleichen Brustumfang usw. herzustellen strebe, daß aber fortwährend allerlei störende Einwirkungen stattfinden, infolge deren diese erstrebte Größe nur mit *zufälligen*, übrigens größeren oder geringeren Fehlern oder Abweichungen vom Normalwert verwirklicht werde, ebenso wie der Astronom bei wiederholten Messungen derselben Größe immer etwas verschiedene Werte findet, die mit zufälligen Fehlern behaftet sind.

Man kann sich nun vorstellen, daß jede dieser zufälligen Abweichungen durch eine große (nach Hagen eine unendlich große) Zahl elementarer Störungen von gleicher absoluter Wirkungsgröße entstehe, jedoch so, daß jede dieser Störungen ebenso leicht im positiven wie im negativen Sinne wirken, d.h. ebenso leicht eine kleine Vergrößerung wie eine kleine Verminderung der Beobachtungsgröße hervorrufen kann. Durch diese gleiche Möglichkeit der positiven und der negativen Wirkung der Elementarstörungen wird eben der Charakter des *Zufälligen* in der Gesamtabweichung der Einzelbeobachtung von dem richtigen oder typischen Wert erzeugt.“ (Lexis 1909, S. 524f)

Es ist deutlich, wie in diesem Gedankengang Vorstellungen und Begriffsbildungen der Theorie der Beobachtungsfehler verwendet, aber zugleich auch umgedeutet werden. An die Stelle von Beobachtungsfehlern treten fiktive Ursachen. Aber das allein genügt nicht; um sie in Analogie zu Meßfehlern interpretieren zu können, müssen diese fiktiven Ursachen als „Störungen“ aufgefaßt werden; und das wiederum setzt voraus, daß es einen „richtigen“ oder „typischen“ Wert gibt, so daß man von Abweichungen sprechen kann. Der entscheidende metaphysische Gedanke wird von Lexis deutlich formuliert: daß „die Natur“ diese „richtigen“ oder „typischen“ Größen *anstrebe*, dabei jedoch von „Störungen“ – somit schon begrifflich in einen Gegensatz zur „Natur“ gebracht – behindert werde.

6. Der Gedankengang ist offenbar vollständig spekulativ, zeigt aber auch, wo Konnotationen zwischen dem „Typischen“ und dem „Normalen“ einsetzen können. Das „Normale“ ist dasjenige, was die Natur anstrebt und auch erreichen könnte, wenn sie in ihren Bemühungen nicht gestört würde. Es ging allerdings nicht einfach darum, was die Natur will;[15] ersetzt man das Wort ‘Natur’ durch das Wort ‘Staat’ versteht

[15] Z.B. bemerkt K. Krümmel in einem Aufsatz über „Anthropometrie und Statistik“ (1921, S. 102): „Vom theoretisch-statistischen Standpunkt aus müßte also das ‘normale’ Individuum eigentlich dasjenige sein, welches am häufigsten vorkommt, doch schließt für die uns interessierende Frage der Begriff der Norm vielmehr ein Werturteil

man z.B. folgende Formulierung in Durkheims „Vorlesungen zur Physik der Sitten und des Rechts“, die im Zeitraum 1896 bis 1915 gehalten wurden:

„Strenggenommen ist der Staat das Organ gesellschaftlichen Denkens. [...] Doch worauf richtet sich dieses Denken? Anders gefragt: Welche Ziele verfolgt der Staat normalerweise, und welche Ziele muß er dementsprechend unter den gegenwärtigen gesellschaftlichen Bedingungen verfolgen?“ (Durkheim 1950, S. 76)

Vor allem in seinen „Regeln der soziologischen Methode“ (1895) hat sich Durkheim bemüht, den Gedanken zu verfolgen, daß das „Typische“ auch als das „Normale“ (im Sinne einer Norm) aufgefaßt werden sollte:

„Wir werden diejenigen Tatbestände normal nennen, die die allgemeinsten Erscheinungsweisen zeigen, und werden den anderen den Namen krankhaft oder pathologisch beilegen. Kommt man überein, als Durchschnittstypus jenes schematische Gebilde zu bezeichnen, das man erhält, indem man die in der Art häufigsten Merkmale mit ihren häufigsten Erscheinungsformen zu einem Ganzen, zu einer Art abstrakter Individualität zusammenfaßt, so wird man sagen können, daß der normale Typus mit dem Durchschnittstypus in eins zusammenfließt und daß jede Abweichung von diesem Schema der Gesundheit eine krankhafte Erscheinung ist. [...] Das was der Physiologe studiert, sind die Funktionen des Durchschnittsorganismus, und bei dem Soziologen ist es nicht anders.“ (Durkheim 1895, S. 148)

Das ist im wesentlichen mit Quetelets „mittlerem Menschen“ identisch, nur daß Durkheim noch expliziter den normativen Sinn betont.

7. Die Idee „typischer Mittelwerte“ spielte in der Sozialstatistik noch längere Zeit eine wichtige Rolle. Eine umfangreiche Ausarbeitung hat sie in einer Arbeit über „Die statistischen Mittelwerte“ von Franz Zizek (1908) gefunden. Dort heißt es zum Beispiel:

„Eine statistische Reihe, deren Einzelwerte sich gemäß dem Gaußschen Gesetz um das arithmetische Mittel verteilen, stimmt äußerlich vollkommen mit einer Reihe von wiederholten Messungen an demselben Objekte überein. Es liegt daher nahe, die Lehrsätze, welche die Theorie der Beobachtungsfehler ausgebildet hat, auf derartige statistische Reihen und die Mittelwerte derselben zu übertragen. Dies tut die mathematische Statistik. Sie untersucht die statistischen Reihen darnach, ob die Verteilung der Einzelwerte um das arithmetische Mittel dem Gaußschen Gesetze entspricht. Reihen, bei denen dies der Fall ist, werden als ‘typische’ Reihen, die Mittelwerte aus solchen Reihen als ‘typische’ Mittelwerte bezeichnet. Liegt ein ‘typisches’ arithmetisches Mittel vor, so wird dessen Genauigkeit festgestellt und die Dispersion der Reihe um das Mittel in derselben Weise gemessen, als ob es sich um bloße Beobachtungsfehler bei wiederholten Messungen an einem und demselben Objekte handeln würde.“ (Zizek 1908, S. 204)

---

ein. Für uns ist ‘Norm’ in gewissem Sinne ein Zielbegriff, d.h. das zu Erstrebende, das wir aus der Masse der Beobachtungen herauszufinden versuchen müssen.“

Zizeks Überlegungen gehen folgendermaßen weiter:

„Natürlich können die Vorstellungen der Fehlertheorie auf die typischen statistischen Mittel in mancher Beziehung nur mutatis mutandis übertragen werden. In dem arithmetischen Mittel einer Reihe von Messungen an verschiedenen Objekten (z.B. Messungen der Körperlänge, der Lebensdauer, des Lohnes einer Anzahl von Personen) kann nicht der wahrscheinlichste Wert für die ‘wahre’ Größe eines bestimmten Objektes erblickt werden, da alle einzelnen Beobachtungseinheiten selbständige reale Erscheinungen und gleich ‘wahr’ sind. Man kann jedoch im arithmetischen Mittel der Messungsdaten einen Normalwert (bzw. die wahrscheinlichste empirische Bestimmung eines theoretischen Normalwertes) erblicken, der dem allgemeinen, auf alle Beobachtungseinheiten einwirkenden Ursachenkomplexe entspricht und von welchem die Einzelfälle lediglich infolge individueller, störender, zufälliger Ursachen abweichen. Demzufolge stellt dann das arithmetische Mittel den ‘Typus’ der beobachteten Erscheinung dar, welcher in den Einzelfällen mit bloß zufällen Abweichungen zum Ausdruck kommt. Aus diesem Gedankengang ergibt sich die große wissenschaftliche Bedeutung, welche die mathematischen Statistiker ‘typischen’ Mittelwerten zuschreiben.“ (Zizek 1908, S. 204f)

Bemerkenswert ist hier auch, wie Zizek seine Vorstellungen über „typische Mittelwerte“ mit einer Unterscheidung zwischen „wesentlichen“ und „zufälligen“ Ursachen verknüpft. Damit werden wir uns in Kap. 6 beschäftigen.

8. Die Übernahme von Vorstellungen aus der Theorie der Beobachtungsfehler in die Sozialstatistik läßt auch verstehen, warum die sog. „Normalverteilung“ in der Anfangszeit eine so zentrale Rolle gespielt hat. Allerdings ist schon bald erkannt worden, daß die meisten „sozialen Massenerscheinungen“ keine Ähnlichkeit mit einer Normalverteilung aufweisen. Schon bei der Besprechung von Fechners Kollektivmaßlehre wurde darauf hingewiesen. Das hat Ende des 19., Anfang des 20. Jahrhunderts zu einer sehr ausgedehnten Diskussion darüber geführt, wie man noch andere, allgemeinere Verteilungsgesetze finden könne, durch die soziale Erscheinungen gedeutet werden können (darauf werden wir in Abschnitt 5.3 noch etwas genauer eingehen). Damit entstand auch ein weiteres Problem, das in der theoretischen Sozialstatistik breit diskutiert wurde: wie bei nicht-symmetrischen Verteilungen ein „typischer Mittelwert" definiert werden sollte, der die gleichen ideologischen Funktionen übernehmen kann, wie der arithmetische Mittelwert bei der Normalverteilung. In seiner Arbeit zur „Anthropometrie" (1909) deutet Lexis die neuen Spekulationsmöglichkeiten an:

„Bei vielen anthropologischen Maßgrößen ist selbst, wenn eine sehr große Zahl von Beobachtungen vorliegt, die symmetrische Gruppierung um einen typischen Wert nicht zu erkennen. [...] Es lassen sich für diese unregelmäßige Verteilung[16] sehr verschiedene Ursachen denken:

1. Es könnte bei der untersuchten Maßgröße eine *typische Unsymmetrie* bestehen. Wir haben oben angenommen, daß positive und negative Störungen des Typus an sich gleichwahrscheinlich seien, und auf dieser Voraussetzung beruht die symmetrische Verteilung der Abweichung zu beiden Seiten des Mittelwertes. Man könnte nun aber auch von der Hypothese ausgehen, daß die eine Art der Störungen in einem bestimmten Verhältnis wahrscheinlicher sei als die andere, womit dann die Analogie eines Glücksspiels an einer Urne gegeben wäre, in welcher schwarze und weiße Kugeln z.B. im Verhältnis von 2 : 3 oder 3 : 4 vorhanden wären. [...]

2. Die unsymmetrische Verteilung kann auch dadurch entstehen, daß die untersuchte Maßgröße als Funktion einer anderen erscheint, die ihrerseits einen typischen Charakter besitzt. Wenn z.B. das Gewicht der Erwachsenen sich nach Quetelet und anderen annähernd wie das Quadrat der Körpergröße verhält, so kann, gerade weil die Messungen der letzteren sich symmetrisch ordnen, die Reihe der Gewichtsbestimmungen keine symmetrische Verteilung der Fälle aufweisen.

3. Es kann eine *Mischung* mehrerer Rassentypen vorhanden sein, die in den Beobachtungen nicht auseinander gehalten sind. [...]

4. Die unregelmäßige Verteilung der Messungsergebnisse kann aber auch daher rühren, daß der Typus einer *Veränderung*, namentlich in der *Entartung* begriffen oder einer solchen schon in bedeutendem Grade verfallen ist. Man kann sich vorstellen, daß ein gewisser Prozentsatz, z.B. die Hälfte der Gesamtheit, sich noch regelmäßig um den ursprünglichen Typus gruppiert, der Rest dagegen durch die äußeren schädlichen Einflüsse oder sonstige Ursachen eine andere Verteilung erhalten hat. [...]

5. Endlich ist es auch möglich, daß für gewisse Maßgrößen oder Größenverhältnisse gleichsam Indifferenzstrecken vorhanden sind, innerhalb welcher die Natur keinen bestimmten Typus bevorzugt, so daß eine ziemlich gleichmäßige

[16] [Die Formulierung ist bemerkenswert, weil sie darauf hinweist, daß das Wort 'regelmäßig' lange Zeit verwendet wurde, um auf eine bestimmte Verteilungsform (die Normalverteilung) zu verweisen.]

Verteilung oder auch infolge irgendwelcher unberechenbarer Einflüsse mehrere ganz unregelmäßige Anhäufungen der Beobachtungsfälle entstehen." (Lexis 1909, S. 530ff)

Man erkennt, welche Spekulationsmöglichkeiten die Rezeption von Begriffsbildungen und Vorstellungen der Wahrscheinlichkeitstheorie für die Sozialstatistik eröffnet.

9. Schließlich sei erwähnt, daß der Gedanke, hypothetische Zufallsgeneratoren zur Begründung von „Typischem" zu verwenden, nicht von allen Autoren unmittelbar mit einer Idee „typischer Mittelwerte" verknüpft worden ist. Als Beispiel sei auf die Idee einer „statistischen Wesensform" hingewiesen, wie sie vornehmlich von Wilhelm Winkler (1929, 1931) vertreten worden ist.[17] Winklers These ist, daß die Statistik „ihrer Natur nach die Wesensform der Massenerscheinungen zum Forschungsgegenstand hat." (1929, S. 422) Was damit gemeint ist, kann im Anschluß an die Überlegungen in Abschnitt 2.1 erklärt werden. Winkler stellt sich vor, daß jedem statistischen Kollektiv ein Zufallsgenerator zugrunde liegt, der die Elemente des Kollektivs bzw. ihre Eigenschaften hervorgebracht hat. Als „Wesensform" einer empirisch ermittelten Häufigkeitsverteilung betrachtet er dann die Wahrscheinlichkeitsverteilung, die den Zufallsgenerator charakterisiert, der – als hypothetische Konstruktion – das statistische Kollektiv erzeugt hat. Hat man z.B. 100 mal gewürfelt, erhält man eine bestimmte Häufigkeitsverteilung; ihre „Wesensform" ist dann die Gleichverteilung, die das Würfelspiel charakterisiert. So heißt es z.B. bei Winkler (1929, S. 421):

„Wenn wir nach diesem Gedankengang eine Begriffsbestimmung der statistischen Wesensform in meinem Sinne geben wollen, so müssen wir sie als diejenige Form eines statistischen Tatbestandes bezeichnen, der die Masse bei wachsender Zahl von Beobachtungseinheiten asymptotisch zustrebt. Wir nennen sie darum *Wesensform*, weil wir der Analogie der Glücksspiele entnommen haben, daß sie das Wesentliche der Erscheinung, bestimmt durch die Grundwahrscheinlichkeit des Vorganges, widerspiegelt."

Aus Überlegungen dieser Art gewinnt Winkler zum Beispiel folgende Fragestellung einer theoretischen Sozialstatistik:

„An zwei verschiedenen Orten oder am gleichen Orte zu verschiedenen Zeiten sind statistische Beobachtungen über einen Tatsbestand gemacht worden. Sind die Ergebnisse so wenig voneinander unterschieden, daß sie als zufällige

[17] Winkler lebte von 1884 bis 1984. Kurze biographische Hinweise findet man bei Johnson und Kotz 1997, S. 322ff.

Abweichungen von einer gemeinsamen Wesensform betrachtet werden können, oder unterscheiden sie sich so stark voneinander, daß notwendigerweise das Vorhandensein zweier voneinander verschiedener Wesensformen angenommen werden muß?“ (Winkler 1931, S. 38f)[18]

Winkler stellt dann dar, wie man die Frage durch einen statistischen Test entscheiden kann. Aber was ist von der Fragestellung zu halten? Hat man zum Beispiel die Einkommensverteilung bei einer bestimmten Menge von Einkommensbeziehern für zwei Zeitpunkte ermittelt, möchte man vermutlich wissen, ob und ggf. welche Veränderungen eingetreten sind. Aber dann sind offenbar alle Veränderungen wichtig, die man feststellen kann. Warum sollte man die Einkommen als „zufällig entstanden“ betrachten? Nur um ggf. sagen zu können, daß auch die Veränderungen „zufälliger Natur“ sind und daß sich an der „Wesensform“ der Einkommensverteilung nichts geändert hat?[19]

2.5.2 Wofür Zufallsgeneratoren?

1. Man kann versuchen, hinter der Suche nach „Typischem“ und insbesondere nach „typischen Mittelwerten“ auch ein praktisches Erkenntnisinteresse zu finden. Folgendes Beispiel von Quetelet macht den Gedankengang deutlich:

„Um zu zeigen, wie weit das Studium der allmäligen Entwicklung des Menschen noch zurück ist, wollen wir annehmen, es handle sich darum, das Alter eines Individuums nach den physischen Merkmalen, die sein Körper an die Hand gibt, zu bestimmen; die Wissenschaft bietet uns keine Hülfsmittel dar, mittelst deren sich die Frage lösen liesse; wir sehen uns hier einem schauerlichen Empirismus preisgegeben. Indessen kommen in der gerichtlichen Arzneikunde nicht selten Fälle vor, wo dergleichen Bestimmungen nöthig werden. Ohne Zweifel kann man sich fragen, ob es wohl jemals möglich seyn werde, dabei zu genügenden Ergebnissen zu gelangen, besonders in Betreff der späteren Lebensalter. So gegründet diese Besorgnis auch erscheinen mag, so wird man

[18]In Fragen dieser Art sehen viele Autoren bis heute ein wichtiges Problem der Statistik. Z.B. heißt es bei Siegel (1956, S. 2): „A common problem for statistical inference is to determine, in terms of probability, whether observed differences between two samples signify that the populations sampled are themselves really different.“ Auf die hier verwendete Rhetorik, in der von Stichproben und Populationen gesprochen wird, werden wir in Kap. 4 etwas näher eingehen.

[19]Man beachte, daß es sich hier nicht darum handelt, zwei Zufallsstichproben zu vergleichen, also nicht um eine Reflexion von „Stichprobenfehlern“; vielmehr geht es um die Prozesse, durch die die beobachteten Sachverhalte, in diesem Beispiel die Einkommen, zustande gekommen sind. Hier kommt es allerdings oft zu einer Verwechslung.

doch dergleichen Untersuchungen nicht ohne Weiteres von der Hand weisen dürfen; es wäre dies sehr unphilosophisch. Lassen sich mit den Daten, welche die tägliche Beobachtung und das Augenmass an die Hand geben, messbare physische Merkmale verbinden, so ist es der Klugheit gemäss, letztere nicht unbeachtet zu lassen. Wird ein Arzt beauftragt, den Körper eines leblos gefundenen Kindes zu untersuchen, und bestimmt er in dem Inspektionsprotokoll das muthmassliche Alter dieses Kindes blos nach dem äusseren Aussehen, so ist es einleuchtend, dass er denjenigen, die das Protokoll lesen, sein Urtheil nur aufdringen kann, so falsch es übrigens auch seyn mag, da nichts zu der Beglaubigung desselben dient. Kämen dagegen zur Unterstützung der aufgestellten Schätzung des Alters noch das Gewicht und die Grösse des Kindes und einige andere messbare physische Merkmale, und hätte man nebenbei genaue Tafeln, auf welchen für die verschiedenen Alter die Werthe dieser physischen Merkmale und die Gränzen, innerhalb welcher sie bei regelmässig gebildeten Subjekten variiren, verzeichnet wären, so wäre es möglich, das über das Alter abgegebene Urtheil näher zu belegen; es würde selbst überflüssig werden, sobald die Elemente eines solchen Beleges eine grosse Genauigkeit zuliessen. Bestimmungen dieser Art darf also die gerichtliche Arzneikunde nicht von der Hand weisen, da sie geeignet sind, an die Stelle von Schätzungen, die blos empirisch und deshalb immer vage und oft falsch sind, bestimmte Merkmale und genaue Daten zu setzen.“ (Quetelet 1835, S. 329)

Die Idee besteht darin, daß der Statistiker auf der Grundlage einer Vielzahl von Beobachtungen Tafeln über den Zusammenhang zwischen einer Reihe beobachtbarer Merkmale und einem für jede Merkmalskombination „typischen“ (mittleren) Alter anfertigen kann und daß dann diese Tafeln verwendet werden können, um in Fällen, in denen das Alter unbekannt ist, darüber anhand der beobachtbaren Merkmale bessere Schätzungen zu finden.

2. Quetelets Beispiel macht auch deutlich, wo die Idee eines Mittelwerts sinnvoll werden kann. Wird die von ihm gewünschte Tabelle angefertigt, findet man zu jeder Merkmalskombination eine Mehrzahl von Menschen mit unterschiedlichem Alter; in moderner Sprechweise: eine durch die Merkmalskombinationen bedingte Häufigkeitsverteilung für das Alter. Das genügt indessen nicht, wenn ein Gerichtsmediziner die Tabelle zur Schätzung des Altes eines Menschen verwenden will, für den nur eine bestimmte Merkmalskombination bekannt ist. Er muß aus der für diese Merkmalskombination aus früheren Beobachtungen ermittelten Altersverteilung einen „Erwartungswert“ bilden, der als Schätzwert für das unbekannte Alter seines Untersuchungsobjekts verwendet werden kann. Er muß also irgendeinen durch die Merkmalskombination bedingten Mittelwert berechnen. Also stellt sich die Frage, wie man dies „am besten“ machen kann; denn es ist klar, daß man Mittelwerte auf viele unterschiedliche Weisen berechnen kann. Diese Frage motiviert bis heute viele Gedankengänge der theoretischen Statistik. Aus dieser Sicht betrachtet war die in der Frühzeit der Sozialstatistik betriebene Suche nach „typischen Mittelwerten“ gewissermaßen ein erster Versuch, um die „Kunst der Vermutungen“ im Bereich der Sozialstatistik mithilfe von Daten zu rationalisieren.

3. Unterstellt man ein praktisches Interesse an Schätzwerten, wird auch in gewisser Weise das Reden von „Fehlern“ verständlich, das ja zunächst nur in der Theorie der Beobachtungsfehler einen nachvollziehbaren Sinn hat. Abweichungen von einem irgendwie gebildeten Schätzwert können dann als Schätzfehler verstanden werden. Beträgt zum Beispiel bei einer bestimmten Merkmalskombination der o.a. Tabelle das (irgendwie definierte) Durchschnittsalter 30 Jahre und beträgt das tatsächliche Alter der zu untersuchenden Leiche 34 Jahre, gibt es dementsprechend einen Schätzfehler, wenn der Gerichtsmediziner das Alter der Leiche auf 30 Jahre schätzt. Auf diese Weise kann eine gewisse Ähnlichkeit mit Gedankengängen der Theorie der Beobachtungsfehler hergestellt werden. Es sollte aber betont werden, daß diese Analogie darauf beruht, daß ein praktisches Interesse an Schätzwerten vorausgesetzt wird. In Quetelets Beispiel ist ein solches Interesse klar erkennbar; aber weder bei Quetelet noch in der nachfolgenden Sozialstatistik bildete ein solches Erkenntnisinteresse den Ausgangspunkt und systematischen Leitfaden. Es kann deshalb bestenfalls verwendet werden, um einen Aspekt der damals betriebenen Suche nach „Typischem“ verständlich zu machen.

4. Schließlich stellt sich die Frage, wofür eine Unterstellung hypothetischer Zufallsgeneratoren benötigt wird. Denn offenbar ist sie nicht erforderlich, wenn man Sozialstatistik als ein empirisches Unternehmen konzipiert. Ihre Aufgabe besteht dann darin, Informationen über empirisch fixierbare Gesamtheiten zu gewinnen und in geeigneter Weise darzustellen. Eine statistische Darstellung würde darin bestehen, Häufigkeitsverteilungen zu ermitteln und sie näher zu charakterisieren. Aber dafür braucht man nicht die Annahme, daß die Mitglieder einer Gesamtheit bzw. ihre jeweils bestimmten Eigenschaften durch einen Zufallsgenerator zustandegekommen sind. Im Gegenteil, diese Annahme führt nur zu fragwürdigen Spekulationen, z.B. über „zufällige“ Abweichungen von „typischen Mittelwerten“, wie die exemplarischen Hinweise in Abschnitt 2.5.1 gezeigt haben. Sie kann zwar verwendet werden, um Verteilungsgesetze zu formulieren, und daran können sich Überlegungen anschließen, wie solche Verteilungsgesetze mithilfe der jeweils verfügbaren Daten am

besten geschätzt werden können. Aber da die unterstellten Zufallsgeneratoren als Fiktionen vorausgesetzt werden, bleibt fragwürdig, ob daraus irgendein Erkenntnisgewinn resultieren kann. Denn der Begriff eines Verteilungsgesetzes bezieht sich auf den unterstellten Zufallsgenerator, nicht unmittelbar auf die sozialen Sachverhalte und Ereignisse, die dem Sozialstatistiker empirisch zugänglich sind.

5. Der kritische Einwand zielt darauf, daß man durch fiktiv unterstellte Zufallsgeneratoren bestenfalls Deutungsschema bekommt, jedoch keine Einsichten in die Prozesse, durch die soziale Sachverhalte und Ereignisse tatsächlich zustandekommen. Allerdings gibt es zwei Überlegungen, die – vielleicht – den Einwand relativieren könnten. Man könnte zugeben, daß sich aus einer bloß hypothetischen Unterstellung von Zufallsgeneratoren keine Einsichten gewinnen lassen, daß man sich jedoch dadurch einen geeigneten methodischen Rahmen verschaffen kann, um empirisch explizierbare Einsichten in Prozesse *formulieren* zu können. Diese Überlegung sollte schon deshalb ernst genommen werden, weil sie auch der neueren statistisch orientierten Sozialforschung zugrunde liegt. Allerdings stellt sich dann die Frage, inwieweit die unterstellten Zufallsgeneratoren nicht nur als ein Deutungsschema verwendet werden, sondern tatsächlich durch Verweise auf Prozesse in unserer Erfahrungswelt empirisch expliziert werden können. Diese Frage sollte also nicht aus dem Auge verloren werden, wenn wir uns später mit sozialstatistischen Modellen beschäftigen.

6. Eine zweite Überlegung kann an das eingangs skizzierte praktische Erkenntnisinteresse anschließen. Denken wir an den Gerichtsmediziner, der das Alter einer Leiche schätzen soll. Er kennt einige Merkmale der Leiche und außerdem eine bedingte Häufigkeitsverteilung für das Alter einer Menge bisher untersuchter Personen mit der gleichen Kombination von Merkmalswerten. Die Frage ist, wie er dieses Wissen verwenden kann, um zu einem Schätzwert für das Alter der Leiche zu gelangen. Dieses Schätzproblem könnte durch die Annahme reflektierbar gemacht werden, daß das Alter der Leiche wie auch das Alter der bisher beobachteten Personen durch den gleichen Zufallsgenerator zustandegekommen sind, wobei die feststellbaren Merkmale als Bedingungen vorausgesetzt werden. Dann kann man Begriffsbildungen und Vorstellungen der Wahrscheinlichkeitstheorie verwenden, um über Schätzverfahren nachzudenken. Der Gedankengang ist in der neueren Literatur zur theoretischen Statistik sehr verbreitet: man unterstellt Zufallsgeneratoren deshalb, um Schätzprobleme (oder auch Tests für Hypothesen) mit Begriffsbildungen der Wahrscheinlichkeitstheorie formulierbar zu machen.[20] Aber der Gedankengang hat ersichtlich eine offene Flanke. Er könnte sinnvoll sein, wenn es explizierbare Gründe für die Unterstellung eines Zufallsgenerators gibt; und zwar müßte es sich um Gründe handeln, die auf die tatsächlichen Entstehungsprozesse der einzuschätzenden Erscheinungen bezug nehmen. Wenn es sich dagegen um eine rein fiktive Unterstellung handelt, wird der Gedankengang zirkulär. Denn der bloße Wunsch, statistische Inferenzverfahren anwenden zu können, liefert natürlich keine hinreichende Begründung für die Annahme, daß auch die dafür erforderlichen Voraussetzungen gegeben sind. Wir werden uns deshalb mit diesem Gedankengang im Rahmen dieses Skripts nicht weiter beschäftigen.

2.5.3 Regelmässigkeiten und Gesetze

1. Ein weiteres Erkenntnisinteresse durchzieht die Entwicklung der theoretischen Sozialstatistik: die Suche nach Regelmäßigkeiten und Gesetzen. Wiederum ist Quetelet ein wichtiger Wegbereiter. Seine Idee war, daß in den Eigenschaften und Verhaltensweisen eines statistisch konstruierbaren „mittleren“, „typischen“ Menschen zugleich die „Gesetzmäßigkeiten“ des gesellschaftlichen Lebens zum Ausdruck kommen. So schrieb er zum Beispiel in seiner „Physik der Gesellschaft“ (1835):

„Auf welche Weise sind die den Menschen betreffenden Gesetze zu erforschen und zu erklären? Wir haben soeben gesehen, dass der Mensch nicht allein in Beziehung auf seine körperlichen Fähigkeiten, sondern selbst in Bezug auf seine Handlungen unter dem Einflusse von Ursachen steht, die grösstentheils etwas Regelmäßiges und Periodisches haben und eben so regelmässige und periodische Wirkungen nach sich ziehen. Durch fortgesetzte Forschungen kann man diese Ursachen und die Art, wie sie wirken, oder die Gesetze, deren Grundlage sie bilden, ausfindig machen; aber, wie schon bemerkt, muss man zu dem Ende die Massen studiren, um aus den Beobachtungen alles Zufällige oder Individuelle zu entfernen. Die Wahrscheinlichkeitsrechnung zeigt, dass unter übrigens gleichen Umständen man sich um so mehr der Wahrheit oder den Gesetzen, die man ergründen will, nähert, eine je grössere Anzahl von Individuen den Beobachtungen als Stütze dienen.“ (S. 9) „Es kommt mir nicht in den Sinn, eine Theorie vom Menschen aufstellen zu wollen, ich will vielmehr blos versuchen, die ihn betreffenden Thatsachen und Erscheinungen auszumitteln, um auf dem Wege der Beobachtung die Gesetze, welche diese Erscheinungen miteinander verketten, ausfindig zu machen. Der Mensch, wie ich ihn hier betrachte, ist in der Gesellschaft dasselbe, was der Schwerpunkt in den Körpern ist; er ist das Mittel, um das die Elemente der Gesellschaft oszilliren; er ist, wenn man so will, ein fingirtes Wesen, bei dem alle Vorgänge den in Beziehung auf die Gesellschaft resultirenden mittleren Ergebnissen entsprechen werden.“ (S. 15)

[20] Die Diskussion ist schwer überschaubar. Es gibt mindestens zwei Gedankengänge, die von D. R. Cox (1990, S. 169) so unterschieden wurden: „Probability is used in two distinct, although interrelated, ways in statistics, phenomenologically to describe haphazard variability arising in the real world and epistemologically to represent uncertainty of knowledge. Discussions of statistical inference focus on the latter, whereas here the emphasis is firmly on the former.“ In der theoretischen Statistik wird allerdings oft gar keine Differenzierung vorgenommen. Z.B. heißt es bei D. V. Hinkley und N. Reid (1991, S. 1): „The representation of experimental and observational data as outcomes of random variables, or stochastic processes, provides a structure for the systematic treatment of inference from data, by which inductive conclusions from the particular to the general can be drawn.“

2. In der Entwicklung der Sozialstatistik hat sich allerdings Quetelets Vorschlag, „soziale Gesetzmäßigkeiten“ durch die Konstruktion eines „mittleren Menschen“ zu untersuchen, nicht durchgesetzt. Stattdessen hat sich als leitende Idee die Auffassung verbreitet, daß man die Prozesse, durch die soziale Sachverhalte und Ereignisse entstehen, durch Verteilungsgesetze charakterisieren könne. Wie wir schon gesehen haben, setzt diese Idee die Vorstellung voraus, daß soziale Sachverhalte und Ereignisse durch Zufallsgeneratoren zustande kommen. Verteilungsgesetze können dann als Wahrscheinlichkeitsverteilungen expliziert werden, die diese Zufallsgeneratoren charakterisieren. So gewinnt auch das Reden von „Gesetzen“ einen zumindest begrifflich nachvollziehbaren Sinn: sie sollen zeigen, wie die Prozesse funktionieren, die soziale Sachverhalte und Ereignisse hervorbringen und deren empirische Verteilungen bestimmen, so wie die Regeln eines Glücksspiels die Verteilung der resultierenden Spielergebnisse bestimmen.

3. Aber gibt es solche Verteilungsgesetze? Die ausgedehnten Diskussionen, die in der Frühphase der Sozialstatistik über diese Frage geführt worden sind, sollen hier nicht verfolgt werden. Durchgesetzt hat sich schließlich eine pragmatische Betrachtungsweise, wie sie am Schluß von Abschnitt 2.4 besprochen worden ist. Außerdem ist die Diskussion um Verteilungsgesetze, so wie sie in der Frühphase der Sozialstatistik geführt worden ist, durch die Entwicklung von Regressionsmodellen gewissermaßen überholt worden. Damit werden wir uns genauer in Kap. 5 beschäftigen. Hier sei zunächst nur auf eine Überlegung von Winkler hingewiesen, die man als eine Art von Rettungsversuch für die Idee der Verteilungsgesetze auffassen kann:

„Es ist das außerordentliche Verdienst Lexis', Quetelets Anschauungen gründlich und wissenschaftlich einwandfrei widerlegt und eine der Wirklichkeit entsprechende Theorie des Zeitverhältnisses der gesellschaftlichen Massenerscheinungen aufgestellt zu haben. Die Widerlegung der Annahme einer absoluten Beständigkeit der allermeisten gesellschaftlichen Erscheinungen schließt natürlich nicht aus, daß es eine verhältnismäßige Beständigkeit in allen biologischen und sozialen Erscheinungen des menschlichen Lebens gibt. [...]
Mit dieser verhältnismäßigen zeitlichen Beständigkeit, die wir im allgemeinen an allen statistischen Entwicklungen beobachten, ist nicht zu verwechseln die Regelmäßigkeit der statistischen Formen, die sich bei der Darstellung *sachlicher* statistischer Reihen (siehe dort[21]) vielfach ergibt. Ihrer ist auch hier Erwähnung zu tun, einmal, weil solche Regelmäßigkeiten erst durch das Gesetz der großen Zahl zum Vorschein kommen, dann, weil der Begriff 'Regelmäßigkeit' in der Statistik vielfach mit demjenigen der Beständigkeit verwechselt wird und dieser wieder, infolge einer Verwechslung der Lexisschen Streuungsuntersuchungen an zeitlichen Reihen mit dem Tatbestande des Gesetzes der großen Zahl, mit diesem Gesetze in Verbindung gebracht wird. Wir haben zwei Dinge streng zu unterscheiden:
1. den Tatbestand des Gesetzes der großen Zahl, durch den die Wesensform der Erscheinungen klargelegt wird. Diese Wesensform kann bei zahlenmäßigen Merkmalen eine regelmäßige Gestalt haben (Beispiel der Körpergrößenverteilung, der Altersgliederung der Heiratenden, der Einkommensverteilung usw.), muß es aber nicht (Beispiel der artmäßigen [Winklers Bezeichnung für qualitative] Merkmale, z.B. Geschlechtsgliederung der Geborenen, Konfessionsgliederung der Bevölkerung, u.dgl.);
2. durch die Lexisschen Streuungsuntersuchungen an zeitlichen Reihen hat sich gezeigt, daß zeitliche Beständigkeit der Wesensform ein seltener Ausnahmefall, Veränderung die Regel ist. Diese Veränderung geht aber meistens nur langsam vor sich." (Winkler 1931, S. 53f)

Winklers Hinweise auf begriffliche Verwechslungen sind durchaus hilfreich. Denn wird in der Sozialstatistik von 'Regemäßigkeit' gesprochen, ist oft nur eine „regelmäßige Form" einer Verteilungsfunktion gemeint. Das aber ist für die Frage nach „sozialen Gesetzmäßigkeiten" ganz unerheblich. Dafür ist vielmehr wichtig, ob bestimmte Sachverhalte oder Zusammenhänge eine zeitliche Beständigkeit aufweisen.[22] Man kann infolgedessen Winklers Ausführungen auch als eine Kritik an der Annahme

[21][Mit „sachlichen statistischen Reihen" meint Winkler statistische Variablen, deren Werte sich auf einen vergleichsweise kurzen Zeitraum beziehen, je nach Art der Sachverhalte z.B. auf einen Tag, einen Monat oder ein Jahr.]

[22]So auch schon Quetelet, wenn er sich nicht in einem Reden über „allgemeine Gesetzmäßigkeiten" ergeht, sondern praktisch wird: „Durch die Kenntnis der Vergangenheit wird es möglich, über die nächste Zukunft zu urtheilen; unsere Konjunkturen können sich oft selbst auf einen Zeitraum von mehreren Jahren erstrecken, ohne dass man befürchten müsste, dass die Zeit Ergebnisse liefern werde, welche gewisse Schranken überschreiten, die sich gleichfalls zum Voraus bezeichnen lassen." (Quetelet 1835, S. 11)

lesen, daß mit Methoden der Sozialstatistik „soziale Gesetzmäßigkeiten" gefunden werden können. Unter praktischen Gesichtspunkten ist aber schließlich nur sein Hinweis wichtig, daß Veränderungen „meistens nur langsam vor sich" gehen. Wenn und insoweit dies der Fall ist, kann man an der Fiktion hypothetischer Zufallsgeneratoren in der Weise festhalten, daß man sie – wie schon von Bernoulli vorgeschlagen wurde – durch einen empirischen Vorbehalt ergänzt: daß ihre Parameter neu geschätzt werden müssen, wenn hinreichend relevante Veränderungen eingetreten sind.

4. Tschuprow hat diesen Gedanken sehr klar in der folgenden Weise zum Ausdruck gebracht:

„Der praktische Wert und somit die Existenz der Statistik als einer besonderen, mit so großem Eifer, wie dies in der Gegenwart geschieht, betriebenen Wissenschaft, beruhen auf der Tatsache der verhältnismäßig geringen Veränderungen der statistischen Zahlen mit der Zeit." (Tschuprow 1906, S. 695)

Tschuprow hat sich auch darum bemüht, die in der Sozialstatistik ermittelbaren „Regelmäßigkeiten" nicht – oder jedenfalls nicht im Sinne von Naturgesetzen – nomologisch zu deuten:

„Denn diese statistischen Gesetzmäßigkeiten haben selber keinen nomologischen Inhalt. Sie sprechen keine allgemeine kausale Beziehungen aus, welche überall und ewig gelten sollen. [...] Man ist zwar vielfach imstande, sie zur Aufdeckung nomologischer Kausalzusammenhänge zu verwerten, und in diesem Sinne bilden sie eine der Grundlagen der statistischen Methode, wie sie im Dienste der nomologischen Wissenschaften vom Naturforscher verwendet wird. An sich sind sie jedoch nicht mehr nomologisch als ähnliche Konstatierungen der Unabänderlichkeit der Verhältnisse etwa auf dem Gebiete der Geographie. Dadurch, daß die Alpen und die Alpentäler sich von Jahr zu Jahr nur höchst unmerkbar verändern, wird zwar deren Beschreibung erleichtert, wenn nicht erst möglich gemacht; diese Tatsache ist jedoch kein Gesetz im Sinne der nomologischen Wissenschaft. Sie bildet die Grundlage für die Existenz des Bädekers für die Schweiz und die Vorbedingung für dessen praktische Anwendbarkeit, wie das Gegenbeispiel der sich stetig und mitunter recht schnell verändernden Gletscher auf das überzeugendste zeigt, für die keine Beschreibung einen Führer ersetzen kann. Hierdurch wird aber der Bädeker nicht zum Range einer naturwissenschaftlichen Untersuchung erhoben, sondern er bleibt das, was er ist: ein wahres Meisterwerk unästhetischer Darstellung des absolut Individuellen auf einem nach allen Richtungen hin sehr eng begrenzten Gebiete.“ (Tschuprow 1906, S. 695f)

Andererseits war Tschuprow nicht der Auffassung, daß die Sozialstatistik sich auf eine Beschreibung gesellschaftlicher Verhältnisse beschränken soll. Sein Ausgangspunkt ist eine Unterscheidung zwischen „nomologischen“ und „idiographischen“ Wissenschaften, wie sie von Autoren der sog. neukantischen Schule ausgearbeitet worden war. Seine Frage war, wie die Sozialstatistik in diese Unterscheidung eingeordnet werden kann. Tschuprow war sich im klaren darüber, daß man die „Regelmäßigkeiten“ der Sozialstatistik nicht im Sinne nomologischer Kausalgesetze verstehen kann. Gleichwohl wollte er ihr nicht nur eine idiographische Aufgabe geben. Er schlug deshalb vor, zwei unterschiedliche Arten nomologischer Aussagen zu unterscheiden. Einerseits nomologische Kausalgesetze, die „Elementarursachen“ mit „Elementarwirkungen“ verknüpfen (Tschuprow 1905, S. 18f), andererseits „losere Arten einer kausalen Beziehung“ (S. 26), bei denen es um Zusmmenhänge zwischen einer „Vielzahl von Ursachen“ und einer „Vielzahl möglicher Wirkungen“ geht:

„... der Inbegriff solcher Forschungsweisen, welche zum Studium der durch die Pluralität der Ursachen und der Wirkungen charakterisierten kausalen Zusammenhänge dienen, ist eben das, was man mit dem Sammelnamen ‘statistische Methode’ zu bezeichnen pflegt.“ (Tschuprow 1905, S. 27)

Diese Andeutungen sollen an dieser Stelle nicht vertieft werden, da wir später genauer besprechen werden, wie in der Entwicklung der theoretischen Sozialstatistik die Idee hypothetischer Zufallsgeneratoren mit Vorstellungen über Kausalität verknüpft worden ist.

5. Tschuprows Überlegungen wurden von Oskar Anderson (1887–1960)[23] fortgesetzt, der ebenfalls eine wichtige Rolle für die Entwicklung einer theoretischen Sozialstatistik gespielt hat; zunächst in seiner „Einführung in die mathematische Statistik“ (1935), dann in seinem Buch „Probleme der statistischen Methodenlehre in den Sozialwissenschaften“ (1957). Der Gedankengang verläuft etwa folgendermaßen:

„Die Gesamtheit unserer wissenschaftlichen Kenntnisse kann in zwei Hauptgruppen zerlegt werden, die der Neukantianer *Windelband*, der diese Unterscheidung besonders schön herausgearbeitet hat, die nomothetische (gesetzesfestlegende) und die idiographische (einzelbeschreibende) nannte. Anstelle von

[23]Kurze biographische Hinweise finden sich bei Johnson und Kotz 1997, S. 176ff, sowie im Vorwort zu Anderson 1963, Bd. 1.

nomothetisch sagt man auch nomologisch oder ontologisch (*Kries*) oder auch nomographisch (*Tschuprow*). [...] Beim ersten Wissenstyp handelt es sich um das Auffinden bzw. den Aufbau von sog. Naturgesetzen, beim zweiten besteht die Aufgabe des wissenschaftlichen Forschers darin, die tatsächliche Verteilung der einzelnen Erscheinungen unserer Umwelt in Raum und Zeit zu beschreiben.“ (Anderson 1957, S. 3)

Wie Tschuprow stellt sich auch Anderson die Frage, wie die Sozialstatistik in diese Unterscheidung eingeordnet werden kann. Dies geschieht zunächst mit folgender Überlegung:

„Man kann in der Hauptsache drei verschiedene Typen des idiographischen Wissens unterscheiden:

a) Unser Interesse konzentriert sich auf das konkrete Bild des zu beschreibenden Objekts, sozusagen auf seine persönliche Geschichte. Dieses ist das historische oder biographische Interesse.

b) Es werden nur gewisse begrenzte Eigenschaften des Objekts ins Auge gefaßt, die ihnen mit einer Menge anderen gemeinsam sind; gleichzeitig wird aber ihre genaue Lage im Raum (manchmal auch in der Zeit) festgesetzt: geographisches oder topographisches Interesse.

c) Es gibt eine sehr große Klasse von idiographischen Kenntnissen, bei welchen nicht bloß gewisse begrenzte Eigenschaften der Objekte registriert, sondern für diese auch nur relativ breite Grenzen in Raum und Zeit festgesetzt werden: dieses ist das eigentliche statistische Interesse. Wie ich bereits angedeutet habe, genügt ein derartiges unvollkommenes Wissen häufig für unsere praktischen Zwecke vollkommen.“ (Anderson 1957, S. 4)

Die statistische Methode wird dann genauer so charakterisiert:

„Mit dem eben geschilderten statistischen Interesse beginnt die statistische Wissenschaft. Wie es Professor *Žižek* (1922) seinerzeit dargelegt hat, verfährt man hierbei im wesentlichen wie folgt:

a) Es werden die Merkmale der betreffenden Beobachtungseinheiten genau definiert. (Das ist lange nicht immer so einfach: Man denke z.B. an solche Begriffe wie Haushalt, Unternehmung, Betrieb, ja sogar Wohnung und dgl.).

b) Es werden – mehr oder weniger willkürlich – gewisse Eingrenzungen im Raum (Stadt, Bezirk, Gemeinde) und Zeit (Jahr, Monat, Tag, zuweilen sogar Stunde, etwa bei der Geburtenstatistik in den großen Städten) festgelegt.

c) Es werden hierauf alle Einheiten, die jene Merkmale besitzen und die in den gegebenen Raum-Zeit-Grenzen liegen, ausgezählt.“ (Anderson 1957, S. 4f)

Soweit handelt es sich um ein idiographisches Verständnis statistischer Methoden, in unserem Fall um deskriptive Sozialstatistik. Andersons Überlegungen gehen dann so weiter:

„Wäre die statistische Methode ein rein idiographisches Verfahren, so würde ihre Darstellung vor allem in der Lehre von der Organisation bzw. Anordnung der statistischen Erhebungen, von der Gruppen- oder Gesamtheitenbildung, von der Auszählung der in ihnen enthaltenen Einheiten und von der Technik der statistischen Aufarbeitung und Aufstellung von Tabellenwerken bestehen. Ferner würde als zweite Aufgabe zu ihr auch die Lehre von der Verdichtung (*Bruns*), Reduktion (*R. A. Fisher*) oder Straffung (*O. Donner*) der Beobachtungsreihen und Tabellen gehören, d.h. die Lehre von der endgültigen (oder auch 'wissenschaftlichen') Aufarbeitung des Zahlenmaterials, die darin besteht, daß ein System von summarischen, durchschnittlichen Charakteristiken der gebildeten Gruppen aufgestellt wird, wie etwa Verhältniszahlen, Durchschnitte, Streuungsmaße, Momente verschiedener Ordnungen, Regressionsgleichungen und dergleichen. [...]

Abgesehen von ihren rein idiographischen Funktionen, ist die statistische Methode dazu berufen, auch einige nomographische Zielsetzungen zu verfolgen. Es ist bemerkenswert, daß im Laufe der letzten 50 Jahre gerade diese immer mehr in den Vordergrund getreten sind und zu großen wissenschaftlichen Erfolgen geführt haben." (Anderson 1957, S. 8f)

Wo sieht nun Anderson die Einsatzstellen, an denen die Sozialstatistik nomographische Aufgaben bekommt? Es sind im wesentlichen drei Überlegungen.

6. Die erste Überlegung bezieht sich darauf, daß man oft nur in der Lage ist, Daten über einen Teil der Elemente einer statistischen Gesamtheit zu gewinnen:

„Gewisse nomographische Elemente dringen zunächst dadurch in die statistische Methode ein, daß unsere statistischen Beobachtungen relativ selten wirklich erschöpfender Natur sind und daß man die entsprechenden statistischen Maßzahlen häufig nur im *Schätzverfahren* gewinnen kann. [...]

Aus dem Gesagten geht hervor, daß eine wichtige Aufgabe der statistischen Theorie darin besteht, Mittel anzugeben, wie man von den Charakteristiken (Maßzahlen) der tatsächlich erhobenen Gesamtheiten, die in Wirklichkeit nur aus Teilen gewisser Gesamtheiten höherer Ordnung bestehen, bis zu den Charakteristiken der letzteren vordringen kann. Diese Gesamtheiten höherer Ordnungen können endweder eine unendliche, besser ausgedrückt, eine unbegrenzte Zahl von Einheiten umfassen, wie es in der Regel von der Schule *R. A. Fishers* angenommen wird, oder eine endliche, aber sehr große Anzahl, oder schließlich eine ganz begrenzte Anzahl derselben." (Anderson 1957, 9ff)

Aber in welcher Weise entsteht hierdurch eine nomographische Aufgabe? Wenn man mithilfe von Daten aus einer Stichprobe Aussagen über eine in unserer Erfahrungswelt real existierende Gesamtheit von Dingen, Menschen oder Situationen machten möchte, kann man dies zwar nur in der Form von Hypothesen über Merkmalsverteilungen (oder daraus ableitbare Maßzahlen) machen; aber diese Hypothesen bekommen dadurch, daß man nur ihre Wahrscheinlichkeit einschätzbar machen kann, keinen nomographischen Charakter. Es sind der Intention nach deskriptive Aussagen über eine statistische Gesamtheit. Allerdings deutet Anderson einen Gedanken an, der – insbesondere durch Fishers Konzeption „hypothetischer unendlicher Gesamtheiten" – in der neueren theoretischen Statistik eine weite Verbreitung gefunden hat: daß man Aussagen über solche fiktiven Gesamtheiten in gewisser Weise nomologisch deuten kann. Mit diesem Gedanken werden wir uns in Kap. 4 beschäftigen.

7. In einer zweiten Überlegung bezieht sich Anderson (1957, S. 12f) auf die Frage nach Regelmäßigkeiten – im Sinne einer relativen zeitlichen Konstanz – statistisch ermittelbarer Erscheinungen und sieht eine Aufgabe, die „nomographische Elemente enthält und ebenfalls zur Wahrscheinlichkeitsrechnung hinüberleitet" (S. 12), darin, statistische Methoden zur Untersuchung und Darstellung solcher Regelmäßigkeiten auszuarbeiten. Es bleibt aber unklar, worin diese Aufgabe wirklich bestehen soll. Man kann vermuten, daß es hierbei um Einsichten in Prozesse gehen soll, durch die soziale Sachverhalte und Ereignisse zustande kommen. Anderson betont jedoch:

„Der Statistiker stellt jedoch nur fest, zu welcher Gruppe eine gegebene Gesamtheit gehört und wie ihre Durchschnittscharakteristiken ausfallen. Eine ganz andere Frage ist es, wie man sich eigentlich den Mechanismus des beständigen 'Sich-selbst-Erneuerns' oder der allmählichen Evolution der betreffenden Gesamtheit zu denken hat. *Diese* Frage sollte prinzipiell nur den Fachleuten des betreffenden Wissensgebietes überlassen werden. Was bei der radioaktiven Strahlung eigentlich vorgeht, hat der Physiker zu untersuchen; wodurch die Konstanz des Verhältnisses der Knaben- und Mädchengeburten hervorgerufen wird, hat der Mediziner bzw. der Erblichkeitsforscher zu entscheiden, und bei der Beurteilung z.B. der Genen- und Chromosomentheorie bleibt der Statistiker als solcher bloß ein Laie. Dasselbe gilt auch für die Darstellung des Mechanismus der Gleichgewichts- und Evolutionserscheinungen im Wirtschaftsleben." (Anderson 1935, S. 16f)[24]

[24]Ich zitiere hier aus Andersons älterer Publikation; etwas weniger prägnant findet sich der gleiche Gedanke jedoch auch bei Anderson 1957, S. 13.

Der Gedankengang führt schließlich zu Andersons dritten Überlegung für das Einsetzen nomographischer Fragestellungen: „die Anwendung der Methoden der statistischen Massenbeobachtung auf die Feststellung kausaler Zusammenhänge.“ (Anderson 1957, S. 13) Damit beschäftigen wir uns in Kap. 6.

Kapitel 3

Wahrscheinlichkeiten und Chancen

Im vorangegangenen Kapitel wurde zu zeigen versucht, wie der theoretischen Sozialstatistik die Idee zugrunde liegt, daß man den Wahrscheinlichkeitsbegriff auch dafür verwenden könne, um Modelle für das Zustandekommen sozialer Sachverhalte und Ereignisse zu konstruieren. Um diese Idee zu verstehen, sind wir – wie in der Wahrscheinlichkeitstheorie – davon ausgegangen, daß sich der Wahrscheinlichkeitsbegriff auf die Vorstellung eines Zufallsgenerators bezieht. Es ist indessen schon deutlich geworden, daß Schwierigkeiten auftreten, wenn man über das Zustandekommen sozialer Sachverhalte und Ereignisse so sprechen möchte, als ob sie durch Zufallsgeneratoren entstehen. Eine besondere Schwierigkeit resultiert daraus, daß Zufallsgeneratoren in der Wahrscheinlichkeitstheorie *als Verfahren* betrachtet werden und sich nicht unmittelbar auf empirische Prozesse beziehen. Die Schwierigkeit besteht darin, daß Prozesse, durch die soziale Sachverhalte und Ereignisse entstehen, in den meisten Fällen nicht als Verfahren verstanden werden können. In diesem Kapitel soll das Problem etwas genauer besprochen werden, um verständlich zu machen, wie sich der Wahrscheinlichkeitsbegriff in der Sozialstatistik zu einem fiktiven Deutungsschema entwickelt hat.

3.1 Schwierigkeiten für den Theorieansatz

In diesem Abschnitt soll zunächst besprochen werden, worin die Schwierigkeiten liegen, wenn man in der Sozialstatistik von Wahrscheinlichkeiten sprechen möchte.

3.1.1 Sinngrenzen der Verfahrensidee

1. Wie schon in Abschnitt 2.1 besprochen worden ist, kann dem Begriff eines Verteilungsgesetzes dann ein nachvollziehbarer Sinn gegeben werden, wenn sich die dafür erforderliche Bezugnahme auf einen Zufallsgenerator empirisch explizieren läßt; das soll heißen: wenn sich der Zufallsgenerator als ein Verfahren beschreiben läßt, das durch empirisch demonstrierbare Prozesse realisiert werden kann. Wenn dies der Fall ist, kann man ein Verteilungsgesetz als Charakterisierung der Funktionsweise eines Verfahrens verstehen. Exemplarisch kann man an Glücksspiele denken, die sich in ihrer Funktionsweise durch ein Verteilungsgesetz charakterisieren lassen. Die bisher zur Diskussion gestellten Vorbehalte gegen das Reden von Verteilungsgesetzen in der Sozialstatistik haben sich hauptsächlich darauf bezogen, daß unklar bleibt, ob und ggf. wie soziale Prozesse als Realisationen von Verfahren verstanden werden können. Das Argument bestand insoweit darin, daß – solange eine solche Explikation nicht erfolgt – die gedankliche Bezugnahme auf Zufallsgeneratoren rein fiktiv bleibt.

2. Aber auch wenn man diesen Einwand zunächst zurückstellt, bleibt noch eine weitere Frage: in welcher Weise durch eine gedankliche Bezugnahme auf Zufallsgeneratoren irgendetwas *erklärt* werden könnte. Denn mit Überlegungen, die sich an der Vorstellung eines Zufallsgenerators orientieren, kann nicht erklärt werden, ob und wann Ereignisse auftreten, die durch den Zufallsgenerator erzeugt werden *könnten*. Wenn man einen Zufallsgenerator durch ein Verteilungsgesetz beschreibt, macht man eine Aussage darüber, wie das Verfahren funktioniert, aber nicht darüber, ob, wann und wo es zu Realisationen des Verfahrens kommt. Man denke z.B. an ein Würfelspiel. Wenn es durch eine bestimmte Wahrscheinlichkeitsverteilung charakterisiert wird, wird in keiner Weise darauf bezug genommen, ob, wo, wann und von wem gewürfelt wird. Die Wahrscheinlichkeitsaussagen beziehen sich nicht auf empirische Ereignisse, die in unserer Erfahrungswelt stattfinden, sondern ausschließlich auf ein Verfahren. Hier knüpft auch die Idee der Wiederholbarkeit an. Zu einem Verfahren gehört die Idee, daß es wiederholt realisiert werden kann. Ohne eine zumindest gedankliche Bezugnahme auf Wiederholungen würde es keinen Sinn machen, sich um ein Verständnis der Eigenschaften eines Verfahrens zu bemühen.

3. Es ist vielleicht hilfreich, zur gedanklichen Orientierung ein Bild zu verwenden:

$$\boxed{\mathcal{G}} \longrightarrow s$$

Der Zufallsgenerator $\mathcal{G}$ kann verwendet werden, um Situationen (oder Sachverhalte oder Ereignisse) eines gewissen Typs $(\mathcal{S}, \tilde{X})$ zu erzeugen. Jedesmal, wenn der Zufallsgenerator aktiviert wird, entsteht eine Situation des Typs $\mathcal{S}$, die durch einen Wert im Merkmalsraum $\tilde{X}$ charakterisiert werden kann. Stattdessen kann man auch von einer Zufallsvariablen X^* sprechen, die Werte im Merkmalsraum $\tilde{X}$ annehmen kann. Ein Verteilungsgesetz charakterisiert den Zufallsgenerator bzw. die ihn repräsentierende Zufallsvariable durch eine Wahrscheinlichkeitsverteilung. Man gewinnt dadurch eine Einsicht in die Funktionsweise des Zufallsgenerators, so daß sich Fragen der folgenden Art beantworten lassen: *Wenn* der Zufallsgenerator aktiviert wird, mit welchen Wahrscheinlichkeiten können verschiedene Arten von Situationen auftreten?

4. Man kann die Frage auch so formulieren: Mit welchen Wahrscheinlichkeiten nimmt eine Zufallsvariable ihre verschiedenen möglichen Werte an? Aber diese verkürzte Formulierung kann leicht irreführend werden. Denn der Begriff eines Verteilungsgesetzes für eine Zufallsvariable impliziert keinerlei Aussage darüber, ob, wann und wo die Zufallsvariable überhaupt irgendeinen Wert annimmt. Man kann exemplarisch an einen Vorschlag von Laplace (1814, S. 51f) denken, daß man sich das Zustandekommen des Geschlechts von Neugeborenen so vorstellen könne, daß bei jeder Geburt eine Ziehung aus einer Urne vorgenommen wird, die zwei Sorten von Kugeln enthält. Kennt man in diesem Fall das Verteilungsgesetz, also die Proportionen, in denen die Urne die beiden Sorten von Kugeln enthält, kann man sagen, mit welcher Wahrscheinlichkeit es ein Junge oder ein Mädchen wird. Aber man gewinnt daraus kein Wissen darüber, ob es überhaupt zu einer Geburt kommen wird. Darüber, ob, wann und wo ein Prozeß inganggesetzt wird, der schließlich dazu führt, daß ein Junge oder ein Mädchen geboren wird, macht das Modell nicht nur keine Aussage, sondern es liefert auch keinerlei Ansatzpunkt, um über diese Frage nachdenken zu können.

5. Werden Begriffsbildungen und Vorstellungen der Wahrscheinlichkeitsrechnung in einem verfahrenstechnischen Kontext verwendet, ist der Sachverhalt leicht durchschaubar: man möchte herausfinden, wie ein Verfahren funktioniert, und man abstrahiert gerade deshalb von der Frage, ob, wann und wo es sinnvoll sein könnte, das Verfahren anzuwenden. Die Einsicht, daß Zufallsgeneratoren zunächst als Verfahren konzipiert sind, geht jedoch leicht verloren, wenn man soziale Prozesse durch Zufallsgeneratoren zu deuten versucht. Da es diese Prozesse in der sozialen Realität immer schon gibt, kann leicht der Eindruck entstehen, daß man durch eine gedankliche Bezugnahme auf Zufallsgeneratoren auch das In-Erscheinung-Treten sozialer Sachverhalte und Ereignisse erklären könne. Diesen falschen Eindruck zu kritisieren, genügt indessen nicht. Denn in der Rhetorik der theoretischen Sozialstatistik geht es ihr ja gerade darum, Modelle zu konstruieren, mit denen *empirisch realisierte* Häufigkeitsverteilungen erklärt werden können. Somit entsteht eine bemerkenswerte Frage: Wie kann der Anschein einer solchen Erklärung erzeugt werden, wenn man für die Modellkonstruktion Zufallsgeneratoren verwendet, die sich nur auf Verfahren beziehen? Offenbar muß man in irgendeiner Weise Annahmen darüber einführen, wie, wo und wann die Zufallsgeneratoren tatsächlich aktiviert werden.

6. Um das Problem genauer zu verstehen, beziehe ich mich auf Überlegungen von L. v. Bortkiewicz (1898), in denen er versucht hat, die Poisson-Verteilung als ein Verteilungsgesetz – von ihm „Gesetz der kleinen Zahlen“ genannt – zu deuten, mit dessen Hilfe seltene Ereignisse erklärt werden können. Das erste Kapitel seines Buches beginnt mit folgenden Bemerkungen:

„Bezeichnet p die Wahrscheinlichkeit des Eintretens eines bestimmten Ereignisses A und $q = 1 - p$ die Wahrscheinlichkeit des Nichteintretens von A bei *einem* Versuch, so stellt

$$\frac{n(n-1)\cdots(n-x+1)}{1\cdot 2\cdots x}\,p^x q^{n-x} \tag{1}$$

die Wahrscheinlichkeit des x maligen Eintretens von A bei n Versuchen dar.

Läßt man die Zahl n *in infinitum* anwachsen und p bis auf Null herabsinken und zwar so, daß das Produkt $np = m$ dabei stets unverändert bleibt, so wird sich (1) dem Grenzwert

$$w_x = \frac{m^x\,e^{-m}}{1\cdot 2\cdots x} \tag{2}$$

nähern. Die Formel setzt voraus, daß die Zahl x im Verhältnis zu der Zahl n klein ist.[1]

Die Größe m, welche, ihrem Begriff nach, positiv sein muß, aber sowohl eine ganze Zahl als ein echter oder unechter Bruch sein kann, drückt die mathematische Erwartung der Zahl der Fälle aus, bei denen das Ereignis A unter n Fällen oder Versuchen vorkommt.“ (v. Bortkiewicz 1898, S. 1)

Soweit lassen sich die Ausführungen einfach nachvollziehen, indem man sich gedanklich auf artifizielle Zufallsgeneratoren bezieht, z.B. auf ein Urnenschema. Darauf deutet auch v. Bortkiewicz’ Verwendung des Wortes ‘Versuch’ hin. Der Übergang von (1) zu (2) wurde zuerst von S. D. Poisson (1781–1840) vorgeschlagen, um bei kleinen Ereigniswahrscheinlichkeiten eine bessere Approximation für die Binomialverteilung (1) zu gewinnen, und (2) wird seither als Poisson-Verteilung bezeichnet.

7. Im weiteren Verlauf des ersten Kapitels seines Buches bespricht v. Bortkiewicz eine Reihe von Eigenschaften der Poisson-Verteilung, die

[1] [Diese Zusatzbedingung ist für den Konvergenzbeweis nicht erforderlich; aber dieser Punkt ist hier nicht weiter von Belang.]

uns hier nicht zu interessieren brauchen. Zu überlegen sind jedoch die sozialstatistischen Anwendungen, die v. Bortkiewicz in seinem zweiten Kapitel vornimmt. Es beginnt mit folgenden Bemerkungen:

„Es soll nunmehr der Versuch gemacht werden, die Formeln des ersten Kapitels auf statistische Reihen absoluter Zahlen anzuwenden, welche für eine Reihe von Kalenderjahren angeben, wieviel Male im Laufe eines jeden Jahres ein bestimmtes Ereignis in einer gegebenen Gesellschaft von Menschen vorgekommen ist. Hierbei werden solche Beispiele gewählt, bei denen die Bedingung erfüllt ist, daß den einzelnen Gliedern der ins Auge gefaßten statistischen Reihe jeweils sehr große Zahlen von Beobachtungen bezw. von beobachteten Menschen entsprechen (mindestens einige Tausend), während die Zahlen selbst, aus denen sich die statistische Reihe zusammensetzt, kleine Zahlen sind (nicht über 20 hinausgehend).“ (v. Bortkiewicz 1898, S. 17)

v. Bortkiewicz bespricht dann insgesamt vier Beispiele. Zur Illustration beziehe ich mich auf das erste Beispiel: Selbstmorde von Kindern in Preußen:

„Nachstehende Tabelle [Tabelle 1] enthält die Zahlen der von Knaben und Mädchen unter 10 Jahren in Preußen im Zeitraume 1869–1893 begangenen Selbstmorde. [...]

Unsere Aufgabe besteht nun darin, zu prüfen, ob die angeführten Ergebnisse der Statistik auf das uns aus dem ersten Kapitel bekannte Schema der Wahrscheinlichkeitsrechnung zurückgeführt werden können. Hierbei hat man sich zu fragen: in wieviel Fällen aus 25 würden sich die Ergebnisse 0, 1, 2 usw. am wahrscheinlichsten einstellen, gesetzt, daß jenes Schema zuträfe? Die Antwort wird durch die Produkte $25 \cdot w_x$ geliefert, wobei w_x dieselbe Bedeutung hat wie in § 1 [vgl.o. die Formel (2)]. Nur hat man in die maßgebende Formel (2) des § 1 anstatt der Unbekannten m den wahrscheinlichsten Werter ihrer, nämlich den Mittelwert

$$m' = \frac{49}{25} = 1.96$$

einzusetzen. Auf diese Weise läßt sich nachfolgende Zahlenreihe berechnen:

x	$25 \cdot w_x$	x	$25 \cdot w_x$
0	3, 4	5	0.9
1	6, 8	6	0.3
2	6, 8	7	0.1
3	4, 5	8	0.0
4	2, 2		

Nennt man, nach Lexis’ Vorgang, *Dispersion* die Art, wie sich die Glieder einer statistischen Reihe um den Mittelwert der Reihe verteilen, oder anders

Tabelle 1 Selbstmorde von Kindern in Preußen.

Jahr	Knaben	Mädchen	Zusammen
1869	2	1	3
70	3	-	3
71	1	1	2
72	4	-	4
73	1	1	2
74	4	-	4
75	-	2	2
76	3	1	4
77	-	-	-
78	1	-	1
79	2	-	2
80	6	-	6
81	3	-	3
82	4	-	4
83	1	-	1
84	-	-	-
85	2	-	2
86	2	-	2
87	1	-	1
88	1	1	2
89	-	-	-
90	2	1	3
91	1	1	2
92	1	1	2
93	4	1	5
Im ganzen	49	11	66

das Bild von den Schwankungen, welche eine statistische Reihe darbietet, so kann man sagen, daß es sich nunmehr darum handelt, die *erwartungsmäßige* Dispersion der Elemente x, wie sie sich in der Reihe der Werte $25 \cdot w_x$ darstellt, der *effektiven* Dispersion der nämlichen Elemente, wie diese in der Reihe der Werte l_x zum Ausdruck kommt, gegenüberzustellen, und zuzusehen, ob beide Dispersionen in dem Maße übereinstimmen, daß die Abweichungen der Größen l_x von den entsprechenden Größen $25\,w_x$ als *zufällige* im Sinne der Wahrscheinlichkeitsrechnung gedeutet werden können.“ (v. Bortkiewicz 1898, S. 17ff)

8. v. Bortkiewicz vergleicht dann die „theoretisch zu erwartenden“ mit den tatsächlich beobachteten Häufigkeiten und prüft auch die Streuung der jährlichen Anzahlen, wobei er sich an Lexis’ Dispersionstheorie orientiert. Uns interessieren hier jedoch nicht diese Goodness-of-Fit-Überlegungen, vielmehr die Frage, in welcher Weise in diesem Beispiel die Idee eines Zufallsgenerators verwendet wird. v. Bortkiewicz’ Gedankengang unterstellt, daß es einen Zufallsgenerator $\mathcal{G}$ gibt, der seine Daten erzeugt haben könnte. Stattdessen kann man sich auch auf eine Zufallsvariable X^* beziehen, als deren Realisationen die Daten betrachtet werden. Die Daten bestehen aus 25 Werten (bezugnehmend auf die Selbstmorde der Knaben, die auch v. Bortkiewicz in der o.a. Rechnung verwendet). v. Bortkiewicz’ Frage ist, ob man für diesen hypothetisch unterstellten Zufallsgenerator eine Poisson-Verteilung annehmen kann; und ergänzend, ob man für alle 25 Jahre den gleichen Zufallsgenerator unterstellen kann. Beide Fagen werden von v. Bortkiewicz bejaht. Aber ob zu Recht, ist hier nicht unsere Frage. Könnten die Daten nicht mit einer Poisson-Verteilung beschrieben werden, würde – aus dem Ansatz des Gedankengangs v. Bortkiewicz’ – bloß folgen, daß man nach einer anderen Verteilung suchen müßte. Hier interessiert vielmehr, welche Vorstellungen über das Zustandekommen von Selbstmorden durch die Annahme eingeführt werden, daß sie als Ergebnisse eines (irgendeines) Zufallsgenerators entstehen.

9. Die Schwierigkeit liegt darin, daß man sich in diesem Fall den Zufallsgenerator nicht als ein Verfahren vorstellen kann, jedenfalls nicht im üblichen Sinne des Wortes ‘Verfahren’. Denn beim üblichen Verständnis setzt das Wort einen Akteur voraus, der das Verfahren anwendet und dadurch Prozesse ingangsetzt, die als Realisationen des Verfahrens betrachtet werden können. Die Frage ist: Wer betätigt den Zufallsgenerator $\mathcal{G}$, der mit einer gewissen Wahrscheinlichkeit Selbstmorde entstehen läßt? Sind es die Kinder, von denen einige schließlich einen Selbstmord begehen? Vielleicht soll man es sich so vorstellen: es gibt einen Zufallsgenerator, und einmal pro Jahr zieht jedes Kind entweder eine weiße oder schwarze Kugel, wodurch dann über das weitere Schicksal entschieden wird. Die Vorstellung erscheint zwar merkwürdig, könnte aber in spekulativer Weise einen Anschluß an den wahrscheinlichkeitstheoretischen Begriff eines Zufallsgenerators herstellen. Es bleibt jedoch eine bemerkenswerte Besonderheit im Vergleich zu anderen Beispielen, in denen von einer Analogie zu Zufallsgeneratoren Gebrauch gemacht wird. Als Beispiel kann man an Laplace’ Urne denken, mit der er sich den Prozeß vorstellbar machen möchte, durch den darüber entschieden wird, ob ein

Mädchen oder ein Junge geboren wird. In diesem Beispiel kann man sich zumindest gedanklich einen Prozeß vorstellen, auf dessen Ablauf sich der Zufallsgenerator beziehen soll; und man kann sich auch vorstellen, wie der Prozeß inganggesetzt werden kann. Aber bei den Selbstmorden von Kindern ist ein Prozeß, auf dessen Ablauf man sich mit der Idee eines Zufallsgenerators beziehen könnte, nicht erkennbar.

3.1.2 Akteure und Beobachter

1. Die Schwierigkeiten entstehen daraus, daß es in der Sozialstatistik erforderlich wird, zwischen Akteuren und Beobachtern zu unterscheiden. Bei der gedanklichen Bezugnahme auf einen Zufallsgenerator ist das nicht notwendig. Der Zufallsgenerator kann als ein Verfahren konzipiert werden, und es kann offenbleiben, ob, wann, wo und von wem der Zufallsgenerator aktiviert wird. Der Statistiker, der Einsichten in die Funktionsweise eines Zufallsgenerators gewinnen möchte, kann sich beliebige Subjekte vorstellen, die ihn betätigen, weil die Theoriebildung darauf keinen Bezug nimmt. Infolgedessen kann er auch mit den Subjekten, die den Zufallsgenerator möglicherweise aktivieren, eine gemeinsame erkenntnistheoretische Perspektive einnehmen.

2. Eine ganz entsprechende Betrachtungsweise kann auch ein Psychologe einnehmen, der mit Menschen oder Tieren Experimente durchführt und sich für die Theoriebildung der Konzeption eines Zufallsgenerators bedient. Die Subjekte, mit denen der Psychologe experimentiert, kommen dann nicht als Subjekte eines Verfahrens in Betracht – das ist vielmehr der Psychologe selbst –, sondern nur als Material, dessen Eigenschaften und Reaktionsweisen mithilfe des Verfahrens geprüft werden sollen. Für die Theoriebildung kann infolgedessen davon abstrahiert werden, daß es sich bei den Subjekten, mit denen ein Psychologe experimentiert, selbst um Akteure handelt. Es sind nicht sie, die – als Akteure – das experimentelle Verfahren ingang setzen, sondern dies geschieht durch den Psychologen.

3. Der Sozialstatistiker kann sich indessen das Geschehen in seinem Gegenstandsbereich nicht in der Weise unterwerfen, wie es ein Techniker mit Maschinen, ein Biologe mit Pflanzen oder ein Psychologe mit Tieren und Menschen machen kann. Er kann nicht ohne weiteres davon abstrahieren, daß die sozialen Sachverhalte und Ereignisse, für deren Zustandekommen er sich interessiert, durch Akteure hervorgebracht werden, die ihrem eigenen Willen folgen. Er kann zunächst nur beobachten, wie die in der Gesellschaft tätigen Akteure sich verhalten und Sachverhalte und Ereignisse hervorbringen. Infolgedessen gibt es keinen unmittelbaren Anknüpfungspunkt für eine Konzeption von Zufallsgeneratoren als Verfahren. Die Vorstellung, daß die Tätigkeiten sozialer Akteure darin bestehen, daß sie Zufallsgeneratoren aktivieren, die dann ihrerseits mit gewissen Wahrscheinlichkeiten Sachverhalte und Ereignisse hervorbringen, liefert in den meisten Fällen kein plausibles Bild sozialer Aktivitäten. — Hier kann jedoch ein Wechsel in der Rhetorik einsetzen. Um dennoch der Vorstellung, daß soziale Sachverhalte und Ereignisse durch Zufallsgeneratoren hervorgebracht werden, einen Sinn geben zu können, kann der Sozialstatistiker versuchen, an ein Reden von „Chancen“ anzuknüpfen.

4. Das Wort hat zunächst einen handlungspraktischen Sinn: Menschen können über Handlungschancen nachdenken und kommunizieren. Das Wort bezieht sich in diesem Kontext auf Handlungsmöglichkeiten, die man ausfindig machen, wahrnehmen und ggf. realisieren kann. Allerdings ist dieses handlungspraktische Verständnis des Chancenbegriffs für den theoretischen Sozialstatistiker ganz unbrauchbar. Denn in diesem Sinn impliziert der Begriff eine Bezugnahme auf Akteure, die als Subjekte der Wahrnehmung von Handlungschancen vorstellbar sind. Der Sozialstatistiker interessiert sich jedoch nicht dafür, wie Subjekte ihre Handlungsmöglichkeiten wahrnehmen und einschätzen; sondern er will berechenbar machen, wie sich soziale Akteure unter jeweils gegebenen Bedingungen verhalten. Er braucht deshalb einen Chancenbegriff, der nicht Subjekten im Sinne von Handlungschancen, sondern den Situationen, in denen sich soziale Akteure befinden, zurechenbar gemacht werden kann. Vorauszusetzen ist natürlich ein theoretisch zu vollziehender Bruch mit einem handlungspraktischen Erkenntnisinteresse. Der Sozialstatistiker ist nicht an handlungspraktischen Fragestellungen interessiert, also nicht an Fragestellungen, die sich auf Handlungsmöglichkeiten aus der Perspektive von Handlungssubjekten richten; sondern er möchte einschätzbar machen, wie sich Menschen unter jeweils gegebenen Umständen tatsächlich verhalten bzw. welche Folgen ihre Handlungen haben werden.[2]

5. Hier kann nun der theoretische Sozialstatistiker an zwei Gedankengänge anknüpfen. Erstens an eine „soziologische Wende“ in der Gesellschaftstheorie, durch die der Gedanke verbreitet worden ist, daß das

[2]Man vgl. das in Abschnitt 2.1 angeführte Zitat von Lindsey. Natürlich wird dadurch ein praktisches Interesse nicht ausgeschlossen. Exemplarisch kann man daran denken, daß Unternehmen wissen wollen, wie die potentiellen Käufer ihrer Waren auf Werbekampagnen reagieren.

Verhalten von Menschen durch ihre gesellschaftlichen Verhältnisse bedingt oder sogar bestimmt wird.[3] Der Gedanke ist auch deshalb bemerkenswert, weil er sich sowohl mit einer handlungspraktischen als auch mit einer sozialstatistischen Betrachtungsweise verbinden läßt. Aus der handlungspraktischen Perspektive erscheint es unmittelbar plausibel, daß Handlungsmöglichkeiten auch davon abhängen, in welchen Situationen sich Handlungssubjekte befinden. Zwar macht aus dieser Perspektive die Vorstellung, daß Handlungsweisen durch Situationen *bestimmt* werden, keinen Sinn. Aber schon die Frage, die sich auf Handlungsmöglichkeiten richtet, impliziert eine Bezugnahme auf Eigenschaften einer Situation. — Andererseits kann auch der Sozialstatistiker an den Gedanken anknüpfen, daß Handlungsmöglichkeiten von Menschen durch Eigenschaften einer jeweils gegebenen Situation – und ggf. auch ihnen selbst zurechenbare Eigenschaften – bedingt oder bestimmt werden.

6. Nun ist der Sozialstatistiker jedoch nicht an Handlungsmöglichkeiten interessiert, sondern daran, wie sich Menschen tatsächlich verhalten, bzw. an den resultierenden Sachverhalten und Ereignissen. D.h. der Sozialstatistiker ist nicht an Aussagen über Möglichkeiten im Sinne von Handlungsmöglichkeiten interessiert, sondern daran, mögliche Verhaltensweisen, Sachverhalte, Ereignisse durch Wahrscheinlichkeitsaussagen einschätzbar zu machen. Und an dieser Stelle liegt es nahe, sich an Begriffsbildungen und Vorstellungen der Wahrscheinlichkeitstheorie zu orientieren. Deren Ausgangspunkt ist ja ebenfalls das Interesse, Möglichkeiten durch Wahrscheinlichkeiten einschätzbar zu machen. Um dieses Interesse theoretisch verfolgen zu können, dient in der Wahrscheinlichkeitstheorie die Konzeption von Zufallsgeneratoren. Sie werden als Verfahren konzipiert, um durch eine begriffliche Abkopplung von Prozessen und Ereignissen, die in der Erfahrungswelt „autonom“ auftreten, eine (mathematische) Theoriebildung zu ermöglichen. Zwar kann die Idee eines Zufallsgenerators, wie sie in der Wahrscheinlichkeitstheorie konzipiert wird, nicht bruchlos übernommen werden, da die intendierten Anwendungen nicht verfahrenstechnisch expliziert werden können. Der Sozialstatistiker kann jedoch versuchen, den Begriff eines Zufallsgenerators für seine Zwecke umzudeuten.

[3] Der Gedanke hat eine lange Tradition, die hier nicht besprochen werden kann. Exemplarisch kann man an folgende Bemerkung von P. L. Berger (1963, S. 106) denken: „Die Gesellschaft schreibt uns nämlich nicht nur vor, was wir zu tun, sondern auch, wer wir zu sein haben. Mit anderen Worten: Unser gesellschaftlicher Ort bestimmt nicht nur unser Verhalten, sondern auch unser Sein.“

7. Der Umdeutungsprozeß beginnt beim Chancenbegriff. Der Begriff spielt deshalb eine zentrale Rolle, weil durch ihn Möglichkeiten mit Eigenschaften einer Situation verknüpft werden. Aus handlungspraktischer Sicht ist das unmittelbar verständlich: was ein Mensch tun *kann*, hängt auch von der Situation ab, in der er sich befindet. Das Erkenntnisinteresse des Sozialstatistikers kann hieran durch eine scheinbar geringfügige Veränderung der Formulierung anknüpfen: was ein Mensch tun *wird*, hängt auch von der Situation ab, in der er sich befindet.[4] Ist man jedoch bei dieser Formulierung, gibt es eine Anschlußstelle für den Wahrscheinlichkeitsbegriff. Eine Situation kann dann als eine Gegebenheit betrachtet werden, in der sich unterschiedliche Möglichkeiten realisieren können; und man kann versuchen, die Realisierungsmöglichkeiten durch Wahrscheinlichkeiten einschätzbar zu machen. Es ist dann auch nicht erforderlich, ausschließlich an Handlungen von Akteuren zu denken, die in einer Situation realisiert werden können. Denn hat man den Situationsbegriff aus seiner ursprünglichen Verklammerung mit einer handlungspraktischen Perspektive herausgelöst, so daß man über Situationen als in der Realität vorhandene Gegebenheiten sprechen kann, kann man ihn auch mit beliebigen Vorstellungen über Prozesse verknüpfen, die in Situationen entstehen und ablaufen können. Man kann z.B. an eine unübersichtliche Straßenkreuzung denken, bei der mit einer gewissen Wahrscheinlichkeit Verkehrsunfälle stattfinden.

8. Bemerkenswert ist, wie auf diese Weise ein spezifisch sozialstatistischer Chancenbegriff entsteht. Er unterscheidet sich einerseits, wie schon deutlich geworden sein sollte, von einem handlungspraktischen Chancenbegriff. Er unterscheidet sich indessen auch von einem Chancenbegriff, wie er in der Wahrscheinlichkeitstheorie durch eine gedankliche Bezugnahme auf Zufallsgeneratoren konzipiert wird. Zwar kann man in beiden Fällen Chancen und Wahrscheinlichkeiten begrifflich identifizieren. Dennoch gibt es einen bemerkenswerten Unterschied in der empirischen Explikation. Bei der gedanklichen Bezugnahme auf einen Zufallsgenerator bleibt offen, ob, wann, wo und von wem der Zufallsgenerator betätigt wird, so daß tatsächlich Sachverhalte oder Ereignisse entstehen. In diesem Kontext beziehen sich Wahrscheinlichkeitsaussagen auf die Funktionsweise von Verfahren. In der Sozialstatistik verändert sich

[4] Der Situationsbegriff sollte hier stets so verstanden werden, daß auch Eigenschaften von Menschen als Merkmale ihrer Situation aufgefaßt werden können. Für den gegenwärtigen Gedankengang ist es nicht erforderlich, genauer auf die Frage einzugehen, ob und ggf. wie eine Unterscheidung zwischen Situationen und Menschen zurechenbaren Eigenschaften gemacht werden kann.

indessen die Art und Weise, wie die Konditionalisierung vorgenommen wird. Der Chancenbegriff bezieht sich nicht länger auf ein Verfahren, sondern auf eine Situation, in der mit gewissen Wahrscheinlichkeiten mögliche Vorgänge realisiert werden. Es ist dann die Situation – z.B. eine Straßenkreuzung –, die mit einer gewissen Wahrscheinlichkeit Unfälle hervorbringt.

9. Auf diese Weise kann durch den sozialstatistischen Chancenbegriff eine neue Rhetorik eingeführt werden: man kann über Chancen so reden, als ob es sich um Eigenschaften einer Situation handelt. Dies ermöglicht der Sozialstatistik auch einen neuen Blick auf soziale Akteure. Es ist nicht länger erforderlich, Menschen als Akteure sozialer Prozesse zu betrachten, sondern man kann sich darauf beschränken, sie als Vollzugsgehilfen für die Realisierung von Chancen zu betrachten, die als Eigenschaften einer jeweils gegebenen Situation vorausgesetzt werden. Dadurch ist man einerseits dem Descartes'schen Ideal näher gekommen: daß jede Vorstellung von Akteuren aus der wissenschaftlichen Naturbetrachtung verbannt werden sollte. Andererseits entsteht die Möglichkeit, in einer nicht handlungspraktischen Perspektive über gesellschaftliche Verhältnisse zu sprechen: man kann sie als einen Komplex von Situationen konzeptionalisieren, die durch ihnen zurechenbare Chancen charakterisiert werden können.

10. Die Bezüge zur Entwicklung der Soziologie, insbesondere zur Entwicklung des Sozialstrukturbegriffs, können hier nicht näher verfolgt werden.[5] Die eben skizzierten Überlegungen lassen indessen erkennbar werden, wie sich durch die Verwendung des Chancenbegriffs auch Vorstellungen über Aufgaben einer theoretischen Sozialstatistik entwickeln konnten. Die auf den ersten Blick merkwürdige Idee, empirisch ermittelbare Häufigkeitsverteilungen als Realisierungen von „zugrunde liegenden" Wahrscheinlichkeitsverteilungen aufzufassen, erscheint zumindest rhetorisch dadurch vermittelbar, daß man die Wahrscheinlichkeitsverteilungen als „Beschreibungen" von Chancen betrachtet, die sozialen Situationen als Eigenschaften zurechenbar sind. So könnte man auch versuchen, in v. Bortkiewicz' Vorgehensweise einen Sinn zu finden, der sich nicht darin erschöpft, für eine gegebene Häufigkeitsverteilung nur nach einer anderen Darstellungsform zu suchen. Die in dem Beispiel von ihm intendierte Aussage könnte dann vielleicht so verstanden werden: die Situation, in der Kinder im Zeitraum 1869–1893 in Preußen gelebt haben, läßt sich durch bestimmte Chancen für das Zustandekommen von Selbstmorden charakterisieren.[6] Zwar erfährt man daraus nichts über die Ursachen der Selbstmorde. Wir werden jedoch später sehen, wie in der theoretischen Sozialstatistik versucht worden ist, auch diesen Vorbehalt durch die Entwicklung einer neuen Rhetorik hinfällig zu machen. Der Grundgedanke läßt sich an dieser Stelle schon erkennen: wenn man soziale Situationen auf eine solche Weise durch Chancen charakterisieren kann, daß eine Bezugnahme auf Akteure überflüssig erscheint, kann es sinnvoll erscheinen, in den Chancen zugleich die Ursachen dafür zu sehen, welche Sachverhalte und Ereignisse in den Situationen realisiert werden.

[5]Nur beiläufig sei auf Max Weber hingewiesen, der als einer der ersten Soziologen versucht hat, den Gegenstandsbereich der Soziologie durch einen sozialstatistischen Chancenbegriff zu konzipieren. Einen Hinweis geben z.B. folgende Bemerkungen: „Soziale 'Beziehung' soll ein seinem Sinngehalt nach aufeinander gegenseitig *eingestelltes* und dadurch orientiertes Sichverhalten mehrerer heißen. Die soziale Beziehung *besteht* also durchaus und ganz ausschließlich: in der *Chance*, daß in einer (sinnhaft) angebbaren Art sozial gehandelt wird, einerlei zunächst: worauf diese Chance beruht." „Eine soziale Beziehung kann ganz vorübergehenden Charakters sein oder aber auf Dauer, d.h. derart eingestellt sein: daß die Chance einer kontinuierlichen *Wiederkehr* eines sinnentsprechenden (d.h. dafür geltenden und demgemäß erwarteten) Verhaltens besteht. *Nur* das Vorliegen dieser Chance: – der mehr oder minder großen *Wahrscheinlichkeit* also, daß ein sinnentsprechendes Handeln stattfindet, und *nichts* darüber hinaus – bedeutet der *'Bestand'* der sozialen Beziehung, was zur Vermeidung falscher Vorstellungen stets gegenwärtig zu halten ist." (Weber 1921, S. 13f)

[6]Eine Fülle von Beispielen für eine solche Rhetorik findet sich in dem Buch „What are the chances?" von B. Siskin, J. Staller und D. Rorvik, 1989.

3.2 Der Theorieansatz von W. Lexis

Der im vorangegangenen Abschnitt skizzierte Gedankengang wurde in systematischer Form zum erstenmal von Wilhelm Lexis (1837–1914) zur Grundlegegung einer theoretischen Sozialstatistik verwendet. Da seine Überlegungen sehr einflußreich geworden sind, sollen sie in diesem Abschnitt etwas genauer besprochen werden.[7]

[7]Biographische Hinweise finden sich bei Johnson und Kotz 1997, S. 305f, und bei Hess, 1978. Über Lexis' Einfluß schrieb Tschuprow (1905, S. 13), daß „die Entwicklung der modernen theoretischen Statistik durchweg an die Lexis'schen Untersuchungen anknüpft." Und weiter: „... daß in allen Ländern, wo sich heute auf dem Gebiete der statistischen Theorie etwas regt, die Anfänge der Bewegung auf Lexis zurückgehen. Der mächtige Aufschwung der statistischen Theorie in England, das nahe daran ist, die Führerschaft auf diesem Gebiete zu übernehmen, ist neben Fr. Galton von Fr. Y. Edgeworth eingeleitet worden, dessen Arbeiten von unmittelbarem Einflusse Lexis' zeugen. Der Däne Westergaard, der Russe v. Bortkiewicz, um von den deutschen Schülern Lexis' nicht zu reden, schulden Lexis die Elemente ihrer Auffassungsweise und wohl auch den ersten Impuls zu theoretischer Arbeit auf dem Gebiete der Statistik." (S. 70)

3.2.1 Bezugnahme auf Lebensverläufe

1. Lexis' erste einflußreiche Publikation ist die „Einleitung in die Theorie der Bevölkerungsstatistik" (1875). Sie beginnt mit folgender Bemerkung:

„Die Aufgabe der wissenschaftlichen Bevölkerungsstatistik besteht in der methodischen Gruppirung und Untersuchung der Thatsachen, die sich aus der exacten Massenbeobachtung der allgemein bedeutsamen Lebensmomente der menschlichen Individuen ergeben." (Lexis 1875, S. 1)

Etwas genauer werden dann die Aufgaben der Bevölkerungsstatistik folgendermaßen angegeben:

„Das bevölkerungsstatistische Material wird durch unmittelbare Beobachtung an den Individuen gewonnen, nicht also durch Beobachtung der Massenergebnisse der menschlichen Thätigkeit (Production, Handel u.s.w.). Jedoch schließen wir diejenigen Untersuchungen aus, welche die menschlichen Individuen wesentlich nach ihrer physischen Seite, als Naturwesen im normalen oder abnormalen Zustand erfassen (anthropologische und medicinische Statistik), sowie diejenigen, welche sich mit gewissen Massenerscheinungen des intellectuellen und moralischen Lebens beschäftigen. Es bleibt dann als eigenthümliche Aufgabe der Bevölkerungsstatistik übrig die zahlenmässige Untersuchung der normalen Bedingungen des Bestandes und des Wechsels einer Bevölkerungsmasse. Eine solche Masse bestimmt und ändert sich dadurch, dass ihre Mitglieder geboren werden, altern, sich verheirathen, sich vermehren, verwittwen, wiederholt heirathen und sterben. Ausserdem sind wegen der Beschränkung der Untersuchung auf ein begrenztes Staatsgebiet die Zu- und Abzüge zu berücksichtigen.

Die Theorie der Bevölkerungsstatistik hat nun die Grundsätze aufzustellen, nach denen diese Massenerscheinungen, die Gesammtergebnisse von zahlreichen, gewissermassen molecularen Einzelprocessen, zu beobachten und wissenschaftlich zu bewältigen sind." (Lexis 1875, S. 2)

Was ist hier mit „wissenschaftlich bewältigen" gemeint? Es gibt zwei unterschiedliche Arten von Fragestellungen. Zunächst widmet Lexis einen Großteil seines Buches der Entwicklung eines Begriffsrahmens, um die Dynamik der Veränderungen einer „Bevölkerungsmasse" durch Ereignisse der angegebenen Art (Geburten, Heiraten, Sterbefälle usw.) deskriptiv darzustellen. Man kann insoweit von einer dynamisch konzipierten Sozialstatistik sprechen, wie sie auch in der neueren empirischen Sozialforschung in Form von „Lebensverlaufsanalysen" betrieben wird.[8]

[8] In der methodischen Begleitdiskussion findet auch immer noch eine Auseinandersetzung mit dem von Lexis entwickelten Begriffsrahmen statt; vgl. exemplarisch Mayer und Huinink, 1990.

2. Mit Details der Lexis'schen Konzeption einer dynamischen Sozialstatistik brauchen wir uns hier nicht zu beschäftigen. Wichtig ist, wie er den Zusammenhang zur Wahrscheinlichkeitstheorie herstellt. Nachdem Lexis seine der Intention nach deskriptiven Begriffsbildungen dargelegt hat, beginnt ein neuer Abschnitt mit folgenden Bemerkungen:

„In den vorhergehenden Abschnitten sind die Methoden entwickelt worden, nach denen Veränderungen einer bevölkerungsstatistischen Masse sei es theoretisch genau oder mit einer dem praktisch zugänglichen Material entsprechenden Näherung beobachtet werden können. Vorzugsweise haben wir die Mittel gesucht, um Reihen von Massen aufzustellen, von denen jede aus der nächstvorhergehenden dadurch hervorgegangen ist, dass eine Anzahl von Angehörigen der letzteren in einen bestimmten neuen Zustand getreten ist oder ein bestimmtes neues Merkmal erhalten hat.

Haben wir aber auch in dieser Weise das theoretische Schema mit empirischem Stoffe ausgefüllt, so ist dadurch doch zunächst nichts weiter erreicht, als eine systematische Gruppirung von Beobachtungsdaten, die zu einander in gewissen thatsächlichen Beziehungen stehen. Von einer Theorie der Erscheinungen im Sinne der physikalischen Wissenschaften sind wir auf diesem Standpunkte noch weit entfernt; denn eine solche würde verlangen, dass wir die beobachteten thatsächlichen Beziehungen in exacter Weise auf ihre Ursachen zurückführten und aus der Kenntnis der Ursachen feststellten, wie sich die analogen Erscheinungen in der Zukunft wiederholen würden." (Lexis 1875, S. 93f)

Theoriebildung beginnt also für Lexis mit der Frage nach Ursachen und mit der Absicht, solche Kenntnisse für Voraussagen nutzen zu können.

3. Aber wie kann der Sinn einer solchen Frage nach Ursachen genauer angegeben werden? Es treten sogleich Schwierigkeiten auf, die Lexis folgendermaßen erläutert:

„Nehmen wir den möglichst einfachen Fall, die Veränderungen, welche eine Masse Gleichaltriger durch die Sterblichkeit ihrer Mitglieder von Altersclasse zu Altersclasse erleidet. Die Ueberlebens- und Sterbeordnung derselben können wir genau oder näherungsweise feststellen, aber über das *System* von Ursachen, vermöge dessen aus L_a Lebenden in der $(a+1)$ten Altersclasse m_{a+1} sterben, sagen uns diese Zahlenreihen nichts. In jedem einzelnen Falle können wir zwar die Todesursache beobachten, aber diese Erkenntnis gewährt uns keinen Aufschluss über die Frage, weshalb *im Ganzen* m_{a+1} Personen gestorben und $L_a - m_{a+1}$ lebend geblieben sind. Sie dient höchstens dazu, um uns von der ausserordentlichen Mannichfaltigkeit der Bedingungen des Sterbens und Ueberlebens zu überzeugen und zu dem Geständnis zu nöthigen, dass wir nicht hoffen dürfen, die Gründe des Zusammentreffens und die Verbindung dieser Bedingungen zu durchdringen und zu analysiren." (Lexis 1875, S. 94)

Die Überlegungen verweisen auf ein Problem, das bereits in Abschnitt 2.3 angesprochen worden ist. Der Sozialstatistiker möchte herausfinden, wie die Elemente einer „Masse“ oder eines „Kollektivgegenstandes“ entstanden sind bzw. wie sie ihre jeweils bestimmten Eigenschaften bekommen haben; aber es gibt keinen begrifflich faßbaren Prozeß, auf den man sich empirisch beziehen könnte. Wollte man empirisch vorgehen, müßte man sich auf die Einzelfälle beziehen, die die Elemente der „statistischen Massen“ bilden. Aber dies widerspricht dem theoretischen Ansatz einer Statistik, die ihren Gegenstand durch Gesamtheiten (und für sie definierte Häufigkeitsverteilungen) definiert.[9]

3.2.2 Soziale Chancensysteme

1. In dieser Sackgasse angelangt, nimmt indessen der Gedankengang eine bemerkenswerte Wendung. Lexis' Ausführungen gehen so weiter:

„Kann aber die innere Natur des Verursachungssystems, welches eine gegebene Massenerscheinung der Sterblichkeit bedingt, nicht aufgedeckt und zergliedert werden, so wird die naturwissenschaftliche Aufgabe der Bevölkerungsstatistik auf die Untersuchung beschränkt, ob die wirkenden unbekannten Ursachensysteme einem regellosen Wechsel unterworfen sind, oder ob sie sich mit einer gewissen Constanz behaupten. Im letzteren Falle werden auch in den beobachteten Resultaten gewisse Regelmässigkeiten hervortreten, und umgekehrt kann man aus solchen beobachteten Regelmässigkeiten auf eine annähernde Constanz der Ursachensysteme zurücksschliessen.

Aus diesen Bemerkungen ergibt sich sofort, dass die naturwissenschaftliche Untersuchung der statistischen Massenerscheinungen nur die Aufgabe hat, die inneren Beziehungen derselben in der Form darzulegen, welche der Wahrscheinlichkeitsrechnung eigenthümlich sind.

Wir haben es zu thun mit einer gewissen Zahl von Fällen eines Anfangszustandes, aus denen eine gewisse Zahl von Fällen eines Endzustandes hervorgeht; von den zwischen Anfangs- und Endzustand liegenden physischen Processen wissen wir als Statistiker ebensowenig, wie von den complicirten Bewegungen eines rollenden Würfels, bei dem uns nur die Lage interessirt, in der er zur Ruhe kommt.“ (Lexis 1875, S. 94)

Ein Ausweg aus der Sackgasse soll also dadurch gefunden werden, daß man sich die statistischen Massenerscheinungen so vorstellt, als ob sie

[9] Gleich zu Beginn seines Buches (1875, S. 1) hat Lexis darauf hingewiesen: „Bei der Bildung von Massen für die statistische Beobachtung verschwindet das Individuum als solches, und es erscheint nur noch als eine Einheit in einer Zahl von gleichartigen Gliedern, die gewisse Merkmale gemein haben und von deren sonstigen individuellen Unterschieden abstrahirt wird.“

durch Zufallsgeneratoren hervorgebracht werden. In dieser allgemeinen Formulierung stammt die Idee zwar nicht von Lexis,[10] bemerkenswert ist jedoch, wie Lexis eine neue Rhetorik einführt, indem er versucht, die natürlichen und gesellschaftlichen Verhältnisse, unter denen Menschen leben, als Chancensysteme vorstellbar zu machen, und zwar so, daß es von der Beschaffenheit dieser Chancensysteme abhängt, wie sich die Lebensverläufe der Menschen entwickeln. Diese zunächst noch unbestimmte Vorstellung wird dann begrifflich dadurch fixiert, daß die Chancensysteme in Analogie zu Zufallsgeneratoren konzipiert werden, wie sie in der Wahrscheinlichkeitsrechnung behandelt werden.

2. In dem Beispiel, auf das sich Lexis bezieht (eine Sterbetafel), nimmt die Betrachtungsweise folgende Form an:

„Wir stellen L_a lebende a-jährige gleichsam ein Lebensjahr auf die Probe, indem wir jeden einer äusserst verwickelten Mannigfaltigkeit von lebensgefährdenden Ursachen oder Todeschancen aussetzen, und wir beobachten, wie viele in der gegebenen Altersstrecke von dem drohenden Schlage wirklich getroffen werden.“ (Lexis 1875, S. 94f)

Bemerkenswert ist nicht nur die Idee eines fiktiven Experiments.[11] Für unseren gegenwärtigen Gedankengang noch wichtiger ist die begriffliche

[10] Z.B. betonte schon Laplace (1814, S. 82) „die Vorteile der Wahrscheinlichkeitsanalyse beim Aufsuchen der Gesetze der Naturerscheinungen ..., deren Ursachen unbekannt oder zu kompliziert sind, um ihre Wirkungen der Berechnung unterwerfen zu lassen.“ Noch vor Quetelet hat auch bereits Laplace die Verwendung der Wahrscheinlichkeitsrechnung in den „moralischen Wissenschaften“ propagiert. Das eben angeführte Zitat geht so weiter: „Das ist fast bei allen Gegenständen der moralischen Wissenschaften der Fall. So viele unvorhergesehene, oder verborgene, oder nicht abzuschätzende Ursachen beeinflussen die menschlichen Institutionen, daß es unmöglich ist, ihre Ergebnisse a priori zu beurteilen. Die Reihe der von der Zeit heraufgeführten Ereignisse bringt diese Resultate zur Entwickelung und zeigt die Mittel an, um die schädlichen unter ihnen abzuwenden. Man hat in dieser Hinsicht oft weise Gesetze gemacht; aber da man versäumt hat, die Beweggründe dafür aufzubewahren, so sind mehrere wieder als nutzlos abgeschafft worden, und es mußten erst unangenehme Erfahrungen das Bedürfnis ihrer Wiederherstellung aufs neue fühlen lassen. Es ist daher sehr wichtig, daß in jedem Zweige der öffentlichen Verwaltung ein genaues Register der Wirkungen geführt werde, welche die verschiedenen angewandten Mittel hervorgebracht haben, und welche eben so viele von den Regierungen im großen gemachten Erfahrungen sind. Wenden wir die auf Beobachtung und Berechnung gegründete Methode, die uns in den Naturwissenschaften so gute Dienste geleistet hat, nun auch auf die politischen und moralischen Wissenschaften an.“

[11] Es sei erwähnt, daß eine solche fiktiv-experimentelle Betrachtungsweise auch in der heutigen empirischen Sozialforschung oft zur Motivation statistischer Modellbildung angeführt wird. Z.B. heißt es in einer Formulierung von N. Longford (1993, S. vii): „A general approach to the analysis of data from an observational study relies on matching the context of the study with a hypothetical experimental design, and then

Identifizierung von „lebensgefährlichen Ursachen“ mit „Todeschancen“. Es wird suggeriert, daß mit dem Chancenbegriff in ähnlicher Weise argumentiert werden kann wie mit einem Ursachenbegriff. Ob bzw. in welcher Weise das möglich ist, wird uns noch mehrfach beschäftigen. Hier soll zunächst betont werden, daß Lexis den Chancenbegriff gerade deshalb einführt, weil man über Ursachen keine Aussagen machen kann oder möchte.

3. Bemerkenswert ist weiterhin, daß vollständig offen bleibt, worauf sich der Chancenbegriff bezieht. Die einzige empirische Bezugnahme besteht in einem Verweis auf beobachtete Todesfälle. Aber was soll es heißen, daß sie sich gewissen „Todeschancen“ verdanken? Tatsächlich sind Lexis' Chancensysteme zunächst nur Deutungschema für empirisch beobachtbare Erscheinungen. Das zeigt sich auch darin, wie Lexis glaubt, zu Aussagen über Chancensysteme kommen zu können. Er erläutert dies mit folgender Überlegung:

„Nehmen wir nun zunächst an, dass sich alle Individuen der betrachteten gleichalterigen Masse dem System der Todeschancen gegenüber *gleichartig* verhalten, so wird das Verhältnis der beobachteten Zahl der Todesfälle zu der Zahl der exponirten Lebenden ein Maass für die effective Wirksamkeit der Todeschancen werden, und zwar wird dieses Maass um so genauer, je grösser die Zahl der gleichartigen Lebenden ist, die dem Chancensystem gegenüber gestellt worden sind. Wäre diese Zahl unendlich gross, so würde jenes Verhältnis die genaue *Sterbenswahrscheinlichkeit* jedes einzelnen der exponirten gleichartigen Lebenden im $(a+1)$ten Altersjahr darstellen.

Diese Sterbenswahrscheinlichkeit ist also der einzige Ausdruck, durch den wir das System der Todeschancen charakterisiren können, und wir betrachten das System als constant, so lange es dieselbe Sterbenswahrscheinlichkeit der Gleichartigen bedingt.

Auch wenn wir diese Sterbenswahrscheinlichkeit nicht genau bestimmen können – und die genaue Bestimmung würde ja unendlich viele Beobachtungen voraussetzen – so nehmen wir doch an, dass sie mit einem *bestimmten* Werth den Erscheinungen zu Grunde liegt.“ (Lexis 1875, S. 95)

Es ist ersichtlich, daß der Gedankengang zirkulär ist bzw. nur auf eine Deutung hinausläuft, für deren Berechtigung keine Begründung angegeben wird. Denn als empirische Tatsachen stehen nur jeweils beobachtete relative Häufigkeiten (für Sterbefälle) zur Verfügung. Daß diese relativen Häufigkeiten durch ihnen zugrundeliegende Wahrscheinlichkeiten – also

---

proceeding with the analysis as if the data were the outcome of this 'experiment'. This approach is referred to as *statistical modelling*.“

durch Zufallsgeneratoren bzw. Chancen – *bestimmt* werden, ist eine nur äußerlich angehängte Deutung.

4. Man könnte vielleicht einwenden, daß es keines Beweises bedarf, daß dem beobachtbaren Geschehen überhaupt irgendwelche Ursachen zugrunde liegen. Dennoch liefert Lexis' Ansatz nur ein Deutungsschema; denn über Ursachen werden explizit gar keine Aussagen (nicht einmal irgendwelche spezifischen Annahmen) gemacht. Damit begann ja der Lexis'sche Gedankengang: daß der Statistiker über Ursachen keine Aussagen machen kann (oder möchte), da ihn dies zwingen würde, sich mit einzelnen Fällen zu beschäftigen. Anstatt Aussagen über Ursachen zu machen, werden vielmehr hypothetische Chancensysteme gedanklich konstruiert, die eine scheinbar vergleichbare argumentative Aufgabe übernehmen sollen. Zwar scheint auch Lexis zu bemerken, daß ein fiktives Chancensystem nicht stellvertretend Einsichten in Verursachungszusammenhänge vermitteln kann. Aber er geht darüber mit folgender kurzen Anmerkung hinweg:

„Mit dem Worte 'Chance' bezeichnen wir im Folgenden eine Bedingung, die sich von einer Ursache im eigentlichen Sinne dadurch unterscheidet, dass sie nicht mit Gewissheit, sondern nur mit einer bestimmten Wahrscheinlichkeit eine Wirkung von bestimmter Art hervorbringt.“ (Lexis 1875, S. 95)

Diese Bemerkung ignoriert jedoch das eigentliche Problem: daß die von Lexis konzipierten Chancensysteme nur als Deutungsschema für ein jeweils beobachtetes Geschehen existieren, eine rein begriffliche Existenz haben und in ihrem logischen Status vollständig vergleichbar sind mit den Verteilungsgesetzen, durch die Fechner empirische Häufigkeitsverteilungen als vermeintlich „gesetzmäßig bestimmte“ gedeutet hat. Der Unterschied besteht nur darin, daß Lexis eine neue Rhetorik einführt, die es erlauben soll, von Chancen so ähnlich sprechen zu können, wie man in anderen Kontexten von Ursachen sprechen kann.

3.2.3 Aufgaben der Sozialstatistik

1. Es ist diese neue Rhetorik der „sozialen Chancensysteme“, die den Leitfaden für Lexis' Konzeption einer theoretischen Sozialstatistik liefert. Folgende zusammenfassende Darstellung macht das noch einmal sichtbar:

„Die unbekannte Verbindung zwischen dem Anfange und dem Endzustande der Elementarmasse, d.h. in unserem Beispiele zwischen der Zahl der im Ganzen beobachteten und der gestorbenen Individuen einer Elementargruppe in einer bestimmten Altersclasse denken wir uns also in der Form eines Chancensystems. Bleibt dasselbe in mehreren gleichartigen Beobachtungsreihen unverändert, so zeigt die Wahrscheinlichkeitsrechnung, dass auch die beobachteten Verhältnisse der Gestorbenen zu den Lebenden nur in gewissen, mit genügender Sicherheit bestimmbaren Grenzen von einander abweichen werden; und umgekehrt kann man, wenn zwei beobachtete Verhältnisse über gewisse angebbare Grenzen hinaus von einander abweichen, praktisch mit Gewissheit aussagen, dass sich das Chancensystem in beiden Reihen geändert hat.

Demnach fasst sich die Aufgabe der socialphysiologischen Statistik in Folgendem zusammen: sie hat möglichst individualisirte Elementarmassen zu bilden und durch Wahrscheinlichkeitsverhältnisse die Chancensysteme zu charakterisiren, welche die bedeutsamen Veränderungen derselben bedingen; sie hat ferner zu untersuchen, wiefern durch die Verschiedenheit der Unterscheidungsmerkmale der Elementarmassen Verschiedenheiten ihrer Chancensysteme entstehen, und endlich festzustellen, ob die einzelnen Chancensysteme im Laufe der Zeit annähernd constant bleiben oder sich in einer bestimmbaren Weise ändern." (Lexis 1875, S. 121)

2. In diesem Zitat werden auch drei zentrale Aufgaben sichtbar, die Lexis für die theoretische Sozialstatistik formuliert:

a) Zunächst soll sie empirisch ermittelbare relative Häufigkeiten (für soziale Ereignisse und Sachverhalte) als Ausdruck zugrunde liegender Chancensysteme deuten, wodurch es möglich werden soll, einen Wahrscheinlichkeitsbegriff zu verwenden, um Chancensysteme zu charakterisieren.

b) Sodann soll sie sich bemühen, die Mitglieder einer „statistischen Masse" nach möglichst vielen Merkmalen zu differenzieren, um auf diese Weise möglichst „gleichartige Teilmassen" zu finden, denen jeweils ein spezifisches Chancensystem unterstellt werden kann.

c) Schließlich soll sie sich bemühen, etwas darüber herauszufinden, ob und wie sich die Chancensysteme im Zeitablauf verändern.

3. Bei Lexis stehen die ersten beiden Aufgaben im Dienst der dritten, an der er in erster Linie interessiert war. Im Vergleich zum seinerzeit verbreiteten Reden von „sozialen Gesetzmäßigkeiten" war seine Position eher vorsichtig:

„Anders aber bei den menschlichen Massenveränderungen, die wir uns auf einem Chancensysteme beruhend denken. Auch hier können wir im Allgemeinen beurtheilen, welche Momente die Grösse der Veränderung wesentlich mit bestimmen, aber wir erkennen zugleich unmittelbar und unzweifelhaft, dass viele dieser Momente, weil nicht durch physische, sondern durch menschlich-sociale Verhältnisse bedingt, ihrem Wesen nach veränderlich sind. Der Schluss von der beobachteten auf die zukünftige Constanz des Chancensystems wird daher unter Umständen ein höchst problematischer, und die erstere begründet keineswegs die Annahme einer eigentlichen Gesetzmässigkeit der betreffenden Massenänderungen im naturwissenschaftlichen Sinne." (Lexis 1875, S. 122f)

Hier interessiert uns jedoch nicht das damals intensiv diskutierte Problem der zeitlichen Beständigkeit sozialstatistischer Häufigkeitsverteilungen, sondern die neue Rhetorik: Lexis' Vorschlag, die Sozialstatistik dadurch für Begriffsbildungen der Wahrscheinlichkeitstheorie anschlußfähig zu machen, daß gesellschaftliche Verhältnisse als Chancensysteme gedeutet werden. Zwar hat Lexis vorgeschlagen, beide Fragen aufeinander zu beziehen, z.B. in folgender Formulierung:

„Die statistischen Zahlenverhältnisse drücken also nicht die herrschende Norm, das Gesetz der Erscheinungen aus, sondern sie geben uns nur einen näheren, exacten Aufschluss über die wirkliche Constitution der Erscheinungen, nämlich über das Chancensystem, welches denselben zugrunde liegt; und wegen der überall erkennbaren und begreiflichen relativen Stabilität der Wahrscheinlichkeiten in den menschlichen Dingen dürfen wir annehmen, dass dieses Chancensystem sich nur langsam ändern, also auch wenigsten in der nächsten Zukunft noch ähnliche Zahlenverhältnisse hervorbringen wird." (Lexis 1875, S. 124)

Aber gleichgültig, wie es sich mit der zeitlichen Beständigkeit verhält, bemerkenswert – und fragwürdig – erscheint mir vor allem der theoretische Anspruch: daß man durch eine Deutung gesellschaftlicher Verhältnisse als Chancensysteme „Aufschluss über die wirkliche Constitution der Erscheinungen" gewinnen kann.

4. Gerade wenn man Lexis' Gedankengang ernst nimmt, wird fragwürdig, welcher Erkenntnisgewinn sich daraus ergeben könnte, daß man die sich historisch wandelnden Erscheinungen des gesellschaftlichen Lebens als Ausdruck fiktiver Chancensysteme deutet. Denn zunächst hat Lexis ja sehr deutlich betont, daß die Sozialstatistik keine Einsichten liefern kann, wodurch – durch welche Ursachen – soziale Sachverhalte und Ereignisse zustande kommen. Man denke exemplarisch an v. Bortkiewicz' „Gesetz der kleinen Zahlen" für die Selbstmorde von Kindern. Insofern wird aber auch der theoretische Anspruch, daß man etwas über „die wirkliche Constitution der Erscheinungen" lernen könne, fragwürdig.

5. Aber diese Frage wird von Lexis nicht ernsthaft verfolgt, sondern durch eine metaphorische Parallelisierung von „Ursachen" und „Chancen" ausgeblendet. Stattdessen gerät eine andere Frage in den Mittelpunkt: wie man mit den Mitteln der Sozialstatistik zu einer immer differenzierteren Ermittlung von Chancensystemen kommen kann. Lexis erläutert den Gedankengang im Anschluß an das Beispiel, das bereits im vorangegangenen Abschnitt angeführt worden ist.

„Wir haben oben angenommen, dass die exponirten Lebenden sich dem System der Todeschancen gegenüber gleichartig verhalten, d.h., dass alle Individuen a priori mit gleicher Leichtigkeit unter dem Einfluss jeder tödtlich wirkenden Ursachencombination gerathen können.

Unter dieser Voraussetzung müsste man nicht nur, wenn man die ganze Gesammtheit rein *zufällig* in Unterabtheilungen zerlegte, sondern auch bei einer planmässigen Gruppenbildung nach bestimmten Merkmalen für jede Unterabtheilung ungefähr denselben Näherungswerth der Sterbenswahrscheinlichkeit finden, und die Abweichungen dieser Näherungswerthe unter sich dürften gewisse theoretisch bestimmbare Grössen nicht übersteigen.

Wenn man aber die a-jährige Bevölkerung eines Landes nach Körperconstitution, Lebensweise, ökonomischer Lage u.s.w. in Gruppen zerlegte, so unterliegt es keinem Zweifel, dass die einzelnen Gruppen erheblich von einander abweichende Sterbenswahrscheinlichkeiten aufweisen würden. Die Individuen sind also dem vorausgesetzten System der Todeschancen gegenüber *nicht* als sämmtlich gleichartig anzusehen, oder, was auf dasselbe hinausläuft, es ist eine unbekannte Anzahl von Gruppen vorauszusetzen, welche *verschiedenen* Systemen von Todeschancen unterworfen sind.“ (Lexis 1875, S. 96)

Hier deutet sich ein weiterer Gedanke an, der für die Entwicklung der Sozialstatistik von zentraler Bedeutung geworden ist: daß man versuchen kann, Menschen aufgrund beobachtbarer Merkmale (sei es der Menschen selbst, sei es der Situationen, in denen sie sich befinden) in Gruppen einzuteilen und dann für jede dieser Gruppen ein spezifisches Chancensystem zu unterstellen. Der Gedanke liegt bereits der Konstruktion einer Sterbetafel zugrunde, die je nach dem Alter der Menschen unterschiedliche Sterbewahrscheinlichkeiten annimmt. Die Überlegung kann indessen fast beliebig fortgesetzt werden. Zum Beispiel kann man in jeder Altersgruppe zusätzlich zwischen Männern und Frauen unterscheiden und für sie unterschiedliche Sterbewahrscheinlichkeiten annehmen, und man kann weiterhin Menschen nach ihren Berufen unterscheiden usw.[12]

[12] So auch schon bei Laplace (1814, S. 111): „So viele veränderliche Ursachen üben ihren Einfluß auf die Sterblichkeit aus, daß die Tabellen, die sie darstellen, nach Ort und Zeit sich ändern müssen. Die verschiedenen Lebensumstände bieten in dieser Hinsicht merkliche Unterschiede dar, die von den mit jedem Stande untrennbaren Beschwerden und Gefahren abhängen und welche bei den auf die Lebensdauer begründeten Rechnungen nicht außer Acht gelassen werden dürfen. Aber diese Unterschiede sind noch nicht genugsam beachtet worden. Eines Tages werden sie es sein; dann wird man wissen, welche Einbuße an Leben jeder Beruf fordert, und man wird aus diesen Kenntnissen Gewinn ziehen, um die Gefahren zu vermindern.“

3.3 Wahrscheinlichkeit als Deutungsschema

1. Mit Lexis’ neuer Rhetorik, daß man von Chancen so ähnlich sprechen könne, wie von Ursachen, werden wir uns in Kapitel 6 näher beschäftigen. Hier soll zum Abschluß nur betont werden, daß Lexis auch für das Problem, wie in der Sozialstatistik sinnvoll von Wahrscheinlichkeiten gesprochen werden kann, keine Lösung anzubieten vermag. Chancensysteme werden nur als Deutungsschema für empirisch realisierte Häufigkeitsverteilungen eingeführt, erfahren jedoch keine selbständige empirische Explikation. Sofern dieser Eindruck entstehen kann, beruht er nur auf einer irreführenden Äquivokation mit Handlungschancen.

2. Lexis hat auch in späteren Arbeiten deutlich gemacht, daß es sich nur um ein Deutungsschema handelt, z.B. in seinen „Abhandlungen zur Theorie der Bevölkerungs- und Moralstatistik“ (1903). In einer dieser Abhandlungen bespricht er zunächst den Begriff einer „Wahrscheinlichkeit a priori“ und bezieht sich dafür – wie in der Wahrscheinlichkeitsrechnung üblich – auf Zufallsgeneratoren (Glücksspiele). Dann gehen seine Ausführungen so weiter:

„Man kann nun auch umgekehrt die Hypothese aufstellen, dass ein Ereignis, das in einer grossen Zahl s von Versuchen p-mal aufgetreten ist, die Wahrscheinlichkeit p/s habe, dass also z.B., wenn bei 600 Zügen aus einer Urne (wobei die gezogene Kugel stets wieder in die Urne zurückzulegen war) 105 mal eine weisse und 495 mal eine schwarze Kugel gezogen worden ist, das Verhältnis der weissen zu den schwarzen Kugeln in der Urne ungefähr 105 : 495 betrage, oder 1 : 5, wenn man wüsste, dass überhaupt nur 6 Kugeln in der Urne seien. Man kann auch noch weiter gehen und solche Hypothesen aufstellen für Ereignisse, auf die das Bild eines Glücksspiels mit Würfeln oder Urnen gar nicht mehr angewandt werden kann. Wenn überhaupt ein Ereignis unter gewissen festgesetzten Bedingungen p mal eingetreten ist, das, soviel wir die Sache beurteilen können, an jedem von s beobachteten Einzelfällen hätte vorkommen können, so nehmen wir hypothetisch an, dass m/s näherungsweise die Wahrscheinlichkeit dieses Ereignisses darstelle. Ob diese Annahme richtig sei, ob sie überhaupt einen Sinn habe, kann erst durch die Untersuchung mehrerer gleichartiger Beobachtungsreihen entschieden werden. Durchaus irrig ist die Meinung, dass überhaupt jedes aus grossen Zahlen gebildete Verhältnis sich wie der Näherungswert einer festen Wahrscheinlichkeit verhalte und sich daher mit steigender Beobachtungszahl einem konstanten Werte nähere. Es ist vielmehr leicht möglich, dass bei fortschreitender Vergrösserung von s durch

Hinzufügung neuer Beobachtungen das Verhältnis m/s fortschreitend grösser oder kleiner wird.“ (Lexis 1903, S. 44f)

Damit ist man indessen erneut bei den Überlegungen, die schon in Abschnitt 2.4 besprochen worden sind. Für die hypothetische Unterstellung eines Zufallsgenerators wird kein Grund angegeben. Es wird auch nicht versucht, diese Unterstellung durch Hinweise auf Eigenschaften der Prozesse, die die fraglichen Sachverhalte oder Ereignisse hervorbringen, sinnvoll erscheinen zu lassen. Anstelle dessen wird ein scheinbar empirisches Argument angeboten: daß die Deutung einer empirischen Häufigkeitsverteilung durch ein Verteilungsgesetz dadurch begründbar sei, daß bei einer zeitlichen Reihe vergleichbarer Häufigkeitsverteilungen nur Unterschiede auftreten, die als „zufällig“ gedeutet werden können. Das Argument verfehlt indessen die Fragestellung, wie schon in Abschnitt 2.4 ausgeführt worden ist.

3. Möglicherweise entgegen den Intentionen von Lexis liefert seine Überlegung eine scheinbare Begründung dafür, daß für jede beliebige empirische Häufigkeitsverteilung ein hypothetisches Verteilungsgesetz angenommen werden kann. Tatsächlich hat sich in der Statistik ein Sprachgebrauch entwickelt, der Wahrscheinlichkeiten im wesentlichen mit relativen Häufigkeiten identifiziert. Z.B. heißt es in einem Buch zur theoretischen Statistik von H. Forcher (1913):

„Durch die intensiven statistischen Maßzahlen wird der Einfluß der einzelnen Ursachen auf die Erscheinungen gemessen. Die bekannteste intensive Maßzahl ist die *statistische Wahrscheinlichkeit*. Am schärfsten hat den Begriff meines Erachtens Blaschke in seinem tiefgründigen Werke: ‘Vorlesungen über mathematische Statistik’ herausgearbeitet, wenn er sagt: ‘Die statistische Wahrscheinlichkeit ist ein Bruch, in dessen Nenner die Anzahl der Individuen einer bestimmten Art (des Anfangszustandes), in dessen Zähler die Summe der daraus binnen der Zeiteinheit unter dem Einfluß der zu messenden Ursache hervorgegangene Änderungen (des Endzustandes) sich befindet.’ Sie heißen intensive Maßzahlen, weil sie die Häufigkeit eines Zustandes oder einer Zustandsveränderung aufdecken.“ (Forcher 1913, S. 203)

Hier deutet sich auch an, wie mit fiktiven Wahrscheinlichkeiten der Anschein einer kausalen Argumentation erweckt werden soll: Wahrscheinlichkeiten werden als Meßgrößen für unbekannte Ursachen hingestellt. Es handelt sich indessen nur um eine rhetorische Variante der Lexis’schen Ausdrucksweise, daß man Wahrscheinlichkeitsgrößen als Chancen und diese wiederum als dem Geschehen zugrundeliegende Ursachen auffassen könne.

4. Hier noch ein weiteres Beispiel aus dem „Grundriß der Statistik“ von Franz Zizek (1921):

„Unter den Verhältniszahlen kommt eine besondere Bedeutung jenen zu, welche als Wahrscheinlichkeitsgröße angesehen werden können. Eine statistische Wahrscheinlichkeitsgröße ist nach Lexis ein Bruch, dessen Zähler eine Anzahl beobachteter besonderer Fälle oder Elemente angibt, die aus der im Nenner angegebenen Anzahl beobachteter Fälle oder Elemente entweder hervorgegangen sind oder einen Teil dieses letzteren bilden. Die Sterbenswahrscheinlichkeit z.B. ist das Verhältnis der Gestorbenen einer bestimmten Altersklasse zur Gesamtheit der Lebenden, die die untere Grenze dieser Altersklasse erreicht haben und unter das betreffende Sterberisiko getreten sind; das Verhältnis der Zahl der Knabengeburten zu der Gesamtzahl der Geburten stellt die Wahrscheinlichkeit einer Knabengeburt dar.

Die *Bedeutung* der Verhältniszahlen, welche in formaler Hinsicht als *Wahrscheinlichkeitsgrößen* (oder Funktionen von solchen) auftreten, besteht darin, daß sie die Behandlung mittelst verschiedener, der Wahrscheinlichkeitstheorie angehörenden Methoden zulassen, während diese Methoden auf Verhältniszahlen anderer Art nicht angewendet werden können. Die Wahrscheinlichkeitstheorie erblickt in Verhältniszahlen, die als Wahrscheinlichkeitsgrößen angesehen werden können, sozusagen empirische Näherungswerte von ihnen zugrunde liegenden theoretischen Wahrscheinlichkeiten, sie unternimmt es, die Zuverlässigkeit (Genauigkeit) solcher Wahrscheinlichkeitsgrößen zu berechnen, d.h. festzustellen, innerhalb welcher Grenzen die dem empirischen Werte zugrunde liegende theoretische Wahrscheinlichkeit mit gegebener Wahrscheinlichkeit liegen dürfte, die Differenz zwischen zwei empirischen Wahrscheinlichkeitsgrößen (z.B. der Sterbenswahrscheinlichkeit zweier Bezirke) wird nach der Richtung untersucht, ob sie noch mit einer gewissen Wahrscheinlichkeit dem Zufalle zugeschrieben werden könne oder ob anzunehmen sei, daß den beiden empirischen Bestimmungen ungleiche theoretische Wahrscheinlichkeiten zugrunde liegen.“ (Zizek 1921, S. 131f)

Die Ausführungen machen deutlich, daß der Wahrscheinlichkeitsbegriff nur noch als ein Deutungsschema für empirische Häufigkeiten verwendet wird; und auch nur deshalb, um Begriffsbildungen und Vorstellungen aus der Wahrscheinlichkeitstheorie übertragen zu können. Der Anspruch, daß das Deutungsschema zu einer Erklärung sozialer Sachverhalte und Ereignisse führen soll, erscheint nur noch indirekt in Formulierungen wie „zugrunde liegen“.

5. Eine rein metaphysische Deutung erhält dieses „zugrunde liegen“ schließlich in den sog. frequentistischen Definitionsversuchen des Wahrscheinlichkeitsbegriffs, bei denen die von Lexis noch empirisch gemeinte Forderung nach einer näherungsweisen zeitlichen Konstanz relativer

Häufigkeiten in die Definition des Wahrscheinlichkeitsbegriffs eingebaut wird. Exemplarisch kann man sich das anhand folgender Ausführungen von Ernst Nagel (1939) verdeutlichen:

“The basic ideas of the frequency conception of probability emerge upon an examination of such a statement as ‘The probability that a person of thirty residing in the United States survives his thirty-first birthday is .945.’ The meaning of such a statement can be ascertained by examining how it is established. That procedure, greatly simplified, is somewhat as follows. Suppose that during a period of years there is no migration to or from the United States, and that during these years exact counts are made of its inhabitants who fall into definite age groups. Thus, suppose that in 1900 there are 2,000,000 persons who have just reached their thirtieth birthday, and that exactly one year later there are 1,890,000 persons who have just reached their thirty-first birthday; that is, of the thirty-year-olds in 1900 a ratio of .9450 survive at least another year. We imagine that similar figures are obtained for the four succeeding years, and that the ratios of thirty-year-olds who survive their thirty-first birthday are .9452, .9456, .9451, and .9454, respectively. We notice that, although these ratios are not constant, the differences do not appear until the fourth decimal is reached. We may say, therefore, that during these five years approximately 945 out of a thousand thirty-year-old residents of the United States live for at least another year; and we may make the further *assumption* that for an *indefinite number of future years* the corresponding ratios of survivals remain in the neighborhood of .945. Accordingly, the statement ‘The probability that a thirty-year-old resident of the United States survives his thirty-first birthday is .945’ means that in the long run the relative frequency with which thirty-year-olds in the United States survive for at least one year is approximately .945.” (Nagel 1939, S. 19)

Soweit es sich in diesen Ausführungen um empirische Aussagen handelt, wird offenbar ein Wahrscheinlichkeitsbegriff gar nicht benötigt, sondern man kommt vollständig mit relativen Häufigkeiten aus. Auch um die in vier Jahren beobachteten relativen Häufigkeiten als näherungsweise gleich oder als kleine Abweichungen von ihrem Durchschnittswert zu betrachten, ist ein Wahrscheinlichkeitsbegriff nicht erforderlich. Ein solcher Begriff wird hier von Nagel vielmehr nur deshalb verwendet, um eine metaphysische und in ihrer Begründbarkeit gar nicht reflektierte Hypothese einzuführen: daß es auch in aller Zukunft so bleiben wird, wie man in den vier Jahren beobachtet hat.

Kapitel 4

Reale und fiktive Gesamtheiten

In den beiden vorangegangenen Kapiteln wurde besprochen, wie der Ansatz der theoretischen Sozialstatistik auf der Vorstellung beruht, daß soziale Sachverhalte und Ereignisse durch hypothetische Zufallsgeneratoren hervorgebracht werden. In der weiteren Entwicklung der Sozialstatistik im 20. Jahrhundert ist diese Vorstellung zunehmend mit einer anderen vermischt worden: daß man sich die jeweils verfügbaren Daten über soziale Sachverhalte und Ereignisse als zufällig zustandegekommene Stichproben aus „Gesamtheiten höherer Ordnung“ (Anderson) vorstellen könne. Das soll in diesem Kapitel etwas genauer besprochen werden.

4.1 Einleitende Bemerkungen

1. Die meisten neueren Lehrbücher zur statistischen Methodenlehre orientieren sich an der Vorstellung, daß sich statistische Daten auf Stichproben aus ihnen zurechenbaren Gesamtheiten beziehen. Eine typische Formulierung ist zum Beispiel:

„Unter einer *Stichprobe* versteht man eine Teilmenge von Einheiten, die aus einer bestimmten Masse ausgewählt wurden, mit dem Ziel, daraus Schlüsse auf die Beschaffenheit dieser Masse zu ziehen. Die Masse selbst, die durch die Stichprobe untersucht werden soll, wird in der Stichprobentheorie als *Gesamtheit* bezeichnet.“ (Pfanzagl 1972, S. 139f)

Weiterhin wird in der Methodenliteratur betont, daß man mit den Daten einer Stichprobe nur dann begründete Aussagen über eine korrespondierende Gesamtheit machen könne, wenn es sich um eine zufällig ausgewählte Stichprobe handelt; zum Beispiel:

„... müssen wir den Begriff der *Stichprobe aus einer Grundgesamtheit* präzisieren. Bisher genügte es zu wissen, daß es sich dabei um eine Auswahl von n Dingen aus einer gegebenen Gesamtheit G handelt oder, da das beobachtete Merkmal und weniger das Ding als solches interessiert, um n Zahlen (bzw. Zahlenpaare bei gleichzeitiger Beobachtung zweier Merkmale usw.).

Die Stichprobe soll Aufschluß über die genannte Gesamtheit G geben, die man aus finanziellen, zeitlichen oder prinzipiellen Gründen nicht als Ganzes untersuchen kann. Damit sich dieser Aufschluß gewinnen läßt, muß die Stichprobe eine *Zufallsauswahl* darstellen. Dies bedeutet, sie muß so entnommen werden, daß jedes Ding in G eine ganz bestimmte angebbare Wahrscheinlichkeit hat, gezogen, d.h. der Stichprobe einverleibt zu werden.“ (Kreyszig 1968, S. 163)

Die Vorstellung ist also folgende: Man hat es mit einer bestimmten Gesamtheit von Dingen, Situationen oder Ereignissen zu tun und kann daraus mithilfe eines angebbaren Verfahrens auf zufällige Weise eine Teilmenge auswählen; insbesondere muß jedes Element der Gesamtheit ein Mitglied der Stichprobe werden können.

2. Diese Vorstellung hat zwei wichtige Implikationen. Erstens muß es sich um eine in unserer Erfahrungswelt real existierende Gesamtheit handeln, denn nur dann kann man einen Teil ihrer Mitglieder mithilfe eines angebbaren Verfahrens auswählen.[1] Dinge oder Situationen, die es heute nicht mehr oder noch nicht gibt, können nicht Mitglieder einer Stichprobe werden, die durch ein heute anwendbares Verfahren definiert werden kann. Zweitens impliziert der Gedankengang, daß Zufälligkeiten nur aus dem Auswahlverfahren resultieren können. Der Gedanke, daß die Gesamtheit aus Elementen besteht, deren Eigenschaften in irgendeinem Sinne „zufällig variieren“, kommt von vornherein nicht in Betracht. Die Elemente der Gesamtheit werden vielmehr mit jeweils bestimmten Eigenschaften vorausgesetzt.

3. Man kann sich den Sachverhalt anhand eines Urnenmodells verdeutlichen, das auch in der Literatur oft zur Erklärung verwendet wird.[2] Man stellt sich vor, daß jedes Mitglied der Grundgesamtheit durch eine Kugel in einer Urne repräsentiert wird und daß jede Kugel den Namen und die Eigenschaften des von ihr repräsentierten Objekts gewissermaßen als Inschrift hat. Das Auswahlverfahren besteht dann darin, einen Teil der Kugeln aus der Urne zufällig herauszuziehen. Das Modell macht deutlich, in welcher Weise man sich die Gesamtheit als eine *für das Auswahlverfahren vorausgesetzte Gegebenheit* vorzustellen hat. Offenbar können nur Kugeln aus der Urne gezogen werden, die in ihr enthalten sind; und es ist auch klar, daß das Auswahlverfahren nur einen Einfluß darauf hat, welche Kugeln ausgewählt werden, aber nicht auf ihre Eigenschaften.[3]

4. Das Urnenmodell macht weiterhin deutlich, wie man sich Auswahlverfahren, durch die Zufallsstichproben definiert werden, als Zufallsgeneratoren vorstellen kann. Es handelt sich um Auswahlgeneratoren. Solche Auswahlgeneratoren erzeugen nicht die Dinge oder Eigenschaften, die man beobachten möchte, sondern sie wählen aus einer vorgegebenen Menge von Dingen einen Teil aus. Die Auswahlgeneratoren, durch die Zufallsstichproben zustande kommen, müssen also begrifflich strikt von den hypothetischen Zufallsgeneratoren unterschieden werden, die in der theoretischen Sozialstatistik als Modelle für Prozesse dienen sollen, durch die Dinge mit jeweils bestimmten Eigenschaften entstehen. Dem entsprechen auch zwei vollständig unterschiedliche Fragestellungen. In der theoretischen Sozialstatistik, so wie wir sie bisher besprochen haben, möchte man Einsichten in Prozesse gewinnen, durch die soziale Sachverhalte und Ereignisse entstehen. Der theoretische Ansatz besteht darin, diese Prozesse als Zufallsgeneratoren zu konzipieren und dann zu versuchen, etwas über die Beschaffenheit dieser Zufallsgeneratoren herauszufinden. Auswahlgeneratoren sind demgegenüber nur ein methodisches Hilfsmittel, um zu – ihrer Intention nach – deskriptiven Aussagen über Merkmalsverteilungen in Gesamtheiten zu gelangen, die als in der Realität gegeben dem Auswahlverfahren vorausgesetzt werden. Es wäre in diesem Fall auch ganz sinnlos, das Ziel darin sehen zu wollen, Einsichten in den Zufallsgenerator zu gewinnen. Denn der Zufallsgenerator ist in der Stichprobentheorie ein Auswahlverfahren, das sich der Statistiker selbst gemacht hat und daß er infolgedessen von vornherein kennt; und auch kennen muß, um hypothetische Aussagen über Merkmalsverteilungen in den Gesamtheiten, auf die sich die ausgewählten Stichproben beziehen,

[1] Im ursprünglichen Verständnis von Stichproben wird das gut deutlich: „Das Wort Stichprobe geht auf den Vorgang der Warenprüfung zurück. So wird man z.B. bei der Prüfung eines Ballens Baumwolle an verschiedenen Stellen hineinstechen und je eine Probe entnehmen. Ein anderes Beispiel: Kommt bei einem Händler eine Sendung von 1000 Apfelsinen an, so kontrolliert er vielleicht nur 40 Stück genau. Sind diese alle einwandfrei, so wird er wahrscheinlich annehmen, daß *fast* alle 1000 Apfelsinen von guter Qualität sind.“ (Kellerer 1960, S. 110)

[2] Kellerer (1960, S. 117) bemerkt: „Die Stichprobentheorie bedient sich der Wahrscheinlichkeitsrechnung, die ihre Entstehung der mathematischen Behandlung von Glücksspielen verdankt. So ist es verständlich, wenn in unseren theoretischen Ausführungen Ausspielungen von Münzen oder Ziehungen von gleich großen Kugeln aus einer Urne eine Rolle spielen. Derartige ‘Modelle’ haben den Vorteil, daß sie den theoretischen Sachverhalt ohne Rücksicht auf das spezielle Anwendungsgebiet wiedergeben.“ Billeter (1970, S. 14f) geht gleichsam selbstverständlich von folgender Überlegung aus: „Es sei angenommen, daß alle Elemente der Grundgesamtheit als Kugeln in einer Urne dargestellt sind. Bei der Stichprobenentnahme geht es also darum, zufällig eine gewisse Anzahl Kugeln herauszugreifen. Dies geschieht dadurch, daß blindlings in die Urne gegriffen wird und so viele Kugeln herausgeholt werden, als Elemente in die Stichprobe einbezogen sein müssen.“

[3] Ein davon zu unterscheidendes Problem entsteht daraus, daß bei einer nachfolgenden Datenermittlung Fehler – u.U. auch „Meßfehler“, wenn sinnvoll von Messen gesprochen werden kann – auftreten können. Aber das hat offenbar nichts damit zu tun, welche Elemente aus einer Gesamtheit für eine nachfolgende Datenermittlung ausgewählt werden.

einschätzbar zu machen.

5. Man kann es auch so sagen: Auswahlverfahren beziehen sich auf „*daten*erzeugende Prozesse", so wie dies Wort in Abschnitt 1.2 erläutert worden ist; dagegen beziehen sich die hypothetischen Zufallsgeneratoren der theoretischen Sozialstatistik ihrer Intention nach auf die Prozesse, durch die soziale Sachverhalte und Ereignisse hervorgebracht werden. Die Unterscheidung ist offenbar wichtig, wenn man ein Verständnis theoretischer Sozialstatistik gewinnen möchte. Leider hat jedoch die Verbreitung des Denkschemas „Stichprobe–Grundgesamtheit" in der Sozialstatistik dazu geführt, daß die Unterscheidung zwischen Auswahlverfahren und hypothetischen Zufallsgeneratoren zunehmend unkenntlich geworden ist. Die begrifflichen Unklarheiten verdanken sich im wesentlichen dem irreführenden Gedanken, daß man sich auch die hypothetischen Zufallsgeneratoren, die sich ihrer Intention nach auf soziale Prozesse beziehen, als Auswahlverfahren vorstellen könne. In den folgenden Abschnitten soll diesen Unklarheiten etwas genauer nachgegangen werden.

4.2 Der Ansatz R. A. Fishers

Für die Verbreitung der Vorstellung, daß man sich auch Zufallsgeneratoren, die sich ihrer Intention nach auf Prozesse beziehen, als Auswahlverfahren aus vorgängig existierenden Gesamtheiten vorstellen könne, hat insbesondere der Statistiker R. A. Fisher (1890–1962) eine wichtige Rolle gespielt.[4] Das soll in diesem Abschnitt besprochen werden.

4.2.1 Hypothetische Populationen

1. In einer einflußreich gewordenen Abhandlung „On the Mathematical Foundations of Theoretical Statistics" (1922) beginnt Fisher mit folgender Klage über die unbefriedigenden theoretischen Grundlagen der Statistik:

"Several reasons have contributed to the prolonged neglect into which the study of statistics, in its theoretical aspects, have fallen. In spite of the immense amount of fruitful labour which has been expended in its practical applications, the basic principles of this organ of science are still in a state of obscurity, and it cannot be denied that, during the recent rapid development of practical methods, fundamental problems have been ignored and fundamental paradoxes left unresolved." (Fisher 1922, S. 310)

[4]Informationen über Fishers Biographie und Arbeit finden sich bei Johnson und Kotz 1997, S. 99ff.

Dann folgt ein Abschnitt über „the purpose of statistical method", der mit folgenden Worten beginnt:

"In order to arrive at a distinct formulation of statistical problems, it is necessary to define the task which the statistician sets himself: briefly, and in its most concrete form, the object of statistical methods is the reduction of data. A quantity of data, which usually by its mere bulk is incapable of entering the mind, is to be replaced by relatively few quantities which shall adequately represent the whole, or which, in other words, shall contain as much as possible, ideally the whole, of the relevant information contained in the original data." (Fisher 1922, S. 311)

Die Formulierung klingt vertraut: der Statistiker hat es mit einer oft großen Menge von Daten zu tun und möchte ihren „Informationsgehalt" möglichst kurz und übersichtlich präsentieren, z.B. in Gestalt von Häufigkeitsverteilungen oder sie charakterisierenden Maßzahlen.[5]

2. Aus dieser Aufgabenbestimmung ergibt sich indessen kein Gesichtspunkt für eine theoretische Statistik. Es ist deshalb bemerkenswert, wie die Überlegungen von Fisher weiter gehen. Im Anschluß an das eben angeführte Zitat heißt es:

"This object is accomplished by constructing a hypothetical infinite population, of which the actual data are regarded as constituting a random sample. The law of distribution of this hypothetical population is specified by relatively few parameters, which are sufficient to describe it exhaustively in respect of all qualities under discussion. Any information given by the sample, which is of use in estimating the values of these parameters, is relevant information. Since the number of independent facts supplied in the data is usually far greater than the number of facts sought, much of the information supplied by any actual sample is irrelevant. It is the object of the statistical processes employed in the reduction of data to exclude this irrelevant information, and to isolate the whole of the relevant information contained in the data." (Fisher 1922, S. 311f)

Die Idee, die Fishers Ansatz einer theoretischen Statistik begründen soll, besteht also darin: daß man sich die jeweils verfügbaren Daten als eine Zufallsstichprobe aus einer Grundgesamtheit vorstellen soll.

[5]Eine der Fisher'schen ganz ähnliche Formulierung findet sich bereits bei F. Galton (1908, S. 33): „The object of statistical science is to discover methods of condensing information concerning large groups of allied facts into brief and compendious expressions suitable for discussion."

3. Um Fishers Idee zu verstehen, ist es zunächst wichtig, sie von der Konzeption eines sozialstatistischen Inferenzproblems abzugrenzen, bei dem man es mit einer Stichprobe aus einer definiten und infolgedessen endlichen Gesamtheit zu tun hat, deren Elemente in unserer Erfahrungswelt fixierbar sind und für die ein Auswahlverfahren effektiv definiert werden kann. Fishers „hypothetical infinite populations“ sind keine definiten Gesamtheiten im Sinne der Sozialstatistik, sondern begrifflich konstruierte *indefinite* Gesamtheiten. In dieser Hinsicht entsprechen sie ziemlich genau Fechners Kollektivgegenständen, mit denen wir uns in Abschnitt 2.2 beschäftigt haben.[6]

4. So wie wir Fechners Konzeption der Kollektivgegenstände als eine etwas obskure Formulierung für die Unterstellung eines hypothetischen Zufallsgenerators zu verstehen versucht haben, kann man auch versuchen, Fishers hypothetische Populationen zu verstehen. Allerdings wird dann ein Problem noch deutlicher, auf das bereits bei der Diskussion von Fechners Ansatz hingewiesen worden ist. Die Redeweise von Populationen impliziert, daß es ihre Elemente mit jeweils bestimmten Eigenschaften in gewisser Weise schon gibt. Die Vorstellung ist, daß man sich eine Gesamtheit *bestimmter* Dinge oder Situationen oder Ereignisse zumindest vorstellen kann, so daß sich die in Gestalt einer Stichprobe vorliegenden Daten auf einen Teil dieser Sachverhalte beziehen. Wenn man also z.B. Daten über Körpergrößen von 500 Studenten hat, impliziert Fishers Konzeption, daß man sich eine Gesamtheit aller Studenten vorstellen könne, die es jemals gegeben hat und geben wird; und es ist diese indefinite „Gesamtheit aller Studenten“, deren Körpergrößenverteilung ermittelt werden soll.[7]

[6] Als Hintergrund für die Entwicklung der Fisher'schen Konzeption kommt allerdings nicht unmittelbar Fechner in Betracht. Die Vorstellung, daß man sich problemlos auf indefinite Gesamtheiten beziehen könne, findet sich indessen auch schon bei F. Y. Edgeworth und Karl Pearson.

[7] Daß es sich um eine obskure Idee handelt, wurde z.B. sehr deutlich von F. P. Ramsey (1931, S. 204) zum Ausdruck gebracht: „The science of statistics is concerned with abbreviating facts about numerous individuals which are interpreted as a random selection from an infinite 'population'. If the qualities concerned are discrete, this means simply that we consider the proportions of the observed individuals which have the qualities, and ascribe these proportions to the hypothetical population. If the qualities are continuous, we take the population to be of a convenient simple form containing various parameters which are then chosen to give the highest probability to the instances observed. In either case the probable error is calculated for such a sample from such a population. (For all this see Fisher.) The significance of this procedure is that we record in a convenient simple form

(1) The approximate proportions having the given qualities in different degrees,

5. Ersetzt man Fishers hypothetische Populationen durch hypothetische Zufallsgeneratoren, kann man zwar ein obskures Reden über indefinite Mengen (die Elemente enthalten, die es noch nicht gibt) vermeiden; es ist jedoch wichtig, daß sich eine Reihe von Gedankengängen der theoretischen Statistik, insbesondere auch bei Fisher, gerade der Vorstellung verdanken, daß indefinite Populationen als Sachverhalte aufgefaßt werden können. Bei Fisher wird das besonders deutlich in seiner Konzeption des Wahrscheinlichkeitsbegriffs. Die Ausführungen in seiner oben zitierten Abhandlung gehen so weiter:

> “When we speak of the *probability* of a certain object fulfilling a certain condition, we imagine all such objects to be divided into two classes, according as they do or do not fulfil the condition. This is the only characteristic in them of which we take cognisance. For this reason probability is the most elementary of statistical concepts. It is a parameter which specifies a simple dichotomy in an infinite hypothetical population, and it represents neither more nor less than the frequency ratio which we imagine such a population to exhibit. For example, when we say that the probability of throwing a five with a die is one-sixth, we must not be taken to mean that of any six throws with that die one and one only will necessarily be a five; or that of any six million throws, exactly one million will be fives; but that of a hypothetical population of an infinite number of throws, with the die in its original condition, exactly one-sixth will be fives.” (Fisher 1922, S. 312)

So soll also die Vorstellung, daß man es mit einer Gesamtheit von Dingen, Situationen oder Ereignissen zu tun hat, dazu dienen, um Wahrscheinlichkeiten als relative Häufigkeiten definierbar zu machen.[8] Aber der Gedankengang ist nicht nur logisch anfechtbar: denn relative Häufigkeiten können nur in endlichen Gesamtheiten definiert werden.[9] Schließlich noch problematischer ist die implizierte Vorstellung, daß die Gesamtheit

(2) The number of instances which we have observed (the weight of our induction) (probable error).

For the use of the figures to give a degree of belief with regard to a new instance no rule can be given. The introduction of an infinite population is a stupid fiction, which cannot be defended except by some reference to proceeding to a limit, which destroys its sense.“

[8] Die Idee, Wahrscheinlichkeiten durch relative Häufigkeiten in indefiniten Gesamtheiten (Klassen) zu definieren, hat bereits Vorläufer im Anschluß an die Kollektivmaßlehre Fechners gefunden; vgl. z.B. Helm, 1902.

[9] Es ist in diesem Zusammenhang bemerkenswert, daß Fisher den Begriff einer Häufigkeitsverteilung, der ja zunächst nur für eine endliche Gesamtheit sinnvoll definiert ist, von vornherein auf hypothetische Populationen bezieht: „No finite sample has a frequency curve: a finite sample may be represented by a histogram, or by a frequency polygon, which to the eye more and more resembles a curve, as the size of

schon dann existiert, wenn man sich tatsächlich nur auf eine Stichprobe beziehen kann, die sich auf mehr oder weniger viele *bisher* beobachtete Situationen bezieht. Man denke z.B. an die Typhus-Fälle, die v. Kries als Beispiel verwendet hat (vgl. Abschnitt 2.4). Hat man zum Beispiel n Typhus-Fälle beobachtet und festgestellt, welche von ihnen einen tödlichen Ausgang hatten, müßte man – dem Ansatz Fishers folgend – sich nicht nur vorstellen, daß es alle Typhus-Fälle heute schon gibt, und zwar auch die, die erst in der Zukunft auftreten werden; man müßte weiterhin glauben, daß auch heute schon bestimmt ist, welcher Anteil dieser Typhus-Fälle einen tödlichen Ausgang nehmen wird.

6. Somit kann Fishers Konzeption einer theoretischen Statistik folgendermaßen charakterisiert werden: Sie beginnt mit der Annahme, daß eine Menge von Daten (welcher Art auch immer) als eine Zufallsstichprobe aus einer hypothetisch unterstellten unendlichen Grundgesamtheit zustandegekommen ist; und die Aufgabe der theoretischen Statistik besteht dann darin, eine Verteilungsfunktion für die hypothetische Grundgesamtheit zu finden, die mit den jeweils verfügbaren Daten vereinbar erscheint. Wie schon bemerkt worden ist, kann man diese Konzeption auch (und weniger obskur) verstehen, wenn man das Reden von hypothetischen Populationen vermeidet und sich stattdessen auf hypothetische Zufallsgeneratoren bezieht, die die Sachverhalte, auf die sich die jeweiligen Daten beziehen, hervorgebracht haben (könnten). Anklänge an diese Sichtweise finden sich auch bei Fisher, zum Beispiel in der folgenden Bemerkung:

the sample is increased. To reach a true curve, not only would an infinite number of individuals have to be placed in each class, but the number of classes (arrays) into which the population is divided must be made infinite. Consequently, it should be clear that the concept of a frequency curve includes that of a hypothetical infinite population, distributed according to a mathematical law, represented by the curve.“ (Fisher 1922, S. 312) Aber der Wunsch, mit im mathematischen Sinne kontinuierlichen Verteilungsfunktionen operieren zu können, kann natürlich nicht die Annahme rechtfertigen, daß eine Stichprobe eine Teilmenge aus einer „hypothetical infinite population“ ist. — Man kann hier an einen vergleichbaren Gedanken bei Laplace (1814) erinnern, der angenommen hat, „daß man die Geburten mit dem Herausziehen von Kugeln aus einer Urne vergleichen kann, die eine unendliche Anzahl von weißen und schwarzen Kugeln enthält, welche so vermischt sind, daß bei jedem Zuge die Chancen des Herauskommens für jede Kugel dieselben sind.“ (S. 51f) Warum unendlich viele? „So vermehrt das Herausziehen einer weißen Kugel aus einer Urne, die eine beschränkte Zahl weißer und schwarzer Kugeln enthält, die Wahrscheinlichkeit, eine schwarze Kugel beim folgenden Zuge herauszuziehen. Aber das findet nicht mehr statt, wenn die Zahl der Kugeln der Urne unbegrenzt ist, wie man es voraussetzen muß, um diesen Fall dem der Geburten analog zu machen.“ (S. 125f)

“It should be noted that there is no falsehood in interpreting any set of independent measurements as a random sample from an infinite population; for any such set of numbers are a random sample from the totality of numbers produced by the same matrix of causal conditions: the hypothetical population which we are studying is an aspect of the totatlity of the effects of these conditions, of whatever nature they may be.” (Fisher 1922, S. 313)

In dieser Bemerkung kommt auch wieder eine typische Neigung mathematischer Statistiker zum Ausdruck: ihre fiktiven Begriffsbildungen für Erscheinungen beliebiger Art – „of whatever nature they may be“ – anwendbar zu halten. Die Fortsetzung des gerade angeführten Zitats macht schließlich noch einmal deutlich, worin Fisher die Hauptaufgabe des Statistikers erblickt:

“The postulate of randomness thus resolves itself into the question, ‘Of what population is this a random sample?’ which must frequently be asked by every practical statistician.” (ebda.)

4.2.2 Bezugnahme auf Experimente

1. Obwohl Fisher stets die allgemeine Anwendbarkeit seines Ansatzes für eine theoretische Statistik betont hat, so daß eine theoretische *Sozial*statistik bloß als ein Spezialfall erscheint, sollte zum besseren Verständnis darauf hingewiesen werden, daß es ihm in erster Linie um die Entwicklung eines Begriffsrahmens für *Experimente* ging. Praktisch hatte er es in erster Linie mit landwirtschaftlichen Experimenten zu tun.[10] Daß Fisher zunächst immer an Experimente gedacht hat, kommt gut in der folgenden Bemerkung aus seinem Buch über „The Design of Experiments“ (zuerst 1935) zum Ausdruck:

“Statistical procedure and experimental design are only two different aspects of the same whole, and that whole comprises all the logical requirements of the complete process of adding to natural knowledge.” (Fisher 1966, S. 3)

Zahlreiche ähnliche Bemerkungen finden sich in Fishers Schriften.[11] Es ist auch dieser Kontext, das Experimentieren, in dem Fisher sich haupt-

[10]Von 1919 bis 1933 arbeitete Fisher als Statistiker an der landwirtschaftlichen Versuchsanstalt in Rothamsted (England). Etwa in diese Zeit hat er auch die Grundgedanken seiner Konzeption einer theoretischen Statistik entwickelt. Vgl. Johnson und Kotz 1997, S. 101ff.

[11]Hier noch ein Beispiel: „It was in the study of experimental data that the characteristics of modern statistics first began to display themselves, notably in that series of methods which are known as the tests of significance.“ (Fisher 1951, S. 188) Denkt man daran, wie sich eine theoretische Statistik in der Geschichte der Sozialstatistik

sächlich bemüht hat, den Sinn der hypothetischen Unterstellungen seines Ansatzes zu einer theoretischen Statistik zu begründen. Im Zentrum steht seine Überzeugung, daß statistische Inferenz generalisierbare Einsichten in die Funktionsweise experimenteller Verfahren liefern kann.

2. Aus zwei Gründen ist der enge Zusammenhang zwischen Fishers Ansatz zu einer theoretischen Statistik und Experimenten hier zu erwähnen. Erstens weil sich durch eine Bezugnahme auf Experimente eine durchaus sinnvolle Analogie zu Zufallsgeneratoren herstellen läßt; zweitens, weil dadurch ein bemerkenswerter Kontrast zur Sozialstatistik entsteht, die es – jedenfalls in ihren primären Betätigungsfeldern – nicht mit experimentell erzeugbaren Prozessen zu tun hat. Verfolgen wir zunächst die Analogie. Drei Gesichtspunkte erscheinen für ein Verständnis von Experimenten wesentlich. Zuerst eine Unterscheidung von passiven Beobachtungen. d'Alembert hat diese Unterscheidung in der zusammen mit Diderot herausgegebenen „Enzyklopädie" (1756, S. 434) so formuliert:

„Die Beobachtung, die weniger originell und tiefgründig ist, beschränkt sich auf die Tatsachen, die man vor Augen hat, das heißt darauf, Erscheinungen aller Art, die uns das Schauspiel der Natur darbietet, gut zu betrachten und ausführlich zu beschreiben; dagegen sucht das Experiment die Natur tiefer zu erforschen, ihr das zu entreißen, was sie verbirgt, und durch mannigfache Kombination der Körper neue Erscheinungen hervorzubringen, um diese wiederum zu studieren – kurz, es beschränkt sich nicht darauf, die Natur zu belauschen, sondern es befragt sie und zwingt sie zur Auskunft."

Ein zweiter Gesichtspunkt entsteht daraus, daß bei Experimenten *planmäßig* vorgegangen wird. Ein Experiment kann durch eine jeweils bestimmte Vorgehensweise charakterisiert werden oder, wie man auch sagen kann, durch ein *Verfahren*. Damit hängt unmittelbar ein dritter Gesichtspunkt zusammen: Experimente können wiederholt werden, denn der Begriff bezieht sich auf ein Verfahren.

3. Die Analogie zu Zufallsgeneratoren entsteht daraus, daß auch diese als Verfahren definiert sind, durch die Sachverhalte oder Ereignisse hervorgebracht werden können. In beiden Fällen charakterisiert das Verfahren die Vorgehensweise. Durch das Verfahren wird *festgelegt*, wie der Zufallsgenerator beschaffen sein und verwendet werden soll. Ganz analog legt bei einem Experiment das Verfahren fest, wie bei der Durchführung

herausgebildet hat, ist die Bemerkung kaum zutreffend; sie wird aber verständlich, wenn man an den experimentellen Kontext denkt, in dem Fisher seine Ideen entwickelt hat.

des Experiments vorgegangen werden soll und wie die Bedingungen, unter denen das Experiment auszuführen ist, und ggf. dafür zu verwendende Apparate und Vorkehrungen beschaffen sein sollen. Dies impliziert auch, daß man bei der Durchführung eines Experiments in der Lage sein muß, die Bedingungen zu kontrollieren, unter denen es durchgeführt wird. Hier muß man allerdings auf die Formulierung achten. Der allgemeine Begriff eines Experiments setzt nicht voraus, daß man in der Lage sein muß, die Bedingungen, unter denen es durchgeführt wird, planmäßig zu variieren. Man muß nur in der Lage sein, das Verfahren richtig anzuwenden, durch das das Experiment definiert ist. So verhält es sich auch bei einem Zufallsgenerator, bei dem überhaupt keine Möglichkeit besteht, durch Variation der Bedingungen auf das Ergebnis Einfluß zu nehmen.

4. Noch etwas ist wichtig für die Analogie zwischen Experimenten und Zufallsgeneratoren. In beiden Fällen ist ein Akteur erforderlich, der das Experiment durchführt bzw. den Zufallsgenerator betätigt. Bei Experimenten ist dies offensichtlich. Es bedarf eines handelnden Menschen, der das Experiment durchführt. Es ist dieser Mensch, der – mithilfe des Experiments – „die Natur befragt und zur Auskunft zwingt". Ganz entsprechend verhält es sich bei Zufallsgeneratoren. Sie sind durch Verfahren definiert; aber ein Verfahren wird nicht von selbst aktiv. Der Würfel springt nicht von allein vom Tisch, sondern jemand muß ihn nehmen und dadurch den Zufallsgenerator betätigen.[12] Dies hat auch eine wichtige Implikation für die Frage, in welcher Weise Zufallsgeneratoren zur begrifflichen Repräsentation von Prozessen verwendet werden können. Sie können verwendet werden, um sich ein Bild davon zu machen, wie Prozesse ablaufen. Aber sie implizieren keine Vorstellungen darüber, wodurch Prozesse zustande kommen. Zum Beispiel bezieht sich die Wahrscheinlichkeitsverteilung, durch die eine Urne charakterisiert wird, nicht darauf, ob überhaupt Kugeln aus der Urne herausgezogen werden. Die Wahrscheinlichkeitsverteilung bezieht sich auf ein Verfahren, nicht darauf, ob mit dem Verfahren Ereignisse erzeugt werden.

[12]Allerdings ist es möglich, Zufallsgeneratoren partiell von menschlichen Akteuren abzulösen. Z.B. kann man sich eine Verknüpfung von Zufallsgeneratoren und Uhren in der Weise vorstellen, daß zu bestimmten Uhrzeiten ein Automatismus einsetzt, der einen Zufallsgenerator auslöst. In dieser Weise werden Uhren und Taktgeber bei der Konstruktion vieler Arten von Maschinen verwendet. Dem entspricht als Verallgemeinerung der Konzeption von Zufallsgeneratoren die Idee stochastischer Prozesse. Modellvorstellungen der theoretischen Sozialstatistik, die von stochastischen Prozessen ausgehen, werden wir jedoch in diesem Skript nicht besprechen.

5. Die eben gemachte Bemerkung ist auch deshalb wichtig, weil sie auf einen problematischen Aspekt in einer Formulierung hinweist, die bisher oft verwendet wurde: daß theoretische Sozialstatistik mit der Annahme beginnt, daß soziale Sachverhalte oder Ereignisse durch hypothetisch unterstellbare Zufallsgeneratoren *zustande kommen*. Die Formulierung trifft zwar in vielen Fällen die Art und Weise, wie Vorstellungen der Wahrscheinlichkeitsrechnung verwendet werden; z.B. wenn Laplace davon spricht, daß man sich die Geburten von Kindern als Ziehungen aus einer Urne vorstellen könne. Aber es ist klar, daß das Wort 'zustande kommen' in diesem Zusammenhang eine Zweideutigkeit aufweist. Bezieht man sich in der Wahrscheinlichkeitsrechnung auf einen Zufallsgenerator in der Form einer Urne, geht es tatsächlich nicht darum, ob, wann und von wem eine Kugel aus der Urne herausgezogen wird.[13] Auch wenn man bereit sein sollte, eine Analogie zwischen der Geburt von Kindern und dem Ziehen von Kugeln aus einer Urne als eine Denkmöglichkeit anzuerkennen, kann deshalb mit dieser Vorstellung nicht erklärt werden, ob, wann und von wem ein Kind geboren wird. Der potentielle Sinn der Analogie kann also nur darin liegen, daß sie sich auf *die Art und Weise des Ablaufs* von Prozessen bezieht.

6. Wir werden später besprechen, daß auch dieses Verständnis nicht ausreicht, um Modelle der theoretischen Sozialstatistik zu verstehen. Es kann aber zunächst verwendet werden, um den Sinn einer Analogie zwischen Experimenten und Zufallsgeneratoren zu verstehen, die erforderlich ist, um Begriffsbildungen und Vorstellungen der Wahrscheinlichkeitstheorie bzw. einer von ihr ausgehenden Statistik zur Charakterisierung von Experimenten zu verwenden. Im Hinblick auf Experimente erscheint die Analogie auch sinnvoll; denn in diesem Fall erfindet der Statistiker keine fiktiven Zufallsgeneratoren, sondern kann sich mit seinen Begriffsbildungen direkt auf Experimente *als Verfahren*, die tatsächlich durchgeführt werden können, beziehen. Natürlich müssen Hypothesen aufgestellt werden. Aber sie beziehen sich in diesem Fall darauf, wie ein Verfahren funktioniert, das empirisch exemplifiziert werden kann. Anders verhält es sich jedoch in der theoretischen Sozialstatistik, deren Zufallsgeneratoren sich nicht auf Experimente – im Sinne realisierbarer Verfahren –, sondern auf „soziale Prozesse" beziehen. In diesem Fall sind die Zufallsgeneratoren selbst hypothetischer Natur, oder genauer gesagt: es handelt sich um *fiktive* Zufallsgeneratoren. Wie schon in Abschnitt 2.4 besprochen worden ist, kann die Annahme, daß sich die Prozesse, durch

[13] Vgl. hierzu auch die Überlegungen in Abschnitt 3.1.2.

die soziale Sachverhalte und Ereignisse entstehen, durch Zufallsgeneratoren konzeptualisieren lassen, nicht als eine Hypothese im üblichen Sinne des Wortes verstanden werden. Insbesondere handelt es sich nicht um eine statistische Hypothese, vielmehr um ein Postulat bzw. um eine Fiktion, die statistische Hypothesen formulierbar machen soll.

7. Diese Bemerkungen beziehen sich natürlich auf ein Verständnis von Sozialstatistik, bei dem nicht von vornherein eine experimentelle Betrachtungsweise eingenommen wird. Es entspricht insoweit dem Selbstverständnis der Sozialstatistik in ihrer historischen Entwicklung. Daß auch im Bereich der sozialen Beziehungen zwischen Menschen eine experimentelle Betrachtungsweise eingenommen werden kann, ist selbstverständlich. Sie beginnt schon, wenn sich der Geschäftsführer eines Supermarkts überlegt, wie er seine Waren plazieren sollte, um ihren Verkauf zu maximieren. Insbesondere in der Psychologie und Sozialpsychologie hat sich die Vorstellung verbreitet, daß es sinnvoll sein könnte, das Verhalten von Menschen durch Experimente zu untersuchen. Statistische Begriffsbildungen und Modelle können daran anschließen, indem sie diese Experimente als Zufallsgeneratoren konzipieren. In diesem Skript beziehen wir uns dagegen auf die traditionelle Sozialstatistik bzw. ihre Fortsetzungen in der gegenwärtigen empirischen Sozialforschung. Ihre Daten entstehen nicht durch Experimente; und infolgedessen können auch die Prozesse, die die Sachverhalte hervorbringen, auf die sich die Daten beziehen, nicht als Experimente – in irgendeinem ernsthaften Sinne des Wortes – konzipiert werden.[14]

4.3 Gesamtheiten in der Sozialstatistik

Fishers Ansatz ist für die Entwicklung der Statistik sehr einflußreich geworden. Uns interessiert hier hauptsächlich, wie seine Konzeption „hypothetischer Populationen" mit Vorstellungen über sozialstatistische Gesamtheiten verbunden und vermischt worden ist.

[14] Als Beispiel für eine bloß metaphorische Verwendung des Experimentbegriffs kann folgende Formulierung von J. Galtung (1957, S. 181) dienen: „By an *experiment* we mean a part of the empirical world that can be analyzed in terms of a set of *variables* (independent and dependent) and a set of *conditions* (relevant and irrelevant) that produce an *outcome* in terms of the dependent variable." Vgl. auch die Hinweise auf „fiktive Experimente" in Abschnitt 3.2.2.

4.3.1 Entwicklung der neuen Rhetorik

1. Daß sich statistische Begriffsbildungen auf Gesamtheiten beziehen, ist natürlich ein Gedanke, der die Entwicklung der Statistik von Anfang an geprägt hat. In der Sozialstatistik war auch mehr oder weniger klar, was gemeint ist: Gesamtheiten von Dingen, Menschen oder Situationen, die in der Erfahrungswelt fixiert werden können. Solche Gesamtheiten sollten mit statistischen Begriffsbildungen *beschrieben* werden. Daraus entwickelte sich das *sozialstatistische* Inferenzproblem: ob man auch anhand von Stichproben statistische Aussagen über Gesamtheiten einschätzbar machen kann, und wie man die Stichproben auf zweckmäßige Weise bilden sollte. Zunächst war keineswegs allgemein anerkannt, daß man nach Möglichkeit Zufallsstichproben verwenden sollte, um Gedankengänge der Wahrscheinlichkeitstheorie anwendbar zu machen. Diese Frage wurde erst seit etwa Anfang der 1920er Jahre intensiv diskutiert.[15] Schließlich setzte sich jedoch in den 30er und 40er Jahren die Auffassung durch, daß nur Zufallsstichproben für wahrscheinlichkeitstheoretisch begründbare Inferenzverfahren verwendet werden können.

2. Eine wichtige Rolle in dieser Diskussion spielte Jerzy Neyman (1894–1981),[16] insbesondere seine Abhandlung „On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection“ (1934). Bemerkenswert ist, daß sich Neyman in dieser Abhandlung auf Fishers Konzeption einer theoretischen Statistik beruft, bei der es, wie wir gesehen haben, gar nicht um sozialstatistische Gesamtheiten und ein sozialstatistisch konzipiertes Inferenzproblem geht, sondern um eine Unterstellung hypothetischer Zufallsgeneratoren für Prozesse (insbesondere Experimente), durch die beobachtbare Sachverhalte und Ereignisse zustande gekommen sein könnten.[17]

[15] Die Diskussion wurde durch Beiträge von A.N. Kiär inganggesetzt; ein deutschsprachiger Beitrag erschien 1898 im „Allgemeinen statistischen Archiv“. Weiterhin bemerkenswert ist eine Arbeit von E. Altschul (1913), in der die Diskussion in die Frage eingebettet wird, was die Sozialstatistik von der Biometrie lernen könne. Die Diskussion entwickelte sich indessen nur langsam. Erst 1924 beauftragte das *International Statistical Institute* eine Kommission, um die „repräsentative Methode“ zu untersuchen. Der Kommission gehörten an: A.L. Bowley, C. Gini, A. Jensen, M.L. March, V. Stuart und F. Zizek. Der Bericht wurde von Jensen (1926) abgegeben. Zur Entwicklungsgeschichte der „repräsentativen Methode“ im Zeitraum von 1895 bis 1939 vgl. Kruskal und Mosteller, 1980.

[16] Biographische Angaben finden sich bei Johnson und Kotz 1997, S. 137ff.

[17] Neyman (1934, S. 562): „... an approach to problems of this type has been suggested by Professor R.A. Fisher which removes the difficulties involved in the lack of knowledge of the *a priori* probability law. Unfortunately the papers referred [gemeint sind die Veröffentlichungen von Fisher] have been misunderstood and the validity of statements they contain formally questioned. This I think is due largely to the very condensed form of explaining ideas used by R.A. Fisher, and perhaps also to a somewhat difficult method of attacking the problem. Avoiding the necessity of appeals to the somewhat vague statements based on probabilities *a posteriori*, Fisher's theory becomes, I think, the very basis of the theory of representative method.“

3. Bemerkenswert ist auch eine etwas spätere Arbeit Neymans, „Outline of a theory of statistical estimation based on the classical theory of probability“ (1937), in der die Unterscheidung zwischen sozialstatistischen Stichproben und wiederholbaren Experimenten explizit irrelevant gemacht wird. Die Unterscheidung wird zunächst aufgegriffen:

“We shall distinguish two aspects of the problems of estimation: (i) the practical and (ii) the theoretical. The practical aspect may be described as follows:

(i*a*) The statistician is concerned with a population, π, which for some reason or other cannot be studied exhaustively. It is only possible to draw a sample from this population which may be studied in detail and used to form an opinion as to the values of certain constants describing the properties of the population π. For example, it may be desired to calculate approximately the mean of a certain character possessed by the individuals forming the population π, etc.

(i*b*) Alternatively, the statistician may be concerned with certain experiments which, if repeated under apparently identical conditions, yield varying results. Such experiments are called random experiments. To explain or describe the machinery of the varying results of random experiments certain mathematical schemes are drawn up involving one or more parameters, the values of which are not fixed. The statistician is then asked to provide numerical values of these parameters, to be calculated from experimental data and upon the assumption that the mathematical model of the experiments is correct. [...]

In both cases described, the problem with which the statistician is faced is the problem of estimation. This problem consists in determining what arithmetical operations should be performed on the observational data in order to obtain a result, to be called an estimate, which presumably does not differ very much from the true value of the numerical character, either of the population π, as in (i*a*), or of the random experiments, as in (i*b*).” (Neyman 1937, S. 333f)

Neyman versucht dann zu zeigen, daß es aus der Sicht des Statistikers keinen Grund gibt, die beiden Fragestellungen zu unterscheiden:

“Sampling randomly from the population π, it is possible to obtain samples, say

$$E_1, E_2, \dots, E_N$$

where each sample is described by means of values of the characters $X, Y, \ldots, Z$, corresponding to each of the individuals forming the sample. The probability of any sample E_i, say $\mathrm{P}(E_i \mid \theta_1, \theta_2, \ldots, \theta_l)$, depends on a certain number, l, of parameters θ_i, the values of which are unknown, describing the properties of the population π. The problem consists in determining how to use the sample which may be actually obtained in order to estimate θ_1.

We see that the conditions of the problem in (i*a*) are expressed in terms of probability. The same holds good with regard to the problem in (i*b*), which shows that the distinction between (i*a*) and (i*b*) is only superficial." (Neyman 1937, S. 335)

Somit wird deutlich, in welcher Weise Neyman versucht, die Unterscheidung zwischen sozialstatistischen Gesamtheiten, die als in der Erfahrungswelt gegebene Sachverhalte verstanden werden können, und Experimenten, bei denen es sich um *Verfahren* zur Erzeugung von Sachverhalten und Ereignissen handelt und zu deren Charakterisierung mithin auf die Konzeption eines Verfahrens bezug genommen werden muß, als *aus statistischer Sicht* überflüssig aufzuzeigen.

4. Somit konnte die Vorstellung entstehen, daß sich alle Fragen statistischer Inferenz gleichermaßen der Idee bedienen können, daß man es mit Zufallsstichproben aus (irgendwie) *gegebenen* Populationen zu tun hat. Schon in einer früheren Gemeinschaftsarbeit von Neyman und E. S. Pearson (1928) wurde diese Vorstellung so formuliert:

"One of the most common as well as most important problems which arise in the interpretation of statistical results, is that of deciding whether or not a particular sample may be judged as likely to have been randomly drawn from a certain population, whose form may be either completely or only partially specified." (Neyman und Pearson 1928, S. 175)

Diese Probemformulierung ist weitgehend identisch mit dem in Abschitt 4.2 behandelten Ansatz von Fisher; jetzt ist jedoch ein Gedanke hinzugekommen, der bei Fisher nur latent vorhanden ist: daß zwischen der Durchführung von Experimenten und der Auswahl von Stichproben aus gegebenen Gesamtheiten keine für die Statistik relevante konzeptionelle Unterscheidung erforderlich sei.[18]

5. Bemerkenswert ist, wie rasch sich diese Vorstellung verbreitet hat. Während in der älteren Literatur Begriffsbildungen der Wahrscheinlichkeitstheorie stets mehr oder weniger explizit durch Vorstellungen

[18]Erst seit etwa 1950 hat sich dagegen eine eigenständige Stichprobentheorie für endliche Gesamtheiten entwickelt. Eine kurze Zusammenfassung der wichtigsten Entwicklungsetappen findet sich bei T. M. F. Smith, 1976. Wir werden darauf in Abschnitt 4.4 etwas näher eingehen.

über Zufallsgeneratoren (oder durch Rückgriff auf Vorstellungen aus der Theorie der Beobachtungsfehler) motiviert und begründet wurden,[19] haben sich die Lehrbücher ab den 30er Jahren zunehmend an der Vorstellung orientiert, daß man Daten als Zufallsstichproben aus irgendwie gegebenen Populationen auffassen könne. Eines der ersten Lehrbücher diese Art ist die „Einführung in die mathematische Statistik“ von O. Anderson (1935), auf die weiter unten etwas genauer eingegangen wird.[20] Bemerkenswert ist ein Lehrbuch von T. C. Fry (1928), das zwar noch keinen systematischen Gebrauch von der Populationsrhetorik macht, diese aber an einer Stelle bereits einführt und kommentiert:

"When an 'experiment' is performed we really sort out one of a group of possibilities. Thus, when a penny turns up 'heads' we sort out one of *two* possibilities; or when we measure the resistance of an electrical device we sort out one of an infinity of possibilities. [...] Thinking along these lines leads us quite naturally to the conception of a vast storehouse of possibilities from which, by our experiment, we have drawn one blindly. This hypothetical storehouse we may call the 'universe' (or 'population') from which our 'sample' was drawn. Naturally, we must imagine it to contain all possible results in proportions equal to their respective probabilities, which will ordinarily require us to think of an infinity of items in the storehouse (or, 'individuals' in the 'population'). [...]

These words 'universe' or 'population', 'individual', and 'sample' are in constant use in statistical literature, and they are really quite valuable, for they enable us to talk quite abstractly about events of such diverse sorts that no sufficiently general word could otherwise be found for them. We have introduced them as relating to a hypothetical storehouse the properties of which are determined by the probabilities, known or unknown, which govern our experiment. It is only fair to say that they did not so originate. They are the product of a statistical school which, if it does not actually deny that a probability exists before an experiment is performed, at least regards the experiment as the only feasible means of arriving at it. If I interpret them correctly, the adherents of this school attribute to a 'population' a degree of logical reality which transcends that of abstract 'probability'. To them, the 'population' or 'universe' is the actuality, while 'probability' is relegated to a shorthand form of statement for 'the relative frequency of like individuals in the population'.

Though I appreciate the practical usefulness of the 'universe' concept in aiding imagination, I do not see that the logic of a subject is in any way bettered

[19]Als Beispiele seien genannt: Czuber 1908, 1921. Coolidge 1927, Fry 1928, Uspensky 1937.

[20]Als Beispiele sei weiterhin verwiesen auf: Elderton 1938, Hagood 1941, Hagood und Price 1952, Quenouille 1950.

by substituting for an abstraction (probability) that sort of quasi-concreteness which dreams possess: that is, the mental image of a 'universe' which would be concrete enough if it existed, but which never does, and frequently could not, exist." (Fry 1928, S. 266ff)

6. Wie rasch sich die Vorstellung verbreitet hat, daß es auch Sozialwissenschaftler stets mit einer stichprobentheoretisch zu reflektierenden Fragestellung zu tun haben, zeigt exemplarisch folgende Bemerkung von Gösta Carlsson aus dem Jahr 1951:

"In the social sciences we usually deal with samples instead of complete populations. There is a body of knowledge relating to the errors engendered by this procedure, but there will be no need to discuss the technical aspects of the statistical theory of sampling here. Suffice is to say that expected sampling errors can be computed by methods derived from the mathematical theory of probability. That this should be done, and that a sampling error should be attached to every statistic descriptive of a sample, are principles firmly embedded in the mind of every research worker with a decent methodological upbringing." (Carlsson 1951, S. 139)

Die Bemerkung trifft allerdings kaum für die Entwicklung der Sozialstatistik in Deutschland zu.[21]

7. In Deutschland ist nicht nur Fishers Konzeption einer theoretischen Statistik zunächst weitgehend ohne Einfluß geblieben; dasselbe gilt für die insbesondere von Neyman und E. S. Pearson verbreitete Idee, daß sozialstatistische Stichproben und experimentell gewonnene Daten durch eine einheitlich konzipierte Theorie statistischer Inferenz behandelt werden können. Die Entwicklung der *Sozial*statistik blieb bis in die 50er Jahre des 20. Jahrhunderts überwiegend von einer deskriptiven Haltung geprägt, wie sie um die Jahrhundertwende vornehmlich von G. v. Mayr kanonisiert worden war (vgl. Abschnitt 1.1). Eine *theoretische* Sozialstatistik, wie sie von Autoren wie Lexis, v. Bortkiewicz, Fechner u.a. anvisiert worden war – also eine Konzeption, bei der Begriffsbildungen der Wahrscheinlichkeitstheorie dienen sollen, um sich auf Prozesse

[21] Auch in der amerikanischen Soziologie vollzog sich diese Entwicklung im wesentlich erst in den 1940er Jahren. Noch 1934 bemerkte F. A. Ross (1934, S. 473f): „It is amazing how few sociologists using statistical methods seem to be aware of the possibility, to say nothing of the desirability or necessity, of applying these measures [Maßzahlen für „Stichprobenfehler"] to the results of their statistical researches. [...] No sociologist should apply statistics to his problem until he has mastered the fundamentals of the theory of sampling, until he understands the nature of standard errors of the statistical measures which he intends to use, and is fully equipped to apply the tests that they afford." Zur historischen Entwicklung in den USA vgl. auch Bernert, 1983. Zu parallelen Entwicklungen in der Ökonometrie vgl. Morgan, 1990.

zu beziehen, durch die soziale Sachverhalte und Ereignisse entstehen –, war während dieses Zeitraums eher eine Randerscheinung.[22] Ein Grund dafür war die in der deutschen Sozialstatistik lange Zeit verbreitete Abneigung gegen eine Verwendung mathematischer Hilfsmittel, die über ein bloßes Zählen und eine Berechnung von Häufigkeiten und Mittelwerten hinausgehen.[23]

8. Eines der ersten deutschsprachigen Lehrbücher, das Fishers Konzeption hypothetischer Populationen zwar nicht systematisch folgt, jedoch seine Rhetorik aufgreift, ist der „Grundriß der theoretischen Statistik" von W. Winkler (1931), auf den schon mehrfach hingewiesen worden ist. Dort heißt es:

„Bekanntlich faßt die moderne Statistik jede statistische Masse, möge sie auch in größter Vollständigkeit erfaßt worden sein, als eine Musterprobe ('sample') aus einer unendlich großen Masse gleicher Wesensbeschaffenheit ('Wesensform') auf, eine Probe, die je nach ihrer absoluten Größe mehr oder weniger dem Einflusse der 'Zufallsstreuung' unterliegt, also eine von der Grundform abweichende Gestalt annimmt." (Winkler 1931, S. 156)

Wie schon früher bemerkt worden ist, geht es Winkler bei dieser Betrachtungsweise nicht um das sozialstatistische Inferenzproblem, bei dem man es mit einer durch ein effektives Auswahlverfahren definierten Stichprobe aus einer definiten Gesamtheit zu tun hat. Winkler weist ausdrücklich darauf hin, daß – wie bei Fisher – *jede* sozialstatistische Gesamtheit als Zufallsstichprobe aus einer umfassend gedachten „Obergesamtheit" zu betrachten sei.

[22] Ich betone dies, weil wir uns in diesem Skript mit Ansätzen zu einer theoretischen Sozialstatistik beschäftigen, uns also in erster Linie mit Autoren beschäftigen, die zu ihrer Entwicklung beigetragen haben. Es wäre aber falsch, anzunehmen, daß diese Konzeptionen die Entwicklung der Sozialstatistik von Anfang an dominiert haben.

[23] So versteht man auch, daß – jedenfalls in Deutschland – theoretische Sozialstatistik, die Begriffsbildungen der Wahrscheinlichkeitsrechnung verwendet, oft auch „mathematische Statistik" genannt wurde. Eines der letzten Statistik-Lehrbücher, die sich – in der Tradition von G. v. Mayr – vehement gegen mathematische und insbesondere wahrscheinlichkeitstheoretische Begriffsbildungen ausgesprochen haben, ist die „Theoretische Statistik" von H. Wolff (1926). Dort heißt es z.B. (S. 347): „'Daß auf formal-mathematischem Wege die tatsächliche Gestaltung der Beziehungsverhältnisse der sozialen Massen abschließend erkannt werden könnte, dürfte wohl ausgeschlossen sein.', sagt richtig, wenn auch einschränkend, G. von Mayr. Ich gehe noch weiter und behaupte, daß keine mathematische Formel geeignet ist, die materielle Auswertung statistischer Beziehungen zu ermöglichen. Nur die formale Darstellung kann mathematisch-zahlenmäßig gefördert werden; die statistische Erkenntnis dagegen nicht." Es sei aber darauf hingewiesen, daß einige Autoren versucht haben, eine solche Konzeption der Sozialstatistik fortzusetzen, man vgl. z.B. Blind, 1953.

9. Wichtig geworden sind dann vor allem Arbeiten von Oskar Anderson, insbesondere seine „Einführung in die mathematische Statistik" (1935), in der Anderson versucht, systematisch ein *allgemeines* Konzept statistischer Gesamtheiten zu entwickeln. Der Gedankengang wird folgendermaßen motiviert:

„Wir haben bereits in der Einleitung festgestellt, daß die Grundform, in welcher man gewöhnlich die Resultate der Auszählung homograder Gesamtheiten [qualitative statistische Variablen] darstellt, die *relative Häufigkeit* (bzw. das 100fache oder 1000fache von ihr) der verschiedenen Merkmale und Merkmalsgruppen in ihr ist. So berechnen wir z.B. die relative Häufigkeit der Personen, die im arbeitsfähigen Alter stehen (ein Merkmal), der arbeitsfähigen Männer (Kombination zweier Merkmale), der arbeitsfähigen deutschen Männer (Kombination dreier Merkmale) usw. Unser letztes Ziel besteht hierbei entweder darin, nur die Gliederung eben der gegebenen konkreten Gesamtheit zu erfassen, oder aber (und dies dürfte der häufigere Fall sein), auf Grund der relativen Häufigkeiten, die eine gewisse Gesamtheit aufweist, zu den uns unmittelbar nicht gegebenen und daher gewöhnlich ganz unbekannten relativen Häufigkeiten in einer verwandten Gesamtheit *höherer Ordnung* vorzudringen. Gegeben sei z.B. die relative Häufigkeit der Knabengeburten unter allen Geburten, die eine Gruppe 'nordischer' Ehepaare im Laufe eines Jahres aufzuweisen hat. Gefragt wird nach der relativen Häufigkeit aller 'nordischen' Knabengeburten desselben Jahres (Gesamtheit höherer Ordnung) oder solcher Geburten in einer längeren Reihe von Jahren (Gesamtheit nächsthöherer Ordnung)." (Anderson 1935, S. 28f)

Die Idee besteht also darin, daß man sich zu jeder Menge von Daten eine oder mehrere „Gesamtheiten höherer Ordnung" bilden kann. Dabei kann es sich, wie Anderson (S. 30) hinzufügt, um Gesamtheiten handeln, die „ein objektives Dasein führen, durch bloße Willkür des Forschers gebildet werden, oder sogar rein gedankliche Konstruktionen darstellen."

10. Ganz ähnlich wie Fisher verwendet auch Anderson diesen Gedanken, um zu erläutern, wie er von 'Wahrscheinlichkeit' sprechen möchte:

„An die Stelle des ungeschickten Ausdrucks 'relative Häufigkeit eines Merkmals in einer Gesamtheit höherer Ordnung als die gegebene' setzen wir nun den Ausdruck: 'statistische Wahrscheinlichkeit des Merkmals'. Wir werden uns im Laufe der weiteren Ausführungen davon überzeugen können, daß diese Definition bereits ausreichend ist, um alle *formal mathematischen Theoreme* der Wahrscheinlichkeitsrechnung auf ihr aufzubauen. Wir schreiben 'formal mathematische', um anzudeuten, daß es von diesen Theoremen einen unmittelbaren Übergang zur konkreten Tatsachenwelt noch nicht gibt, denn nach unserem Dafürhalten kann das sogenannte 'Gesetz der großen Zahlen' allein mit Hilfe mathematischer 'tautologischer Umformungen' *nicht* bewiesen werden.

Um diesen Übergang zur Tatsachenwelt zu schaffen, müssen gewisse zusätzliche Annahmen eingeführt werden, die nicht mathematischer, sondern empirischer Natur sind und darin bestehen können, daß man sich die betreffende Gesamtheit höherer Ordnung als ein *'statistisches Kollektiv'* denkt. Unter einem 'statistischen Kollektiv' verstehe ich – in gewisser Anlehnung an Mises – eine solche *'gut durchmischte'* Gesamtheit, aus welcher die gegebene Gesamtheit auf dem Wege einer 'blinden Auswahl' etwa nach dem Muster von Kugelziehungen aus einer geschlossenen Urne entstanden sein könnte. Unsere 'statistischen Kollektive' bilden also nur einen Sonderfall unter den statistischen Gesamtheiten, und sie können, wie diese, ein objektives Dasein führen, durch bloße Willkür des Forschers gebildet werden, oder sogar rein gedankliche Konstruktionen darstellen; ferner können sie, ebenfalls wie die Gesamtheiten überhaupt, entweder Bestandsmassen sein, die zu einem gewissen Zeitpunkte wirklich in ihrem ganzen Umfange existieren, oder aber Ereignismassen, deren Elemente nur allmählich im Laufe der Beobachtung entstehen und nicht alle gleichzeitig in Erscheinung treten." (Anderson 1935, S. 29f)

Hier setzt die *theoretische* Sozialstatistik also mit dem Gedanken ein, daß die jeweils verfügbaren Daten sich auf zufällig zustande gekommene Teilmengen aus Gesamtheiten höherer Ordnung beziehen. Und wie bei Fisher soll sich das Wort 'Wahrscheinlichkeit' auf relative Häufigkeiten in den höheren Gesamtheiten beziehen.[24]

11. Allerdings versucht Anderson, den Ansatz von Fisher allgemeiner zu fassen. Wir wir gesehen haben, betrachtet Fisher die jeweils gegebenen Daten stets als eine Zufallsstichprobe aus einer hypothetisch unterstellten unendlichen Grundgesamtheit. Anderson betont demgegenüber, daß dies zwar zulässig, jedoch nicht immer sinnvoll sei. Anderson (1935, S. 85f) stellt zunächst den Standpunkt Fishers dar und fährt dann fol-

[24] In einer späteren Arbeit hat Anderson auch vorgeschlagen, daß man sich die höheren Gesamtheiten stets anhand eines Urnenschemas vorstellen könne: „Man darf ferner nicht vergessen, dass der Statistiker, wenigstens der Sozialstatistiker, nur bei der Erhebung in unmittelbare Berührung mit dem zu beobachtenden Objekt kommt, und dabei spielt die Wahrscheinlichkeitstheorie keine Rolle. Bei der weiteren Aufarbeitung und wissenschaftlichen Auswertung des gesammelten Materials, wobei eben die Wahrscheinlichkeitstheorie gebraucht wird, hat der Statistiker es schon allein mit dem *Zählpapier* und dessen Inhalt bzw. mit der *Zählkarte* zu tun. Eben deshalb liegt es auch besonders nahe und ist, soviel ich urteilen kann, mit keinem Verlust an Allgemeingültigkeit der theoretischen Ergebnisse verbunden, wenn man in der wahrscheinlichkeitstheoretisch orientierten, d.h. 'stochastischen' Theorie der Statistik wenigstens im Bereiche der sozialen Massenerscheinungen *immer von einem Urnenschema* ausgeht. Man hat sich eben eine Urne mit Zählkarten als Vertreterin der Gesamtheit höherer Ordnung zu denken, aus der eine Anzahl von Karten so oder anders 'gezogen' wird oder 'selbst herausfällt'. Diese bilden dann die Gesamtheit niederer Ordnung, die 'Stichprobe'." (Anderson 1947, S. 493)

gendermaßen fort:

„Somit führt Fisher jede tatsächlich gegebene Gesamtheit [Stichprobe] sofort auf eine höhere Gesamtheit von *unendlichem* Umfange zurück und hat hierbei nur den 'Fall mit Zurücklegen' im Auge. Sein Standpunkt wird in vielen Fällen zu denselben praktischen Konsequenzen führen wie der unsere, doch gibt es auch solche, wo beide stark differieren können. Stellen wir uns z.B. ein Spiel Karten vor, aus dem eine gewisse Anzahl Karten ohne Zurücklegen entnommen wird. *Unsere* Gesamtheit höherer Ordnung ist eben das ganze Spiel Karten, während man sich nach Fisher hier ebenfalls eine unendliche Population von Karten vorstellen müßte. Oder ein anderes Beispiel, welches für den Statistiker größeres praktisches Interesse besitzt: gegeben sei die Bevölkerung eines gewissen Landes; täglich sterben aus ihre gewisse Personen weg und andere werden hinzugeboren. Es ist sehr wohl möglich, daß hierdurch auch einige durchschnittliche Charakteristiken der Bevölkerung, wie z.B. das Geschlechtsverhältnis, die Geburten- und Sterblichkeitsziffer usw. Veränderungen erleiden. Bei einer jeden solchen Veränderung müssen wir uns, nach Fisher, eine *andere* unendliche Population vorstellen, aus der die gegebene auf dem Wege einer zufälligen Auswahl entstanden sein könnte [. . .]. Im Interesse der größeren Allgemeingültigkeit der theoretischen Konstruktionen und insbesondere im Interesse der Theorie der Stichprobenerhebungen müssen wir darauf bestehen, daß auch der Fall von *endlichen* Gesamtheiten höherer Ordnungen in der statistischen Methodenlehre genügend berücksichtigt wird.“ (Anderson 1935, S. 86)

Auf den ersten Blick erscheint Andersons Idee, den Ansatz von Fisher nur als einen Spezialfall einer allgemeinen Theorie höherer Gesamtheiten, die je nach Anwendungsfall unterschiedlich konzipiert werden können, zu betrachten, durchaus sinnvoll. Eine genauere Überlegung zeigt jedoch, daß die wichtige Unterscheidung zwischen Auswahlverfahren (für die Stichprobentheorie) und hypothetischen Zufallsgeneratoren (für Modelle, die sich auf Prozesse beziehen) dadurch eher noch unklarer gemacht wird. Denn folgt man dem Andersonschen Gedankengang, erscheint es schließlich so (wie schon oben am Beispiel Neymans dargestellt), als ob es sich in beiden Fällen nur um Spezialfälle einer allgemein konzipierbaren Vorstellung handelt; daß man in beiden Fällen die jeweils verfügbaren Daten *als eine Auswahl* aus einer real existierenden oder so denkbaren Gesamtheit höherer Ordnung auffassen könne.

12. Damit wird jedoch der irreführende Gedanke, daß man sich auch Zufallsgeneratoren, die sich ihrer Intention nach auf Entstehungsprozesse sozialer Sachverhalte oder Ereignisse beziehen, als ein Auswahlverfahren aus einer schon existierenden Gesamtheit vorstellen könne, nicht vermieden, sondern verfestigt. Es ist natürlich richtig, daß sich diese Kritik

zunächst gegen Fishers Konzeption hypothetischer unendlicher Populationen richten muß. Aber die Kritik sollte nicht so geführt werden, daß Fishers Ansatz nur als ein Spezialfall einer allgemeinen Stichprobentheorie erscheint. Vielmehr sollte gezeigt werden, daß Fisher nur eine obskure Rhetorik verwendet, die fälschlicherweise das Bild einer Zufallsauswahl aus einer Population suggeriert. Löst man sich von dieser irreführenden Rhetorik, zielt Fishers Ansatz nicht auf eine Stichprobentheorie, sondern macht einen Vorschlag, wie man hypothetische – oder, für Fisher primär, durch Experimente definierbare – Zufallsgeneratoren zur Konstruktion statistischer Modelle verwenden kann.

13. Die in Deutschland durch W. Winkler und O. Anderson begonnene Verwendung des Schemas „Stichprobe–Grundgesamtheit“ als Basiskonzeption auch für eine theoretische Sozialstatistik fand zunächst keine systematische Fortsetzung.[25] Das erste Lehrbuch, das die inzwischen hauptsächlich in England und den USA entwickelten Überlegungen zur theoretischen Statistik systematisch verarbeitet, ist die „Einführung in die mathematische Statistik“ von Leopold Schmetterer (1956, 2. Auflage 1966).[26] In der Einleitung knüpft Schmetterer zunächst an ein traditionelles Statistikverständnis an:

„Man kann somit als ein Charakteristikum der Statistik das Studium der Massenerscheinungen betrachten. Es ist eine Erfahrungstatsache, daß bei Massenerscheinungen Gesetzmäßigkeiten nachgewiesen werden können, die bei Einzelerscheinungen kein Gegenstück haben. Das Studium dieser empirischen Gesetze, die man bei Massenerscheinungen beobachten kann, ist die Aufgabe der

[25] Über die Entwicklung der Statistik während des NS-Regimes erhält einige Hinweise aus dem von Friedrich Burgdörfer herausgegebene zweibändige Sammelwerk „Die Statistik in Deutschland nach dem heutigen Stand“ (1940), das zum 70. Geburtstag von Friedrich Zahn, dem Nachfolger von G. v. Mayr in der Herausgabe des „Allgemeinen statistischen Archivs“, erschien.

[26] Im Vorwort zur ersten Auflage heißt es: „Die mathematische Statistik hat in den letzten 25 Jahren in Verbindung mit der modernen Wahrscheinlichkeitsrechnung einen enormen Aufschwung genommen, der allerdings fast ausschließlich von Gelehrten außerhalb des deutschsprachigen Raumes getragen wurde. Die Ungunst der Zeit brachte es mit sich, daß die deutschsprachigen Länder von dieser Entwicklung ziemlich unberührt blieben. Die Folge davon ist, daß es wohl eine große Anzahl fremdsprachiger und vielfach ausgezeichneter Werke über den Gegenstand der mathematischen Statistik gibt, jedoch kein einziges modernes Lehrbuch in deutscher Sprache, wenn man von der Monographie von A. Linder absieht, welche sich vor allen Dingen an den statistisch arbeitenden Naturwissenschaftler und nur in geringem Ausmaße an den Mathematiker wendet.“ Über den Inhalt seines Buches sagt Schmetterer an der gleichen Stelle: „Die Darstellung umfaßt hauptsächlich jenen Bestand der mathematischen Statistik, den man heute bereits als klassisch bezeichnen könnte und der mit den Namen *Fisher*, *Pearson* und insbesondere *Neyman* verknüpft ist.“

formalen Statistik.[27] Als Aufgabe der mathematischen Statistik, die uns hier ausschließlich beschäftigen wird, könnte man die Aufstellung eines Kalküls bezeichnen, dessen Sätze bei geeigneter Interpretation in Übereinstimmung mit den Feststellungen der formalen Statistik stehen.

Ein Beispiel möge zeigen, wie die empirischen Gesetze der Statistik in roher Fassung zu fast selbstverständlichen Denkgewohnheiten geworden sind. Die Tatsache, daß das Geschlechtsverhältnis bei Neugeborenen immer ungefähr 1 : 1 beträgt, ist für alle Gebiete des menschlichen Lebens von größter Bedeutung, und eine ernste Störung dieses Zustandes hätte für unser Dasein gewichtige Folgen. Es wird aber niemandem einfallen, sich diesbezüglich Befürchtungen hinzugeben, nur, weil er vielleicht soeben eine Familie kennengelernt hat, wo sämtliche (z.B. 5) Kinder Knaben sind. Man weiß eben aus Erfahrung, daß bei Betrachtung nur eines Falles auch große Abweichungen von den Verhältnissen, wie sie in der Masse herrschen, vorkommen können, ohne daß dies im einzelnen begründet werden kann. Man spricht in diesem Zusammenhang von zufälligen Abweichungen von den Verhältnissen in der Gesamtheit und meint damit, daß keine Veranlassung bestehe, anzunehmen, daß diese Abweichungen von einer bestimmten gemeinsamen Ursache hervorgerufen sind. So hat man etwa keine zufälligen Abweichungen vor sich, wenn man auf der Straße das Geschlechtsverhältnis der Passanten beobachtet und feststellt, daß es in der Nähe einer Knabenschule vor 8 Uhr morgens stark gestört ist. Die Ausschaltung solcher systematischer Abweichungen sowie die Beurteilung von Abweichungen als zufällig oder systematisch gehören zu den wichtigsten Aufgaben des Statistikers.“ (Schmetterer 1966, S. 21)

Das Beispiel deutet an, daß sich Schmetterer gedanklich auf Prozesse beziehen möchte, durch die Sachverhalte und Ereignisse entstehen. Solche Prozesse werden als „zufällig“ bezeichnet, wenn und insoweit wir nicht wissen, wie sie zustande kommen und ablaufen. Also könnte man vermuten, daß sich Schmetterer in der Begründung seiner Begriffsbildungen auf solche Prozesse bezieht. Das ist jedoch nicht der Fall. Nachdem er im ersten Kapitel seines Buches die Grundbegriffe der Wahrscheinlichkeitsrechnung im wesentlichen rein formal (axiomatisch) eingeführt hat, wobei zur Motivation die Idee eines „Zufallsexperiments“ verwendet wird, beginnt das zweite Kapitel mit einer Einführung in die „elementare Stichprobentheorie“:

„Anläßlich der Einführung des Wahrscheinlichkeitsbegriffes haben wir folgende Situation beschrieben: Es liege eine Gesamtheit von Beobachtungen vor, die sich auf gewisse Merkmale beziehen. Diese Gesamtheit wird als unendlich

[27][Schmetterer verweist an dieser Stelle auf die zweite Auflage von Winklers „Grundriß der Statistik“ (1931), die 1947 erschienen ist.]

angesehen, und zwar in dem Sinne, daß nach einer gleichbleibenden Vorschrift die Beobachtungen stets reproduzierbar sind, also z.B. eine unendlich ausgedehnte Serie von Würfelversuchen. Aus dieser Gesamtheit wählt man ‘zufällig’ eine Reihe von Beobachtungen aus. Ist deren Anzahl hinlänglich groß, dann weichen die relativen Häufigkeiten von Ereignissen, die sich auf die unter Beobachtung stehenden Merkmale beziehen, im allgemeinen nur unbedeutend von einem stets konstanten Wert ab, den wir als empirische Wahrscheinlichkeit bezeichnet haben.[28] Es ist nicht leicht, empirische Kriterien dafür anzugeben, wann eine Auswahl aus einer Gesamtheit als zufällig angesehen werden kann. Oft begnügt man sich mit der etwas vagen Formulierung, daß eine Zufallsauswahl realisiert ist, wenn kein Grund vorliegt, der eine bevorzugte Auswahl irgendeiner Beobachtung erkennen ließe. Man zieht in diesem Zusammenhang vielfach das ‘Urnenschema’ heran. Die Urne oder – besser gesagt – deren Inhalt (z.B. gleichgestaltete Kugeln) repräsentiert die Gesamtheit, und nun werden dieser Urne Kugeln entnommen, wobei stets auf ‘gute Durchmischung’ des Inhalts zu achten ist. Die gezogenen Kugeln sieht man als Zufallsauswahl aus der Urne an. [. . .]

Ziel solcher zufälliger Auswahlen aus der Gesamtheit ist es, über die Struktur der Gesamtheit etwas auszusagen, insbesondere über die Größe der mit ihr verknüpften empirischen Wahrscheinlichkeiten. Man läßt sich hierbei etwa von der Vorstellung leiten, daß man die empirische Wahrscheinlichkeit ‘genau bestimmen könnte’, wenn man unendlich viele Versuche macht. [. . .]

Es hat sich folgende Terminologie eingebürgert: Die (hypothetische) unendliche Gesamtheit der möglichen Beobachtungen nennt man die *Grundgesamtheit* und die aus dieser Menge (zufällig) herausgegriffenen Beobachtungsergebnisse eine *Stichprobe* aus dieser Grundgesamtheit. Die Anzahl der in der Stichprobe enthaltenen Beobachtungen bezeichnet man als deren *Umfang*. Die Vorstellung einer unendlichen Grundgesamtheit stellt eine Idealisierung der tatsächlichen Verhältnisse dar, auch wenn man nur an eine ‘potentiell unendliche’ Gesamtheit im Sinne der unbegrenzten Reproduzierbarkeit der Zufallsversuche denkt. Der Praktiker sieht jede Gesamtheit, die im Verhältnis zum Umfang der entnommenen Stichprobe ‘sehr groß’ ist, als unendlich an.“ (Schmetterer 1966, S. 147f)

Mit geringfügigen Varianten findet man eine Darstellung dieser Art in den meisten Einführungen in die theoretische Statistik, die seit den 1950er Jahren veröffentlich worden sind. Eine Ausnahme bildet die Spezialliteratur zur Stichprobentheorie, die sich explizit mit in der Realität fixierbaren und infolgedessen endlichen Gesamtheiten beschäfigt. Darauf

[28][Schmetterer (S. 23) folgt einer Variante der frequentistischen Deutung des Wahrscheinlichkeitsbegriffs, die sich Wahrscheinlichkeiten als fiktive Grenzwerte für relative Häufigkeiten bei der wiederholten Anwendung eines Zufallsgenerators vorstellt.]

werden wir in Abschnitt 4.4 genauer eingehen. Hier ist dagegen bemerkenswert, wie in den Ausführungen Schmetterers jede Erinnerung an eine Unterscheidung zwischen einerseits Prozessen, oder auch Experimenten, durch die Sachverhalte und Ereignisse zustande kommen, und andererseits irgendwie *gegebenen* Gesamtheiten, aus denen Stichproben *ausgewählt* werden können, nicht nur vermieden wird, sondern – wie schon oben am Beispiel von Neyman dargestellt wurde – als für die Begriffsbildungen einer theoretischen Statistik überflüssig erklärt wird.

4.3.2 Bemerkungen zu einer Verwechslung

1. Die Vorstellung, daß man auch Zufallsgeneratoren, die sich ihrer Intention nach auf Prozesse beziehen, als Auswahlverfahren auffassen könne, führt nicht nur zu Unklarheiten, sondern gelegentlich auch zu irreführenden Verwechslungen. Auf Unklarheiten, die daraus resultieren, daß hypothetische Populationen indefinite Begriffsbildungen implizieren, wurde schon mehrfach hingewiesen. Um zu illustrieren, wie daraus auch begriffliche Verwechslungen entstehen können, beziehe ich mich auf ein einflußreiches Lehrbuch der Statistik von G. U. Yule und M. G. Kendall (1950):[29]

> "By the logical extension of the idea of a population of concrete objects, which we shall call an *existent population*, we are able to construct the idea of a *hypothetical population*. Consider the throws of a die. Each throw will be regarded as an individual. There is an infinite number of throws which can be made with the die, provided that is does not wear out. Let us then define as our population of discussion all the *possible* throws of the die. [...] Such a population is called a *hypothetical* population. We may define it formally as the aggregate of all the conceivable ways in which a specified event can happen." (Yule und Kendall 1950, S. 367)

Das Problem besteht nicht nur darin, daß es die „Menge aller möglichen Würfe mit einem Würfel“ nicht gibt; so wenig wie die „Menge aller potentiellen Kunden eines Versicherungsunternehmens“ oder die „Menge aller Fußballspiele“, die bereits stattgefunden haben oder in der Zukunft vielleicht noch stattfinden werden. Problematischer ist, daß infolgedessen das Reden über – zum Beispiel – Würfelspiele ganz unklar wird. Natürlich kann man einen Würfel verwenden, um mit ihm „beliebig oft“

[29] Ich zitiere aus der 14. Auflage, die erste Auflage erschien 1911. George U. Yule lebte 1871 – 1951, Maurice G. Kendall 1907 – 1983; biographische Angabe finden sich bei Johnson und Kotz 1997, S. 130ff, 168ff.

bestimmte Ereignisse zu erzeugen. Man kann in diesem Fall auch von einem „aggregate of all the conceivable ways in which a specified event can happen“ reden. Aber *diese* Menge hat genau sechs Elemente, entsprechend den 6 möglichen Ergebnissen eines Würfelwurfs; sie ist weder hypothetisch noch indefinit, und sie darf natürlich nicht mit der indefiniten „Menge aller Würfe des Würfels“ verwechselt werden.

2. Die Menge der möglichen Ergebnisse bei einem Würfelspiel hat tatsächlich mit der Idee einer statistischen Gesamtheit überhaupt nichts zu tun, sondern fixiert den Merkmalsraum einer statistischen Variablen.[30] Angenommen, wir werfen den Würfel 100 Mal und erzeugen dadurch die Situationen $\omega_1, \dots, \omega_{100}$; dann hat man eine Gesamtheit

$$\Omega = \{\omega_1, \dots, \omega_{100}\}$$

und man kann eine statistische Variable

$$X : \Omega \longrightarrow \tilde{X} = \{1, 2, 3, 4, 5, 6\}$$

definieren, die jeder Situation ω_i die in ihr erzielte Augenzahl $X(\omega_i)$ zuordnet. Man hat es dann mit einer definiten Gesamtheit Ω zu tun und außerdem mit einem Merkmalsraum $\tilde{X}$ für eine statistische Variable X.

3. In diesem Beispiel besteht sowohl die Gesamtheit Ω als auch der Merkmalsraum $\tilde{X}$ aus einer endlichen Menge von Elementen. Indem man Merkmalsräume in Zahlenmengen einbettet, kann man auch Merkmalsräume konzipieren, die in einem mathematisch explizierbaren Sinn unendlich viele Merkmalswerte enthalten. Demgegenüber sollte jedoch betont werden, daß definite statistische Gesamtheiten stets endlich sind; denn Mengen von Dingen, Menschen oder Situationen, die in unserer Erfahrungswelt fixiert werden können, haben immer nur eine endliche Anzahl von Mitgliedern. Dagegen sind „unendliche Populationen“ schon begrifflich nicht explizierbar und suggerieren Verwechslungen mit Merkmalsräumen. Zur Illustration kann wieder das Lehrbuch von Yule und Kendall dienen:

> "By a *finite population* we shall mean a population which contains a finite number of members. Such, for instance, is the population of inhabitants of Great Britain and the population of books in the British Museum. Similarly, by an *infinite* population we shall mean a population containing an infinite number of members. Such, for instance, is the population of pressures at various points in the atmosphere, or the population of possible sizes of the wheat

[30] Oder einer Zufallsvariablen, die verwendet werden kann, um einen Zufallsgenerator formal zu repräsentieren.

crop, for, although there are limits to the size, the actual tonnage can take any numerical value within those limits." (Yule und Kendall 1950, S. 367)

Spricht man von den möglichen Werten, die das Gewicht einer Weizenernte annehmen kann, bezieht man sich offenbar auf einen Merkmalsraum, nicht auf eine statistische Gesamtheit. Und da die Autoren von ihrer Weizenernte im Singular sprechen, ist auch ganz unklar, wie man in diesem Fall eine statistische Gesamtheit ausfindig machen kann.

4. In der statistischen Literatur kommt es oft zu einer Verwechslung zwischen einerseits statistischen Gesamtheiten und andererseits Merkmalsräumen für statistische oder Zufallsvariablen. In gewisser Weise ist die Verwechslung in einem Grundbegriff der Statistik, dem Begriff eines Stichprobenraums, schon angelegt. Damit werden wir uns in Abschnitt 4.4.3 beschäftigen.

4.3.3 Populationen und Variabilität

1. Statistische Methoden werden oft mit dem Hinweis motiviert, daß man mit ihrer Hilfe die Variabilität von Merkmalen untersuchen könne. In gewisser Weise ist das unmittelbar einleuchtend: die unterschiedlichen Werte einer statistischen Variablen zeigen, wie Merkmalswerte in einer Gesamtheit variieren; und durch die Bildung einer Häufigkeitsverteilung erhält man eine Beschreibung dieser Variabilität. Durch ein Interesse, Einsichten in solche Variabilität zu gewinnen, kann man z.B. Fechners Kollektivmaßlehre verstehen. Ein entsprechendes Interesse läßt sich in vielen Anwendungen statistischer Methoden feststellen. J. Neyman hat dazu folgende allgemeine Bemerkung gemacht:

"With an overwhelming frequency, modern science is concerned with categories of objects which, while satisfying a common definition, vary considerably in those characteristics that are relevant to the study performed. As a result, any attempt at a mathematical treatment must involve assumptions that some given characteristic X of the objects studied is a random variable with some particular, known or unknown, distribution." (Neyman 1971, S. 1)

Hier kommt auch zum Ausdruck, daß empirisch feststellbare Variabilität die Begründung liefern soll, daß eine Anwendung von Begriffsbildungen und Vorstellungen der Wahrscheinlichkeitstheorie sinnvoll sein kann. Das war ja auch schon Fechners Gedanke: daß die Elemente einer statistischen Gesamtheit „zufällig variieren".

2. Dann entsteht jedoch eine Frage: Wodurch kommt die Variabilität zustande? Denn ein Hinweis auf Zufälligkeiten ist für sich genommen

nichtssagend. Um in irgendeiner Weise weiter zu kommen, muß man sich gedanklich auf Prozesse beziehen, durch die Dinge, Menschen oder Situationen jeweils bestimmte Eigenschaften bekommen. Schon die Fragestellung wird indessen unklar, wenn man sich in der theoretischen Statistik auf unspezifische Weise des Schemas „Stichprobe–Grundgesamtheit" bedient. Es entsteht der Eindruck, daß das Entstehen von Variabilität grundsätzlich mit der Vorstellung eines Auswahlverfahrens verbunden werden kann. Aber dort, wo dieser Gedanke einen Sinn macht, nämlich bei der Auswahl von Elementen aus einer vorgegebenen Gesamtheit, hat man es gerade nicht mit einem Prozeß zu tun, durch den Dinge, Menschen oder Situationen jeweils bestimmte Eigenschaften bekommen.

3. Zur Illustration der Unklarheiten, die durch die Verbreitung des Redens von Stichproben und Populationen in der theoretischen Statistik entstanden sind, sei auf eines der ersten Lehrbücher hingewiesen, die die neue Rhetorik übernommen haben: „Introductory Statistics" von M. H. Quenouille (1950). Das Buch beginnt mit folgenden Bemerkungen:

"*Sampling* – If a series of measurements is made, it is implied that there is a variability in the results which necessitates taking more than one measurement. This variability may be due to the measurements being taken on different objects or under different conditions, as with a set of children's weights, or due to errors or changes in the method of measurement, as might occur in the evaluation of a physical constant such as g, or due to a combination of these, as with plot yields in an agricultural lay-out. If a set of children's weights is taken, this is commonly described as a 'sample' from a particular population, unless every child in the population is measured. This terminology is extended in statistical work so that a set of measurements of the physical constant g is said to be a sample from the population of possible measurements of g, and plot yields are said to be samples from the population of plot yields grown on similar soils under similar conditions. In this wider sense the population does not necessarily exist. For example, it is impossible to measure the yield of a field-plot more than once under the same conditions, since measurements will be affected by the drain on the ressources of the land caused by previous measurements. It is nevertheless useful to think of a plot yield as a sample from the population of possible plot yields." (Quenouille 1950, S. 1)

Diese Ausführungen machen gut deutlich, welche Unklarheiten entstehen, wenn man sich das Entstehen von Variabilität stets als Folge von „sampling" aus einer Population vorstellen möchte.[31] Es ist auch bemer-

[31]Dagegen sind folgende Unterscheidungen von A. L. Bowley (1926, S. 248) erheblich klarer: „Variation is a general law of nature and is found in most human affairs, so that large scale observations usually lead to frequency groups. Four classes can

kenswert, wie Quenouille schließlich bei einer Population endet, die es nicht gibt. Aber wie kann man aus einer nicht-existierenden Gesamtheit eine Stichprobe ziehen? Etwas später heißt es:

"It was explained at the outset that a sample set of measurements is normally taken in order to draw conclusions about the population." (Quenouille 1950, S. 8)

Aber dann kann man gleich weiter fragen: Welchen Sinn könnte es machen, Einsichten in eine nicht-existierende Gesamtheit gewinnen zu wollen? Der Hinweis auf „convenience“ liefert kaum eine Begründung.[32]

4. Halten wir demgegenüber fest: die Sozialstatistik hat es zunächst immer nur mit in der Erfahrungswelt fixierbaren definiten Gesamtheiten zu tun. Auf sie bezieht sich die sozialstatistische Datengewinnung, sei es in der Form von Stichproben oder durch Totalerhebungen. Diese Tatsache schließt es natürlich nicht aus, sich gedanklich auf Prozesse zu beziehen, durch die die Mitglieder einer sozialstatistischen Gesamtheit enstanden sind bzw. jeweils bestimmte Eigenschaften bekommen haben. Hier setzt die theoretische Sozialstatistik an, indem sie diese Prozesse durch Zufallsgeneratoren reflektierbar zu machen versucht. Fishers Idee hypothetischer unendlicher Populationen kann als eine nur in der Rhetorik unterschiedliche Konzeption dieses Grundgedankens einer theoretischen Statistik verstanden werden. Die Rhetorik ist indessen nicht nur deshalb unglücklich, weil sie unklar ist und Verwechslungen nahelegt. Sie legt auch den irreführenden Gedanken nahe, daß es gewissermaßen zwei Arten von Populationen gibt: einerseits die empirisch bestimmbaren Gesamtheiten der Sozialstatistik und andererseits gedanklich konstruierbare hypothetische Populationen. Infolgedessen scheint eine sehr merkwürdige Fragestellung zu resultieren: Kann man sich die empirischen

be distinguished: (a) where every member of a group has been measured, e.g. the wages of every adult male working in a trade; (b) observations of samples selected from a group, e.g. the number of children in each of 1,000 families chosen in a town where there are 50,000 families, or the measurement of leaves of a tree of a particular kind; (c) repeated measurements of a physical quantity (e.g. of the declination of a star) where the variations are due to instrumental *errors*; (d) the mathematical probabilities of various numbers of successes (e.g. the chances of obtaining 1, 2, 3, ... heads when 50 coins are tossed) or the frequencies of events whose magnitude depends on an unknown complex of causes.“ Nur im letzten Fall geht es um Prozesse, durch die Variabilität in der Realität entsteht.

[32] Der Autor wiederholt jedoch nur dieses Argument: „At the beginning of the previous chapter it was explained that it is convenient to regard a set of measurements as a sample from a population of possible measurements, although the population may not exist.“ (S. 21)

Gesamtheiten der Sozialstatistik als Teilmengen – oder mehr noch: als Zufallsstichproben – aus hypothetischen Populationen vorstellen?

5. Daß die Frage in eine Sackgasse führt, wird besonders deutlich, wenn man sich überlegt, wie der Idee einer hypothetischen Population in der Sozialstatistik ein Sinn gegeben werden könnte. Zur Verdeutlichung beziehe ich mich auf Überlegungen von M. J. Hagood in ihrem Buch „Statistics for Sociologists“ (1941).[33]

"Let us consider the situation in sociology requiring the different methods of statistics. [...] If a survey has as its purpose the description of a set of conditions for a certain area at a certain time, usually with the implied purpose of measuring them in order to plan some program of reform to alter them, then descriptive statistics may be adequate for the analysis of such quantitative data as are collected in this type of project.

Practical sampling situations requiring inductive statistics. If, on the other hand, either of the following situations is the case, inductive statistics is indicated. The first situation is the same as that described above for the survey, except that the area is too great for every unit to be counted or measured, and only a sample of units is observed. Here one uses inductive statistics to form estimates of various summarizing measures for the whole area (universe) with accompanying measures of unreliability of the estimates. This function of inductive statistics is simply a substitute for descriptive statistics when all units cannot be surveyed.

Hypothetical sampling situations requiring inductive statistics. The other use of inductive statistics is not so easy to state or explain. In this second situation, from observations made either from a sample or from complete survey of some limited universe, we generalize to a hypothetical universe which is difficult to define. It is the universe of all possible samples (which may be limited universes) which could have been produced under similar conditions of time, place, culture, and other relevant factors." (Hagood 1941, S. 65f)

Die zweite Hälfte des Gedankengangs, in der an Fishers hypothetische Populationen angeknüpft wird, ist offenbar merkwürdig und schwer verständlich. Die Autorin versucht, der Idee zunächst dadurch einen Sinn zu geben, daß sie sich auf Experimente bezieht. Dann geht es so weiter:

"Sociologists, especially those in the field of social psychology, are developing experimental techniques for attack on certain of their problems where they have the advantage of more or less control over relevant factors. In general,

[33] Ich zitiere nach einem Auszug aus diesem Buch, der in einem Sammelband von Morrison und Henkel (1970) abgedruckt wurde. Im wesentlichen gleiche Ausführungen finden sich auch noch in einer zweiten Ausgabe des Buches, vgl. Hagood und Price, 1952.

however, this is not the case, and we are concerned here with the situation where the sociologist can only observe at one time measures or attributes of a series of units. We are concerned especially with the case where the observations are made for all of a series of units as of a certain time, as, for instance, on all of the rural counties in the United States for 1930. Let us suppose that for 1930 one observes a negative correlation coefficient of value r between fertility ratio and level of living for all rural counties in the United States. With the same formulas and procedures used by the experimentalists, the sociologist tests the significance of his observed r and finds it 'significantly' different from zero. What does this test mean to the sociologist? In any test of significance there is a testing of some hypothesis about a universe from which the set of observations (a limited universe itself in this case) may be considered a random sample. That is, the logical structure of a superuniverse and the variation expected in random samples from it is the same for the observer sociologist as for the experimentalist. Imagining any experimental counterpart of the logical model is a more difficult matter, however. It is easy enough for the experimentalist to imagine repeated experiments under identical conditions, whether or not he can actually perfect his technique to the degree that he can reproduce conditions identically. His universe of possibilities can therefore be put into meaningful terms; it can at least be imagined, even if it cannot actually be produced. It is not so easy for the sociologist to imagine a set of observations repeated under conditions identical with those of one date. The fact of change in social and cultural phenomena renders unrealistic any conception of identical repetition of the complex of factors conditioning characteristics such as fertility and level of living. The concept of the universe of possibilities – that is all possible sets of measures on fertility and level of living that could possibly be produced in the thousand rural counties of the United States under conditions exactly similar to those of 1930 – the concept has neither a realistic counterpart nor a readily imaginable counterpart. To what, then, does the variation expected from random sampling from such a universe of possibilities correspond? Only a feat of imagination involving an infinite prolongation of a present moment, where conditioning factors remain the same but 'chance' factors continue to produce random variation can supply the answer." (Hagood 1941, S. 70f)

Die Ausführungen zeigen, in welche Schwierigkeiten man gerät, wenn man versucht, die Idee einer hypothetischen Superpopulation ernst zu nehmen. Die Vorstellung, daß das, was an einem bestimmten Ort und zu einer bestimmten Zeit geschieht, unter gleichen Bedingungen beliebig oft wiederholt werden könnte, ist schon als ein Gedankenexperiment obskur; ganz abgesehen von dem Einwand, daß vermutlich das gleiche herauskommen wird, wenn sich etwas unter *identical conditions* wiederholt. Ein Ausweg aus den logischen Schwierigkeiten ist nur möglich, wenn man sich nicht auf hypothetische Populationen, sondern auf Prozesse bezieht, durch die Sachverhalte und Ereignisse zustandegekommen sein könnten. Dies ermöglicht auch eine sinnvolle Reformulierung der Fragestellung: Für welche Erkenntnisinteressen kann es sinnvoll sein, solche Prozesse als Realisationen von Zufallsgeneratoren zu konzipieren?

4.3.4 Sozial- vs. Naturwissenschaft?

1. Es wurde schon darauf hingewiesen, daß in der Entwicklung der Sozialstatistik in Deutschland im Zeitraum bis etwa 1950 nur einige wenige Autoren sich mit Fishers Idee hypothetischer Superpopulationen auseinandergesetzt haben. Soweit sich in diesem Zeitraum eine Kritik an den Ansätzen zu einer theoretischen Sozialstatistik artikuliert hat, bezog sie sich im Kern darauf, daß die durch den Begriff eines Zufallsgenerators vorausgesetzte Wiederholbarkeit bei sozialen Sachverhalten und Ereignissen, die sich im Rahmen eines historischen Prozesses entwickeln, nicht gegeben sei. Ein wichtiger Vertreter dieser Kritik war Paul Flaskämper. Der Gedankengang wird z.B. in folgender Passage in seinem „Beitrag zu einer Theorie der statistischen Massen“ (1928) deutlich:

„Wenn der Naturwissenschaftler mit der 'statistischen Methode' und mit statistischen Massen arbeitet, so hat er einen ganz bestimmten Fall von Massen im Auge, der in den Sozialwissenschaften nur ganz selten verwirklicht ist, an dem aber – und das ist ein unerträglicher Widerspruch der meisten statistischen Theorien – der Begriff der statistischen Massen überhaupt gemeinhin orientiert ist. Er hat es nämlich zu tun mit gleichartigen, aber variablen Fällen. Gleichartig sind die Fälle solcher Massen insofern, als sie bestimmt werden von einem einheitlichen, allen gemeinsamen Ursachenkomplex, variabel insofern, als in jedem einzelnen Falle zu den gemeinsamen Ursachen (oder allgemeinen oder konstanten Ursachen, wie Quetelet sagte, oder gesamtheitlichen Ursachen, wie neuerdings Winkler sagt) hinzukommen individuelle oder störende oder zufällige Ursachen, die in den einzelnen Individuen den gesamtheitlichen Ursachenkomplex überdecken können. Wenn es sich nun um echte zufällige Ursachen handelt, so werden sie in ihrer Wirkung ebensooft in der einen wie der anderen Richtung von den gesamtheitlichen abweichen und sich bei einer genügend großen Zahl von Fällen gegenseitig aufheben.

Das ist in großen Zügen der Sinn des Gesetzes der großen Zahlen, dessen eingehende Begründung in das Gebiet der Mathematik, im besonderen der Wahrscheinlichkeitsrechnung gehört. [...] In der Tat sind auch in den Naturwissenschaften, in der Physik und vor allem in der Biologie (Variationsstatistik), mit dieser Methode die schönsten Erfolge erzielt worden, die auf anderem Wege nicht zu gewinnen gewesen wären. Schon von vornherein aber ist zu erwarten, daß derartige Massenerscheinungen auf die die wahrscheinlichkeitstheoretischen Gesichtspunkte des Gesetzes der großen Zahlen angewendet werden können, im Reiche der sozialen Erscheinungen sehr viel seltener sind. Viel zu stark sind hier die individuellen Ursachen gegenüber den gesamtheitlichen, die überdies in den meisten Fällen überhaupt nicht eindeutig vorhanden sind.“ (Flaskämper 1928, S. 552f)

Mit dem Reden von „Ursachen“ und insbesondere der Unterscheidung zwischen „allgemeinen“ und „zufälligen Ursachen“ werden wir uns erst in Abschnitt 3 etwas genauer beschäftigen. Hier soll zunächst nur deshalb auf Flaskämper hingewiesen werden, weil er seine Vorbehalte gegenüber dem Ansatz der theoretischen Sozialstatistik mit einer Unterscheidung zwischen Sozial- und Naturwissenschaften verknüpft. Der Gedanke läuft darauf hinaus, daß eine Deutung des Zustandekommens *sozialer* Sachverhalte und Ereignisse durch Zufallsgeneratoren – als Voraussetzung für die Anwendbarkeit eines „Gesetzes des großen Zahlen“ – in den meisten Fällen nicht sinnvoll sei, so daß sich sozialwissenschaftliche Statistik im wesentlichen auf deskriptive Aufgaben beschränken sollte.[34]

2. Überlegungen dieser Art wurden auch von anderen Autoren aufgegriffen, u.a. von A. Blind (1953) und H. Hartwig (1956).[35] Im folgenden soll der Gedankengang anhand einer Arbeit von Günter Menges (1959) besprochen werden. Menges’ Buch trägt den Titel „Stichproben aus endlichen Gesamtheiten“ (1959) und ist hier auch deshalb interessant, weil es sich ausführlich mit Fishers Konzeption hypothetischer unendlicher Gesamtheiten beschäftigt. Soweit ich beurteilen kann, handelt es sich auch um die erste deutschsprachige Auseinandersetzung mit der seit Ende der 1940er Jahre entstandenen englischsprachigen Literatur zur Stichprobentheorie für reale Gesamtheiten.

3. Menges steckt zunächst, wie viele andere Statistikbücher, den gedanklichen Rahmen so ab:

[34] Weitere Überlegungen hierzu finden sich u.a. bei Flaskämper, 1927.

[35] R. Wagenführ bemerkt im ersten Band seiner „Wirtschafts- und Sozialstatistik“ (1970, S. 22): „Der italienische Statistiker Benedetto Barberi, der jahrelang das statistische Zentralamt Italiens geleitet hat und heute Professor für Statistik an der Universität Rom ist, führt in diesem Zusammenhang eine doppelte Unterscheidung ein: die nach Wiederkehr oder Nichtwiederkehr der Phänomene und die nach ihrer Reproduzierbarkeit oder Nichtreproduzierbarkeit. Die Phänomene des Wirtschafts- und Soziallebens sind sowohl nicht-wiederkehrend als auch nicht-wiederholbar; dies unterscheidet sie von den Phänomenen, mit denen z.B. viele Zweige der Naturwissenschaft zu tun haben. Es leuchtet ein, daß diese doppelte Besonderheit auf Fragestellung, Methode und Arbeitsergebnis des Wirtschafts- und Sozialstatistikers nicht ohne Einfluß bleiben kann.“

„*Der direkte Schluß.* Seine Aufgabe im Rahmen der Stichprobentheorie besteht in der Bestimmung von Zufallsverteilungen, unter der Voraussetzung, daß die Ausgangsverteilung sowie die Art der Entnahme der Stichprobe bekannt sind.“ (Menges 1959, S. 15)

„*Der Rückschluß.* Seine Aufgabe im Rahmen der Stichprobentheorie besteht darin, aus der Stichprobe die unbekannten Parameter der Ausgangsverteilung zu schätzen.“ (S. 20)

Im Mittelpunkt seiner Überlegungen steht vor allem die Frage, wie der „Rückschluß“ zu verstehen ist. Dies hängt, so Menges, davon ab, wie man die Gesamtheit konzipiert, für die man Aussagen machen möchte. Menges schlägt folgende Unterscheidung vor:

„Als realitätstranszendent bezeichnen wir den Schluß von der Stichprobe auf eine unendlich große, hypothetische oder nur gedachte Grundgesamtheit, eine solche also, bei der – in der Symbolik des ersten Kapitels – $N \to \infty$. Als realitätsimmanent dagegen bezeichnen wir den Schluß von der Stichprobe auf eine endliche, empirische, realiter vorhandene Grundgesamtheit.“ (Menges 1959, S. 38)

4. Zur Erläuterung der ersten Variante, des realitätstranszendenten Rückschlusses, knüpft Menges an die Konzeption Fishers an. Er skizziert zunächst einige Grundgedanken Fishers und stellt dann folgende Frage:

„Was aber hat es mit der ‘Reduktion der Daten’, der ‘hypothetischen unendlichen Population’, auf sich; welches sind die Prinzipien, die uns erlauben, dieses Gebäude der stochastischen Metrik auf die Realität zu übertragen?“ (Menges 1959, S. 45)

Menges’ Überlegungen gehen dann so weiter:

„Zur Beantwortung dieser Fragen gehen wir von einem einfachen Beispiel aus. Wir stellen uns vor, wir hätten die Absicht, die Wuchshöhe des gemeinen Habichtkrautes (Hieracium pilosella) zu bestimmen.

Auch ohne eingehendere Prüfung vermögen wir zu fixieren, daß die Wuchshöhe x eine Zufallsvariable darstellt, d.h. daß die Wuchshöhe des gemeinen Habichtskrautes ‘zufällig’ verschiedene Werte $x_1, x_2, \dots, x_n$ annehmen kann. Wir stehen also vor einer *statistischen* Frage.

Alsdann messen wir die Wuchshöhe von – sagen wir – 12 Kräutern und erhalten etwa die folgenden Werte x_i mit den zugehörigen Häufigkeiten p_i. [Es folgt eine Tabelle mit Daten, die wir hier auslassen.]

Und nun stellen wir die Frage: Welche Folgerungen können wir aus diesem statistischen Befund bezüglich der ganzen species ziehen? Zur Beantwortung genau dieser Frage hat uns nun Fisher und seine Schule jenes umfangreiche Arsenal an Methoden zur Verfügung gestellt. [...]

Wir definieren unsere 12 ausgewählten Pflanzen als eine Zufallsstichprobe aus der unendlich groß gedachten hypothetischen Ausgangsverteilung *aller* Pflanzen von der Gattung des gemeinen Habichtkrautes, die gegenwärtig vorhanden sind, in der Vergangenheit einmal vorhanden waren und in Zukunft vorhanden sein werden; oder: auf das *Gesetz* des gemeinen Habichtkrautes.“ (Menges 1959, S. 45f)

Menges zeigt dann, wie man mit seinen Beispieldaten die durchnittliche Wuchshöhe des gemeinen Habichkrautes in der Menge *aller* Pflanzen dieser Art, die es jemals gegeben hat oder noch geben wird, schätzen und den „Schätzfehler“ bestimmen kann.

5. Menges setzt seine Überlegungen mit der Frage fort, unter welchen Voraussetzungen das Verfahren einen Sinn machen könnte:

„Woher nimmt man aber die Berechtigung [. . .], eine kleine Gesamtheit von 12 Pflanzen als Zufallsstichprobe aus einer unendlich großen, hypothetischen Population zu deklarieren?

In der Tat ist die Fisher'sche Schule die Antwort auf diese Frage schuldig geblieben, was ihr die Kritik einiger Wahrscheinlichkeitstheoretiker, besonders diejenige von R. von Mises, eingetragen hat. Der Fisher'sche Ansatz läßt sich aber durchaus legitimieren, wenn man nur von folgender Überlegung ausgeht:

Wenn wir Naturzusammenhänge zu erkennen trachten, so intendieren wir stets die *Erkenntnis allgemeiner Gesetze*; ein idiographisches (einzelbeschreibendes) Wissen von Naturzusammenhängen hat keine Aussicht, jemals zu einer wissenschaftlichen Erkenntnis geführt zu werden.

Die Erkenntnis von Naturzusammenhängen kann aber nun auf zweierlei Weise vor sich gehen, und zwar je nachdem, ob gewissen realen Bedingungen eine bestimmte reale Folge zugeordnet ist oder nicht. Im Falle, daß gewissen realen Bedingungen eine eindeutige reale Folge zugeordnet ist, denken wir die Folge unter dem Prinzip der Kausalität. [. . .]

Im Falle jedoch, daß gewissen realen Bedingungen *keine* eindeutige reale Folge zugeordnet ist, wie im Beispiel der Wuchshöhe des gemeinen Habichtkrautes, können wir von einer kausalen Abfolge nicht mehr reden, und es ist deshalb nur konsequent, wenn Hartwig (1956) an die unsichere Stelle des 'Gesetzes der großen Zahlen' oder der 'Cournot'schen Brücke' ein dem Kausalprinzip analoges *'Ätialprinzip'* gesetzt hat. Denn freilich können wir im Falle, daß gewissen realen Bedingungen keine reale Folge eindeutig zugeordnet ist, die 'möglichen' Folgen mit den ihnen entsprechenden Wahrscheinlichkeiten ebenfalls nur unter einem Prinzip und hier also dem der *allgemeinen Bedingungen* und der 'möglichen' Folgen denken. Genauso, wie wir im Falle der Kausalität mit einer Veränderung der Bedingungen auch eine Veränderung der eindeutigen realen Folgen herbeiführen, führen wir im Falle der Ätialität mit einer Veränderung der allgemeinen Bedingungen eine Veränderung der möglichen realen Folgen

und ihrer Wahrscheinlichkeiten herbei.

Bezogen auf den Fisher'schen Ansatz bedeutet dies zunächst, daß wenn wir gewiß sein dürfen, daß die allgemeinen Bedingungen konstant sind, dann mit der gleichen Berechtigung wie den Kausalschluß im Falle konstanter Bedingungen hier bei den allgemeinen Bedingungen den Wahrscheinlichkeitsschluß auch durchführen können.“ (Menges 1959, S. 47ff)

Die Ausführungen erzeugen eine Reihe von Fragen. Kann man sich sinnvoll auf eine Menge aller Exemplare des gemeinen Habichtkrauts beziehen, einschließlich derjenigen, die „erst in Zukunft vorhanden sein werden“? Aber auch wenn man das tun könnte, mit welcher Berechtigung kann man annehmen, daß die durchschnittliche Wuchshöhe auch in der Zukunft so bleiben wird, wie sie sich anhand heute beobachtbarer Exemplare feststellen läßt? Menges beruft sich auf ein Erkenntnisinteresse: man strebt eine „Erkenntnis allgemeiner Gesetze“ an. Aber wie kann ein Erkenntnisinteresse garantieren, daß sich die durchschnittliche Wuchshöhe des gemeinen Habichtkrauts nicht ändern wird? Sicherlich kann man sich eine Reihe von Prozessen vorstellen, die zu Änderungen im Wachstums dieser Pflanze führen können. Hier hilft auch die Bemerkung nicht weiter, mit der sich Menges auf eine „Gewißheit“ dafür beruft, daß „die allgemeinen Bedingungen konstant“ bleiben. Denn nicht nur gibt es für diese Annahme keine Begründung; es bleibt auch unklar, was mit „allgemeinen Bedingungen“ gemeint sein könnte. Menges denkt vermutlich – wie auch Flaskämper in dem oben angeführten Zitat – an eine Unterscheidung zwischen „allgemeinen“ und „besonderen“ Bedingungen. Aber dann könnte sich die durchschnittliche Wuchshöhe des gemeinen Habichtkrauts doch auch infolge einer Veränderung der „besonderen“ Bedingungen seines Wachstums verändern?

6. Ich will diese Fragen hier nicht vertiefen, sondern durch sie nur zeigen, wie Menges durch eine bestimmte Interpretation der Fisherschen Rhetorik dazu gebracht wird, eine merkwürdige Gegenüberstellung von Sozial- und Naturwissenschaften vorzunehmen. In den Naturwissenschaften könne man angeblich davon ausgehen, daß die „allgemeinen Bedingungen“, unter denen sich die von ihr betrachteten Erscheinungen entwickeln, im Zeitablauf konstant bleiben. Ergänzend beruft sich Menges darauf, daß man in den Naturwissenschaften (oder zumindest in einigen Bereichen) experimentieren könne:

„Einen Prüfstein für die Richtigkeit des Wahrscheinlichkeitsschlusses [der von einer Beobachtung der 12 Exemplare zu Aussagen über die Menge aller Exemplare des gemeinen Habichtkrauts führen soll] haben wir allemal im Experiment. Denn wenn wir unter den gleichen allgemeinen Bedingungen, wie sie beim Zustandekommen der für einen transzendenten Schluß bestimmten Masse geherrscht haben, Experimente durchführen, so müssen die möglichen realen Folgen mit den ihnen entsprechenden Wahrscheinlichkeiten wieder zum Vorschein kommen [...].“ (Menges 1959, S. 51)

Aber so wenig wie ein Erkenntnisinteresse garantieren kann, daß die durchschnittliche Wuchshöhe des gemeinen Habichtkrauts immer gleichbleiben wird, so wenig kann dies durch Experimente garantiert werden. Oder genauer gesagt: mithilfe technischer Vorkehrungen kann man vielleicht gewisse Bedingungen sicherstellen, unter denen das gemeine Habichkraut wachsen *soll*; nicht jedoch sicherstellen, wie es unter diesen Bedingungen wachsen wird.

7. Schließlich soll die ganze Argumentation jedoch nur dazu dienen, einen Unterschied zur Statistik in den Sozialwissenschaften zu begründen:

„Wir sind von dem Fisher'schen oder dem ‘eigentlich stochastischen’ Ansatz ausgegangen und haben zu zeigen versucht, daß es unter stochastischem Gesichtspunkt nicht nur berechtigt ist, sondern daß die einzige Möglichkeit darin liegt, eine Messungsreihe oder allgemein eine statistische Masse, sei sie ‘groß’ oder ‘klein’, als eine Zufallsstichprobe aus einer unendlich großen Grundgesamtheit aufzufassen. Eventuell sind die allgemeinen Bedingungen zu konkretisieren bis zu der (freilich immer noch unendlich groß zu denkenden) Masse all derjenigen Elemente, die unter den gleichen spezifischen (aber allgemeinen) Bedingungen stehen; im Falle der Wuchshöhe des gemeinen Habichtskrautes etwa den spezifischen Bedingungen der Lichtverhältnisse, des Klimas, der Bodenbeschaffenheit usw. Selbst wenn die allgemeinen Bedingungen variieren, wird der stochastische Ansatz Fishers nicht in Frage gestellt, weil auf dem Weg über das Experiment Veränderungen der allgemeinen Bedingungen kontrolliert werden können.

Diese Argumentation kann zweifellos für die ganze Physik und Chemie und weite Bereiche der Biologie und Geographie Geltung beanspruchen, für die Bereiche nämlich, auf die R. A. Fisher seinen Ansatz anzuwenden gedachte, indessen wohl nicht ausnahmslos für die Medizin und kaum für die Sozialwissenschaften. Hier haben wir Prozesse, deren allgemeine Bedingungen sich in steter Veränderung befinden und die sich einer experimentellen Behandlung schlechthin entziehen.“ (Menges 1959, S. 53)

Hiervon ausgehend wird dann für die sozialwissenschaftliche Statistik ein unterschiedliches Erkenntniszie1 formuliert. Menges' Überlegungen gehen so weiter:

„Müssen wir darum resignieren, eine solche Situation als unwissenschaftlich abtun? Nein! Es verschiebt sich lediglich das wissenschaftliche Interesse, ein anderes Prinzip tritt an die Stelle desjenigen, das der stochastischen Fragestellung

zugrundeliegt. Wir können eine beliebige statistische Masse zwar nicht mehr als Zufallsstichprobe, gezogen aus einer unendlich großen Grundgesamtheit, aber doch als gezogen aus einer *endlichen oder realiter existierenden Grundgesamtheit* auffassen. Die Schlußweise ist nicht länger mehr transzendent, aber sie ist nunmehr immanent: Wir müssen die Ausgangsverteilung, das Ziel unseres Schlusses, in ihrer *Stellung in Raum und Zeit* definiert und abgegrenzt haben, bevor wir Stichprobenziehungen überhaupt vornehmen können. Mit der Stichprobe verbinden wir konsequenterweise alsdann nicht die Frage nach den allgemeinen Bedingungen, sondern einfach nach den Eigenschaften der *hic et nunc* vorhandenen Gesamtmasse. Das Erkenntnisziel liegt in der Deskription!“ (Menges 1959, S. 55)

Menges' Vorschlag läuft also darauf hinaus, daß man sich in der Sozialstatistik darauf beschränken solle, empirisch in Raum und Zeit fixierbare Gesamtheiten mithilfe statistischer Begriffsbildungen zu beschreiben, daß es dagegen nicht sinnvoll sei, solche Gesamtheiten (oder durch sie definierbare Teilmengen) als Zufallsstichproben aus hypothetischen Superpopulationen zu denken.

8. Ich möchte dagegen zur Diskussion stellen, daß der Gedankengang auf zwei problematischen Voraussetzungen beruht. Zunächst darauf, daß Menges versucht, die Fishersche Rhetorik wörtlich zu nehmen. Infolgedessen kann die Problemstellung, auf die sich die Rhetorik bezieht, nicht deutlich gemacht werden: nämlich ein Verständnis für Prozesse zu gewinnen, durch die Sachverhalte und Ereignisse in unserer Erfahrungswelt hervorgebracht werden. Eine zweite fragwürdige Voraussetzung besteht darin, wie Menges Natur- und Sozialwissenschaften gegenüberstellt: daß man im ersten Fall von gleichbleibenden „allgemeinen Bedingungen“ ausgehen könne, wohingegen sich die Bedingungen des gesellschaftlichen Lebens fortwährend verändern. Aber insbesondere an Menges' Beispiel, der durchschnittlichen Wuchshöhe des gemeinen Habichtkrauts, kann man sich überlegen, daß auf diese Weise kein grundsätzlicher Unterschied gemacht werden kann. Denn es gibt nicht nur eine Sozialgeschichte, sondern auch eine Naturgeschichte (und im Lebensraum der Menschen sind beide ziemlich eng miteinander verbunden). In beiden Fällen kann man sich gedanklich auf Prozesse beziehen, durch die Sachverhalte und Ereignisse zustande kommen. Natürlich unterscheiden sich, zum Beispiel, die Wachstumsprozesse des gemeinen Habichtkrauts von denjenigen, durch die Autounfälle entstehen. Aber daraus folgt nicht, daß es nur im ersten, nicht auch im zweiten Fall sinnvoll sein könnte, sich zur begrifflichen Bezugnahme auf die Prozesse der Vorstellung eines Zufallsgenerators zu bedienen. Daß man dies in beiden Fällen machen *kann*, läßt sich kaum bestreiten. Es ist deshalb verständlich, daß Diskussionen, die sich auf einen grundsätzlichen Unterschied zwischen Natur- und Sozialwissenschaften berufen, aus der statistischen Methodenlehre weitgehend verschwunden sind. Bedauerlicherweise ist dadurch aber auch eine andere Fragestellung weitgehend unsichtbar gemacht worden: welche Erkenntnisinteressen es sinnvoll machen könnten, soziale (und ebenso natürliche) Prozesse als Realisationen von Zufallsgeneratoren zu deuten.

4.4 Konzeptionen der Stichprobentheorie

Es wurde darauf hingewiesen, daß die Verbreitung der Fisherschen Rhetorik auch dazu geführt, daß – jedenfalls in der allgemeinen Einführungsliteratur zur Statistik – die Unterscheidung zwischen empirisch in unserer Erfahrungswelt fixierbaren Gesamtheiten und hypothetischen Populationen, womit man sich in begrifflich unklarer Weise auf Prozesse bezieht, weitgehend verwischt worden ist. Es mag deshalb hilfreich sein, hier noch einen kurzen Blick auf die Spezialliteratur zur Stichprobentheorie zu werfen, die sich fast ausnahmslos auf reale Gesamtheiten bezieht und sich dadurch von der Fisherschen Idee hypothetischer unendlicher Populationen abgrenzt.[36] In diesem Kontext bewegt sich schließlich auch die Arbeit von Menges, auf die im vorangegangenen Abschnitt hingewiesen wurde.

4.4.1 Bezugnahme auf reale Gesamtheiten

1. Ein frühes Standardwerk zur Stichprobentheorie wurde 1953 von M. H. Hansen, W. N. Hurwitz und W. G. Madow veröffentlicht. Der Ansatz wird folgendermaßen motiviert:

"Whenever the problem of obtaining information concerning a large aggregate of people, farms, firms, records, or other units occurs, a sample survey or sample study is likely to provide results of the desired precision at relatively low cost." (Hansen, Hurwitz und Madow 1953, I, S. 1)

Die Gesamtheiten, auf die sich solche Stichproben beziehen, werden folgendermaßen erläutert:

[36]In einer Arbeit über die Entwicklung der Stichprobentheorie schreiben Hansen, Dalenius und Tepping (1985, S. 327): „This paper discusses selected topics in the theory and methods of sample surveys which have as their objectives to provide inferences about some characteristics of a finite population (as distinguished from inferences about a causal system which may have produced certain characteristics of the finite population)."

"*The population for which information is desired.* Information may be needed concerning the people living in a specified city, the families in the United States, the grocery stores in a given state, the electric light bulbs in a given shipment, the farms producing wheat, and so on. These are populations (or universes) of elementary units, such as people, families, grocery stores, electric light bulbs, or farms." (Hansen, Hurwitz und Madow 1953, I, S. 2)

Schließlich beziehen sich die Autoren in ihrer Stichprobentheorie ausdrücklich auf endliche Gesamtheiten:

"By a finite *population*, $\mathcal{A}$, is meant any well-defined set or class containing a finite number of *elements*, $A_1, A_2, \dots, A_N$. These elements may be plants, farms, persons, blocks, counties, businesses, electric light bulbs, insects, and so on. [...] Thus, to define a population we must be able to state the kind of elements of which it consists and to give rules for including or excluding any particular element. These rules may take the form of an enumeration of the elements of the set or may be a statement of the conditions the elements must satisfy." (Hansen, Hurwitz und Madow 1953, II, S. 1)

2. Ein weiteres Standardwerk sind die „Sampling Techniques" von W. G. Cochran (1977, 3. Aufl.). In der Einleitung heißt es:

"This book contains an account of the body of theory that has been built up to provide a background for good sampling methods. In most of the applications for which this theory was constructed, the aggregate about which information is desired is finite and delimited – the inhabitants of a town, the machines in a factory, the fish in a lake." (Cochran 1977, S. 1)

Gesamtheiten, mit denen sich die Stichprobentheorie beschäftigt, werden folgendermaßen erläutert:

"The word *population* is used to denote the aggregate from which the sample is chosen. The definition of the population may present no problem, as when sampling a batch of electric light bulbs in order to estimate the average length of life of a bulb. In sampling a population of farms, on the other hand, rules must be set up to define a farm, and borderline cases arise. These rules must be usable in practice: the enumerator must be able to decide in the field, without much hesitation, whether or not a doubtful case belongs to the population." (Cochran 1977, S. 5)

Die Bemerkung impliziert, daß es sich um real existierende Gesamtheiten handeln muß. Der Statistiker oder von ihm beauftragte Personen müssen „ins Feld" gehen können und dort im Prinzip alle Mitglieder einer Gesamtheit finden können. Dinge oder Situationen, die es nicht mehr oder noch nicht gibt, können infolgedessen nicht Elemente einer statistischen Gesamtheit im Sinne der Stichprobentheorie sein.

3. In einem neueren Lehrbuch über „Model Assisted Survey Sampling“ von C.-E. Särndal, B. Swensson und J. Wretman (1992) wird gleich zu Beginn definitiv festgestellt:

“A survey concerns a finite set of *elements* called a *finite population*. An enumeration exists that unequivocally defines the elements belonging to the population. The goal of a survey is to provide information about the finite population in question or about subpopulations of special interest, for example ‘men’ and ‘women’ as two subpopulations of ‘all persons’.” (Särndal, Swensson und Wretman 1992, S. 4f)

Die Autoren betonen auch, daß das Ziel in statistischen Informationen besteht, die sich auf Gesamtheiten beziehen:

“Although a survey involves observations on individual population elements, the purpose of a survey is not to use such data for decision-making about individual elements, but to obtain summary statistics for the population or for specific subgroups.” (Särndal, Swensson und Wretman 1992, S. 6f)

Das wurde natürlich auch schon früher von vielen Autoren betont, z.B. von W. Cochran, F. Mosteller und J.W. Tukey in einer Abhandlung über „The Principles of Sampling“ (1954, S. 17):

“Another principle which ought not to need recalling is this: By sampling we can learn only about collective properties of populations, not about properties of individuals.”

4. Es sei schließlich erwähnt, daß auch in der statistischen Literatur zu Stichprobenverfahren, die in Deutschland nach dem 2. Weltkrieg entstanden ist, stets davon ausgegangen worden ist, daß sich diese Verfahren auf reale, in der jeweils gegenwärtigen Realität fixierbare Gesamtheiten beziehen. So hat z.B. H. Kellerer in seiner „Theorie und Technik der Stichprobenverfahren“ (1953, S. 11) betont:

„Sowohl bei einer Vollerhebung wie bei einer Teilerhebung – und hier insbesondere bei einer Stichprobe – ist es wichtig, die zugrundeliegende Gesamtheit, die Grundgesamtheit (im Englischen ‘universe’ oder ‘population’ genannt) örtlich, zeitlich und sachlich genau abzugrenzen.“

4.4.2 Stichprobentheoretische Inferenz

1. Die Stichprobentheorie bezieht sich also auf reale, endliche, in der Erfahrungswelt fixierbare Gesamtheiten. Das Erkenntnisinteresse richtet sich auf Eigenschaften solcher Gesamtheiten oder genauer gesagt:

auf Eigenschaften der Verteilungen statistischer Variablen, die für solche Gesamtheiten definiert werden können. In der Stichprobentheorie spricht man von „Populationsparametern“, die als durch die Gesamtheit fest vorgegebene Größen konzipiert werden; z.B. die durchschnittliche Körpergröße in einer bestimmten real existierenden Menge von Personen. Das Ziel besteht darin, die Werte solcher als vorgegeben definierbaren Größen mithilfe der Informationen aus einer Stichprobe zu schätzen.[37]

2. Um dieses Ziel zu erreichen, werden – soweit dies möglich ist – Zufallsstichproben verwendet. Dafür gibt es hauptsächlich zwei Argumente. Erstens soll dadurch erreicht werden, daß man eine „repräsentative“ Stichprobe erhält, die keine verzerrten Schätzwerte liefert. Dieses Argument wird weitgehend anerkannt. In der Literatur zur Stichprobentheorie wird weiterhin fast ausnahmslos davon ausgegangen, daß man durch die Verwendung von Zufallsstichproben statistische Inferenzverfahren anwenden könne. Hierüber gibt es Kontroversen, die uns hier nicht zu beschäftigen brauchen. Wichtig ist jedoch, daß man versteht, wie statistische Inferenz im Rahmen der Stichprobentheorie konzipiert wird. Statistische Inferenz bezieht sich in diesem Fall auf ein Auswahlverfahren, durch das mögliche Stichproben mit einer jeweils angebbaren Wahrscheinlichkeit erzeugt werden können. Durch das Auswahlverfahren wird festgelegt, welche Stichproben erzeugt werden können und welche Wahrscheinlichkeiten ihnen durch das Verfahren zurechenbar sind. Die Wahrscheinlichkeitsaussagen, die infolgedessen gemacht werden können, beziehen sich also auf das Auswahlverfahren, jedoch *nicht* auf die tatsächlich realisierte Stichprobe. Das ist auch unmittelbar einsichtig: welche der durch das Auswahlverfahren als möglich definierten Stichproben man nach einer einmaligen Anwendung des Verfahrens tatsächlich bekommt, kann man nicht wissen. Trotzdem sind wahrscheinlichkeitstheoretische Untersuchungen über Auswahlverfahren nicht sinnlos. Man kann durch sie Eigenschaften der Auswahlverfahren und Eigenschaften von Schätzfunktionen untersuchen, die durch das Auswahlverfahren definiert werden.

[37] Was im Rahmen der Stichprobentheorie mit „schätzen“ gemeint ist, wurde von W.E. Deming (1950, S. 290) so erläutert: „The problem of estimation [... is] the question of how best to perform calculations on the data of the sample, perhaps in combination with prior data, to determine *within calculable limits* what answer would have been obtained from a complete coverage had one been taken with the same definitions, the same procedures of selecting and training the interviewers or assistants, the same questionnaire, and the same thoroughness as was possible with the sample.“

3. In der Literatur zur Stichprobentheorie wird auch meistens anerkannt, daß sich ihre Wahrscheinlichkeitsaussagen auf Auswahl*verfahren* bzw. dadurch definierte Schätz*verfahren* (im Unterschied zu jeweils spezifischen Schätzwerten) beziehen. Zum Beispiel heißt es bei Cochran:

"The purpose of sampling theory is to make sampling more efficient. It attempts to develop methods of sample selection and of estimation that provide, at the lowest possible cost, estimates that are precise enough for our purpose. [...] In order to apply this principle, we must be able to predict, for any sampling procedure that is under consideration, the precision and the cost to be expected. So far as precision is concerned, we cannot foretell exactly how large an error will be present in an estimate in any specific situation, for this would require a knowledge of the true value for the population. Instead, the precision of a sampling procedure is judged by examining the frequency distribution generated for the estimate if the procedure is applied again and again to the same population." (Cochran 1977, S. 8)

Särndal, Swensson und Wretman schreiben:

"We mentioned that two reasons for randomized sample selection are protection against selection bias and the fact that randomly selected samples are viewed as objective. In this book, the probability distribution associated with the randomized sample selection has another very important function. It provides the basis for probabilistic conclusions (inferences) about the target population. For instance, a 95 % confidence interval will have the property of covering the unknown population quantity in 95 % of all samples that can be drawn with the given probability sampling design. The term used in the literature for this kind of conclusion is *design-based inference* or *randomization inference*. Design-based inference is objective; nobody can challenge that the sample was really selected according to the given sampling design. The probability distribution associated with the design is 'real', not modeled or assumed." (Särndal, Swensson und Wretman 1992, S. 21)

Die letzte Bemerkung ist wichtig: es geht in diesem Fall nicht um die hypothetische Unterstellung eines Zufallsgenerators, über dessen Funktionsweise man mithilfe von durch ihn erzeugten Sachverhalten etwas herausfinden möchte. Die in der Stichprobentheorie diskutierten Auswahlverfahren hat man sich selbst gemacht, und man kennt infolgedessen ihre Eigenschaften und insbesondere die sie charakterisierenden Wahrscheinlichkeitsverteilungen. Aus der Kenntnis dieser Wahrscheinlichkeitsverteilungen können dann auf deduktive (mathematische) Weise Wahrscheinlichkeitsaussagen über Schätz*verfahren* abgeleitet werden, die durch das

Auswahlverfahren definiert sind.[38]

4. Es ist also nicht ganz klar, in welcher Weise man davon sprechen kann, daß man mit Daten aus einer Zufallsstichprobe Aussagen über eine korrespondierende Gesamtheit begründen kann.[39] Hält man sich jedoch an das in der Stichprobentheorie dominierende Selbstverständnis, ist jedenfalls folgendes klar: daß man bestenfalls Aussagen über diejenige Gesamtheit begründen kann, deren Elemente durch das jeweils verwendete Auswahlverfahren tatsächlich erreicht werden können.[40] Und dies kann natürlich nur eine reale, nicht eine hypothetische oder nur gedanklich konstruierte Gesamtheit sein. Infolgedessen kann auch das Ziel nur darin bestehen, schließlich zu einer Begründung von ihrer Intention nach deskriptiven Aussagen über reale Gesamtheiten zu gelangen.

5. Man mag dies für unbefriedigend halten und schließlich an „allgemeineren" Aussagen interessiert sein. Leslie Kish hat das in einem Buch über „Survey Sampling" (1965) folgendermaßen angedeutet:

"The *elements* of a population are the units for which information is sought; they are the *individuals*, the *elementary units* comprising the population about which inferences are to be drawn. [...]
We avoid using *universe* as a synonym for population. It denotes a hypothetical infinite set of elements generated by a theoretical model. This may be an ideal operation repeated endlessly, such as the endless tossing of a perfect coin. Behind every survey population stands some hypothetical universe, explicit or

[38] „How was the sample selected? In the probability sampling approach, we know the answer. The sample is selected according to the given probability sampling design. We know or can determine the probabilities given to the different possible samples. This is the key to describing the sample-to-sample variation of a proposed estimator." (Särndal, Swensson und Wretman 1992, S. 21) „There is no such thing as a 'representative', 'fair', or 'unbiased' sample. These terms apply only to the procedure by wich a sample is drawn or to the expected value of the estimators over repeated trials and not to the sample itself. Only by a knowledge of how the estimate derived from the sample can vary from one possible outcome of the sampling procedure to another can the procedure be judged." (O'Muircheartaigh 1977, S. 14)

[39] Denn die Theorie bezieht sich auf Schätz*verfahren*, also auf Regeln, mit denen man aus Daten Schätzwerte machen kann. „The statistician cannot predict the consequences of any single application of the rule. He forecasts merely how the rule will work *in the long run*. If the choice prescribed by the rule is tried many times in such and such circumstances, it will lead to such and such result in such and such percentage of the trials, in the long run." (Polya 1954, S. 82f)

[40] „The population to be sampled (the *sampled* population) should coincide with the population about which information is wanted (the *target* population). Sometimes, for reasons of practicability or convenience, the sampled population is more restricted than the target population. If so, it should be remembered that conclusions drawn from the sample apply to the sampled population." (Cochran 1977, S. 5)

implicit, definite or indefinite. For example, a statistic about the U.S. adult population on a fixed date is expected to have relevance for other dates and, often hopefully, for other cultures. The sample provides statistical inference to the survey population. Inference to hypothetical universe is more complex and not statistical (at least, not in the same sense)." (Kish 1965, S. 6ff)

Kish schränkt also die stichprobentheoretisch reflektierbare Inferenz ausdrücklich auf die reale Gesamtheit ein, aus der die Stichprobe tatsächlich durch ein Auswahlverfahren gebildet worden ist. Wenn aus den Ergebnissen weitergehende Schlußfolgerungen gezogen werden, können diese jedenfalls nicht mit stichprobentheoretischen Überlegungen begründet werden.

6. Allerdings sollte man sich überlegen, ob – wenn man „weitergehende" Schlußfolgerungen ziehen möchte – Begriffsbildungen sinnvoll sind, die sich am Denkmodell „Stichprobe–Grundgesamtheit" orientieren. Kish setzt das in dem eben angeführten Zitat voraus, indem er sich gedanklich auf hypothetische Populationen bezieht, von der die reale Gesamtheit, aus der man die Stichprobe gebildet hat, ein Teil ist. Das entspricht der Rhetorik von Fisher oder auch Andersons Stufenleiter höherer Gesamtheiten. Aber die Analogie ist, wie schon in vorausgegangenen Abschnitten deutlich zu machen versucht wurde, irreführend. Man kann sich das noch einmal anhand des Beispiel der Glühbirnen verdeutlichen, auf das in einigen der Zitate hingewiesen wird, die in Abschnitt 4.4.1 angeführt worden sind. Vergleicht man die in den Zitaten verwendeten Formulierungen, geht es darum, daß man zu Aussagen über real existierende Mengen von Glühbirnen kommen möchte. Eine solche Gesamtheit kann aber offenbar nur aus Glühbirnen bestehen, die bereits produziert worden sind und mithin für ein Auswahlverfahren zur Verfügung stehen. Was wäre in diesem Fall eine „weitergehende" Aussage? Sich gedanklich auf eine Menge von Glühbirnen zu beziehen, die mit einem bestimmten Produktionsverfahren erzeugt werden *könnten*, führt nicht zu einer definiten Menge, über die man sinnvolle Aussagen machten könnte. Es ist aber klar, woran man in diesem Fall interessiert ist: wie das Produktionsverfahren funktioniert, mit dem die Glühbirnen erzeugt werden. Und es könnte sinnvoll sein, das Produktionsverfahren als einen Zufallsgenerator aufzufassen und zu versuchen, das Verfahren z.B. durch eine Wahrscheinlichkeitsverteilung für die Lebensdauer von Glühbirnen zu charakterisieren, die durch das Verfahren erzeugt werden können.

4.4.3 Abstrakte Stichprobenräume

1. Ein Grund dafür, daß die Unterscheidung zwischen Zufallsgeneratoren, die sich auf Prozesse beziehen, und Auswahlverfahren in der statistischen Literatur selten kenntlich gemacht wird, liegt darin, daß in der mathematischen Wahrscheinlichkeitstheorie meistens von einem konzeptionellen Ansatz ausgegangen wird, der – bei rein formaler Betrachtung – eine solche Unterscheidung nicht erforderlich macht. Zur Verdeutlichung beziehe ich mich auf ein Lehrbuch von R. Schlittgen über „Statistische Inferenz" (1996). Es beginnt mit folgender Bemerkung:

„In der Statistik geht es darum, von Daten, den Ergebnisses wiederholter Beobachtungen eines Sachverhaltes, auf die zugehörige Grundgesamtheit oder den diese Daten generierenden Mechanismus zurückzuschließen. Die zugrundeliegenden Beobachtungsvorgänge werden im folgenden als Experimente bezeichnet." (Schlittgen 1996, S. 1)

Schon im ersten Satz werden zwei unterschiedliche Fragen vermischt: einerseits Fragestellungen, die sich auf eine real existierende Gesamtheit beziehen, andererseits Fragen, die sich auf einen Prozeß beziehen, durch den Sachverhalte zustande kommen.[41] Nicht ohne weiteres verständlich ist auch die Verwendung des Wortes 'Experiment', denn Daten entstehen nicht nur durch Experimente, auch z.B. durch Beobachtungen. Dazu gibt Schlittgen folgende Erläuterung:

„Unsere eingangs getroffene Vereinbarung, Beobachtungsvorgänge als Experimente zu bezeichnen, ist sehr großzügig. Danach sind naturwissenschaftliche Experimente genauso enthalten wie einfache Alltagsbeobachtungen. Um die Bezeichnung 'Experiment' zu rechtfertigen, setzen wir wenigsten die Wiederholbarkeit unter gleichen Bedingungen voraus. Erhalten wir bei solchen Wiederholungen jeweils das gleiche exakt vorhersagbare Ergebnis, so ist der induktive Schluß leicht. Wir haben es dann mit einer eindeutigen kausalen Beziehung zu tun. Unser Interesse gilt jedoch Experimenten, bei denen das Ergebnis gerade nicht exakt vorhersagbar ist." (Schlittgen 1996, S. 1)

Durch diese Erläuterung wird allerdings eine Voraussetzung eingeführt, die es ausschließt, beliebige Beobachtungen, insbesondere einfache Alltagsbeobachtungen, als Experimente aufzufassen. Denn was wir beobachten, verändert sich, und wenn man von Beobachtungen im Plural

[41] Schlittgen verwendet hier die Formulierung „Daten generierender Mechanismus", auf deren Zweideutigkeit bereits in Abschnitt 1.2 hingewiesen worden ist.

spricht, beziehen sie sich fast immer auf unterschiedliche Situationen.[42] Deshalb macht im allgemeinen die Vorstellung, daß man Beobachtungen „unter gleichen Bedingungen wiederholen“ kann, keinen Sinn. Verwendet man diese Forderung, um Experimente zu definieren, was man natürlich tun kann, werden infolgedessen Experimente in ganz spezifischer Weise als artifizielle Verfahren definiert: nämlich Verfahren, die so wiederholt werden können, daß eine angebbare Menge von Bedingungen im Wiederholungsfall erhalten bleibt.

2. Schlittgen definiert dann genauer, in welcher Weise er von Zufallsexperimenten sprechen will:

„Ein Experiment heißt *Zufallsexperiment*, sofern es folgende Forderungen erfüllt:

- Es ist unter gleichen Bedingungen wiederholt durchführbar.
- Bei der Durchführung ist nicht exakt vorhersagbar, zu welchem Ergebnis der Beobachtungsprozeß führt. Dies ist jedoch eindeutig.
- Es ist vorab angebbar, welche Ergebnisse überhaupt möglich sind.

Die Menge der möglichen Ergebnisse ω wird als *Stichprobenraum* Ω bezeichnet.“ (Schlittgen 1996, S. 2)

Die Begriffsbildung ist ersichtlich sehr allgemein. Die Bezugnahme auf „Zufallsexperimente“ ist auch im wesentlichen identisch mit einer Konzeption von Zufallsgeneratoren (abgesehen davon, daß eine unspezifische Konnotation mit Experimenten hervorgerufen wird). Man kann dies generell annehmen, wenn in der Literatur zur Wahrscheinlichkeitsrechnung und Statistik auf „Zufallsexperimente“ bezug genommen wird. Hier noch eine Formulierung von J. Neyman (1950, S. 9):

“Consider an experiment capable of yielding one or the other of some n different outcomes $E_1, E_2, \dots, E_n$. If a mathematical treatment of this experiment is adopted postulating that to each of the outcomes $E_1, E_2, \dots, E_n$, there corresponds a definite probability, whether known or unknown, whether the same for each E_k or not, then the experiment in question will be described as *random*. The adjective *random* will also be applied to describe the outcomes $E_1, E_2, \dots, E_n$.”

3. Zu überlegen ist, was man sich unter einer „Menge möglicher Ergebnisse“ vorstellen kann. Betrachten wir zu diesem Zweck einige der von Schlittgen angeführten Beispiele. Schlittgen beginnt mit Glücksspielen:

[42] Eine *Ausnahme* bilden Beobachtungen in der Theorie der Beobachtungsfehler, in der davon ausgegangen wird, daß sich eine Mehrzahl von Beobachtungen auf den gleichen Sachverhalt bezieht.

„Die eingängigsten Beispiele für Zufallsexperimente sind Glücksspiele wie das Roulettespiel, Würfeln, Münzwürfe, das Ziehen einer Karte aus einem Stapel gut gemischter Spielkarten. Hierbei handelt es sich um Vorgänge, die in unserem Kulturkreis wohlbekannt sind und die sich wegen ihrer einfachen Struktur besonders zur Illustration von neuen Konzepten eignen.

Beim Würfeln läßt sich der Stichprobenraum mit den Zahlen Eins bis Sechs identifizieren: $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Beim Werfen einer 2 DM-Münze kann die Kopfseite oder die Seite mit dem Adler nach oben zu liegen kommen. Wir können also setzen: $\Omega = \{K, A\}$.“ (Schlittgen 1996, S. 2)

In diesen Beispielen beziehen sich die Elemente des Stichprobenraums nicht auf Objekte, auch nicht auf Ereignisse, vielmehr auf Kennzeichnungen, die verwendet werden können, um die durch den Zufallsgenerator erzeugbaren Ereignisse zu charakterisieren.[43] In den angeführten Beispielen ist auch klar, inwiefern die Verfahren, auf die man sich gedanklich bezieht, beliebig oft „unter gleichen Bedingungen“ (die durch das Verfahren festgelegt werden) wiederholt werden können.[44] Daß man sich hier auf *Verfahren* bezieht, zeigt auch, daß in diesen Fällen die Vorstellung einer Population – im Sinne einer indefiniten Menge aller möglichen Ereignisse, die mit dem Zufallsgenerator erzeugt werden könnten – gar nicht sinnvoll anwendbar ist. Man würde sonst nur erneut bei der Verwechslung enden, wie sie in Abschnitt 4.3.2 am Beispiel der Ausführungen von Yule und Kendall besprochen worden ist.

4. Die Konzeption eines Stichprobenraums *kann* aber auch verwendet werden, um Auswahlverfahren als Zufallsexperimente zu deuten. Schlittgen weist darauf mit folgendem Beispiel hin:

[43] Dies sollte betont werden, denn in der Wahrscheinlichkeitstheorie wird oft auf eine unklare Weise von Ereignissen gesprochen. Wird die Wahrscheinlichkeitstheorie als eine Theorie der Zufallsgeneratoren entwickelt, beziehen sich ihre Aussagen überhaupt nicht auf Ereignisse, sondern auf Arten von Ereignissen (Ereignistypen), die mit einem Verfahren erzeugt werden können. Sie beziehen sich insbesondere nicht darauf, ob und mit welchen Häufigkeiten Ereignisse, die mit dem Verfahren erzeugt werden könnten, tatsächlich auftreten.

[44] Diese Bedingung wird zwar von den meisten Autoren als eine explizite Forderung an Zufallsexperimente gestellt, bei der Angabe von Beispielen und Anwendungen dann aber oft ignoriert. Zum Beispiel findet sich bei A. M. Mood, F. A. Graybill und D. C. Boes (1974, S. 16) folgendes Beispiel: „The experiment is to record the number of traffic deaths in the state of Colorado next year. Any nonnegative integer is a conceivable outcome of this experiment; so $\Omega = \{0, 1, 2, \dots\}$.“ Diese Betrachtungsweise führt direkt in die Konfusionen, die in Abschnitt 4.3.3 am Beispiel von Hagoods fiktiven Populationen besprochen worden sind.

„Ein Zufallsexperiment besteht in der Auswahl einer Person aus einer vorgegebenen, genau spezifizierten Personengruppe mittels Losverfahren. Jede Person stellt hier ein mögliches Ergebnis ω der Zufallsauswahl dar. Ω ist die endliche Gesamtheit aller Personen der betrachteten Gruppe.“ (Schlittgen 1996, S. 2)

Offenbar gibt es einen Unterschied zu den zuvor betrachteten Glücksspielen. Denn jetzt besteht der Stichprobenraum aus einer dem Zufallsexperiment als gegeben voraussetzbaren Menge von Dingen (Personen). Infolgedessen kann in diesem Fall der Stichprobenraum als eine Population gedeutet werden. Dies impliziert allerdings, daß man die Elemente des Stichprobenraums als namentlich unterscheidbare Individuen auffassen kann.

5. Offenbar handelt es sich um unterschiedliche Konzeptionen eines Stichprobenraums. Bezieht man sich auf Experimente oder allgemein auf Zufallsgeneratoren, faßt der Stichprobenraum die Menge der Arten von Sachverhalten oder Ereignissen zusammen, die durch ein Verfahren erzeugt werden können. In diesem Fall entspricht der Begriff eines Stichprobenraums dem Begriff des Merkmalsraums einer Zufallsvariablen.[45] In der zweiten Interpretation ist jedoch der Stichprobenraum kein Merkmalsraum, sondern eine statistische Gesamtheit, die aus einer Menge bestimmter Individuen besteht. Bezieht man sich in diesem Fall auf einen Zufallsgenerator in Form eines Auswahlverfahrens, kann man in einigen Fällen die vorausgesetzte Gesamtheit mit einem durch das Auswahlverfahren definierbaren Stichprobenraum identifizieren.[46] Es sei indessen darauf hingewiesen, daß einige Autoren auch bei der Betrachtung von Stichproben aus endlichen Gesamtheiten von einer anderen Definition ausgehen.[47]

[45]Zufallsvariablen werden oft als Abbildungen eines Stichprobenraums in einen Merkmalsraum definiert; als Spezialfall kann man indessen auch eine identische Abbildung auf dem Stichprobenraum konzipieren. Wichtig ist, daß in diesem Fall der Stichprobenraum ein Merkmalsraum ist.

[46]Dies setzt ein „Ziehen ohne Zurücklegen“ voraus, andernfalls muß man den Stichprobenraum durch Ω^n definieren, wenn Ω die zunächst konzipierte Gesamtheit bezeichnet.

[47]Zum Beispiel P. Whittle (1992, S. 5f): “Suppose a country of N people has a 100 per cent census and that, as a result of the information gathered, individuals can be assigned to one of K mutually exclusive categories or cells, which we shall label $\omega_1, \omega_2, \ldots, \omega_K$. [...] The value of ω attachted to an individual is what will more generally be termed the *realization*, the description of the individual at the level adopted. The motivation for the term is that, if one picks an individual ‘at random’, i.e. off the street or blindly from a directory, then examination of the value of ω tells one exactly in which category the individual fell in the particular case. The possible realization values $\omega_1, \omega_2, \ldots, \omega_K$ can be regarded as points in an abstract space Ω. This is termed the *sample space*, because the sampling of an individual can be regarded as the sampling of a point ω from Ω. Any property of an individual which is determined by the realization is termed a *random variable* – these are the variables which are meaningful at the level of description adopted.” Hier ist Ω also nicht eine Menge von Individuen, keine statistische Gesamtheit, sondern Merkmalsraum einer Zufallsvariablen, die durch ein Auswahlverfahren definiert wird. Bemerkenswert ist vor allem die Formulierung im letzten Satz, daß Eigenschaften von Individuen dadurch „determiniert“ werden, daß eine so definierte Zufallsvariable ihre Werte annimmt. Die Formulierung könnte einen Sinn machen, wenn sich die Zufallsvariable auf einen Prozeß bezieht, durch den Individuen ihre Eigenschaften erhalten. Aber diese Interpretation kann nicht gemeint sein, wenn Zufallsvariablen durch ein Auswahlverfahren definiert werden, für das eine Menge von Dingen, Individuen oder Situationen mit *gegebenen* Eigenschaften vorausgesetzt werden muß.

6. Jedenfalls zeigen die Ausführungen von Schlittgen, daß der Begriff eines Stichprobenraums, wie er in der Wahrscheinlichkeitstheorie verwendet wird, als eine formale Repräsentation sowohl für Zufallsgeneratoren als auch für Auswahlverfahren verwendet werden kann. Dies gilt infolgedessen auch für den Begriff eines Wahrscheinlichkeitsraums, der an die Konzeption eines Stichprobenraums anschließt. In der Wahrscheinlichkeitstheorie versteht man darunter eine Zusammenfassung von drei Begriffen:

$$(\Omega, \mathcal{A}, \mathrm{P})$$

Dabei ist Ω ein Stichprobenraum, $\mathcal{A}$ ein System von Teilmengen von Ω, das den Anforderungen einer Mengenalgebra genügt, und P ist ein Wahrscheinlichkeitsmaß, das jedem Element von $\mathcal{A}$ eine numerisch bestimmte Wahrscheinlichkeit zuordnet.

7. An diesen Begriff eines Wahrscheinlichkeitsraums knüpft zum Beispiel C. R. Rao in seiner Einführung in die Statistik (1995) an:

„In der statistischen Inferenz, d.h. beim Treffen von Aussagen über eine Grundgesamtheit auf der Basis einer aus ihr gezogenen Stichprobe, ist es notwendig, die heilige Dreiheit zu bestimmen, d.h. den Stichprobenraum Ω, die auf Ω definierte Borelsche σ-Algebra B [in unserer Notation $\mathcal{A}$] und eine auf B definierte Familie von Wahrscheinlichkeitsverteilungen P [P wird hier von Rao als eine Menge möglicher Wahrscheinlichkeitsmaße konzipiert]. Statistische Analyse beschäftigt sich mit dem Herstellen einer Übereinstimmung zwischen einer Stichprobe (einer Teilmenge von Ω) und einem Element (oder einer Teilmenge von Elementen) von P.“ (Rao 1995, S. 137f)

Indem Rao davon ausgeht, daß eine Stichprobe als eine Teilmenge eines Stichprobenraums konzipiert werden kann, muß er eigentlich voraussetzen, daß es sich bei Ω um eine Population im Sinne der Stichprobentheorie handelt.[48] Tatsächlich bezieht er sich aber unterschiedslos auf Stichproben im Sinne der Stichprobentheorie und Experimente. Das Zitat geht so weiter:

> „Ein wichtiger Teil der Dreiheit ist die Spezifikation von P. Eine falsche Spezifikation kann zu falscher Inferenz führen, die in der statistischen Fachsprache manchmal auch *Fehler 3. Art* genannt wird. Das Problem der Spezifikation ist kein leichtes. Detailliertes Wissen über das bei der Datengewinnung tatsächlich angewandte Verfahren ist essentieller Bestandteil der Suche nach der korrekten Spezifikation. Die Situation ist im Falle von Feldbeobachtungen und nichtexperimentellen Daten, bei denen die Natur Ereignisse gemäß einem bestimmten stochastischen Modell produziert und die Ereignisse dann von einem Untersucher beobachtet und aufgezeichnet werden, wesentlich komplizierter. Es existiert nicht immer ein geeigneter Versuchsaufbau zur Planung einer Stichprobenerhebung, der sicherstellt, daß die eingetretenen Ereignisse festgelegte (in der Regel die gleichen) Wahrscheinlichkeiten besitzen, in die Stichprobe zu gelangen.“ (Rao 1995, S. 138)

Offenbar denkt Rao in erster Linie an Experimente oder an fiktive Zufallsgeneratoren, wie sie in der theoretischen Statistik angenommen werden. Aber seine vorausgesetzte Definition von Stichproben und die sich anschließende Rhetorik machen eigentlich nur in einem stichprobentheoretischen Kontext einen Sinn.

8. An dieser Stelle kann man sich noch einmal Fishers (und auch Andersons) Vorschlag überlegen, Wahrscheinlichkeiten durch relative Häufigkeiten zu deuten. Das erscheint möglich, wenn es sich beim Stichprobenraum um eine endliche Gesamtheit von Objekten oder Situationen handelt. Dann kann man *statistische* Variablen definieren, in unserer Notation

$$X : \Omega \longrightarrow \tilde{X}$$

und, wenn man das Auswahlverfahren als eine einfache Zufallsauswahl konzipiert, das Wahrscheinlichkeitsmaß P mit der Häufigkeitsverteilung von X formal identifizieren. Dies setzt allerdings voraus, daß es sich bei Ω um eine endliche, dem Auswahlverfahren voraussetzbare Gesamtheit handelt. Eine Deutung von Wahrscheinlichkeiten durch relative Häufigkeiten ist dagegen nicht möglich, wenn sich der Begriff eines Wahrscheinlichkeitsraums auf einen Zufallsgenerator bezieht, durch den Experimente oder fiktive Prozesse repräsentiert werden sollen, die nicht als Auswahlverfahren konzipiert werden können.

[48] Wollte man auch den Fall eines „Ziehens mit Zurücklegen“ berücksichtigen, müßte man von Ω^n ausgehen.

9. Schließlich sollte noch einmal betont werden, daß den beiden unterschiedlichen Deutungsmöglichkeiten eines Wahrscheinlichkeitsraums unterschiedliche Fragestellungen korrespondieren. Einerseits kann sich die Fragestellung auf Prozesse beziehen, durch die Sachverhalte oder Ereignisse zustande kommen. Dann kann man versuchen, diese Prozesse als Zufallsgeneratoren zu deuten und Einsichten in ihre Funktionsweise zu gewinnen. Soweit es sich in diesem Fall nicht um durch ein Verfahren explizierbare Zufallsgeneratoren handelt, zielt die Theoriebildung auf eine gedankliche Konstruktion fiktiver Zufallsgeneratoren, durch die man Prozesse in unserer Erfahrungswelt deuten möchte. Andererseits kann sich die Fragestellung auf Häufigkeitsverteilungen von Eigenschaften in real existierenden Gesamtheiten richten. Wenn in diesem Fall eine vollständige Untersuchung der Elemente der Gesamtheit nicht möglich oder sinnvoll ist, kann man Auswahlverfahren verwenden, um Teilmengen der Gesamtheit einer Beobachtung zugänglich zu machen. Auch solche Auswahlverfahren können als Zufallsgeneratoren konzipiert werden. Aber in diesem Fall entstehen die Sachverhalte, über die man Kenntnisse gewinnen möchte, nicht durch den Zufallsgenerator, sondern dieser wählt nur aus, was einer näheren Beobachtung unterzogen werden soll. Auch richtet sich die Fragestellung in diesem Fall nicht darauf, wie der Zufallsgenerator funktioniert (eine solche Kenntnis wird vielmehr vorausgesetzt), sondern er dient nur als ein methodisches Hilfsmittel für die Organisation des Beobachtungsprozesses.

Kapitel 5

Hypothesen und Modelle

Sozialstatistik zielt zunächst darauf, Daten über soziale Sachverhalte und Ereignisse zu gewinnen und sie in geeigneter Weise darzustellen. Theoretische Sozialstatistik setzt mit der Frage ein, wie die Sachverhalte und Ereignisse, auf die sich sozialstatistische Daten beziehen, entstanden sein könnten. Jedenfalls zielt der theoretische Anspruch darauf, Fragen dieser Art reflektierbar zu machen. Diesem Zweck dient die Konzeption hypothetischer Zufallsgeneratoren, die die Sachverhalte und Ereignisse erzeugt haben könnten. In einer verbreiteten Rhetorik wird davon gesprochen, daß Hypothesen gebildet und Modelle konstruiert werden. In diesem Kapitel soll damit begonnen werden, den Charakter sozialstatistischer Hypothesen und Modelle etwas genauer zu verstehen.

5.1 Deskriptive und generative Hypothesen

In diesem Abschnitt soll zunächst besprochen werden, daß man auf unterschiedliche Weisen von Hypothesen sprechen kann. Ich möchte auch zeigen, daß die im vorangegangenen Kapitel diskutierte Unterscheidung zwischen Prozessen und Auswahlverfahren zu unterschiedlichen Arten von Hypothesen führt.

5.1.1 Worauf beziehen sich Hypothesen

1. Folgt man dem in der theoretischen (Sozial-) Statistik üblich gewordenen Sprachgebrauch, beziehen sich statistische Hypothesen auf Merkmalsverteilungen in Gesamtheiten (Populationen) oder auf aus solchen Merkmalsverteilungen ableitbare Größen.[1] Was mit Hypothesen dieser Art gemeint ist, kann man verstehen, wenn sie sich auf eine real existierende Gesamtheit beziehen. Denn dann haben Hypothesen einen empirisch explizierbaren Sinn: sie betreffen Eigenschaften einer realen

[1] Z.B. heißt es bei Kreyszig (1968, S. 203): „Unter einer *Hypothese* versteht man in der Statistik eine Annahme über die Verteilung einer Zufallsvariablen, z.B. die Annahme, daß die betreffende Verteilung den Mittelwert 20,3 hat, usw. Und ein *Test* einer Hypothese ist ein Prüfverfahren, das man anwendet, wenn man wissen will, ob man die Hypothese *'annehmen'* kann oder *'verwerfen'* soll, d.h. ob man der Hypothese Vertrauen schenken und sie für richtig halten kann oder nicht.“

Gesamtheit. Diesem Verständnis von Hypothesen korrespondiert das sozialstatistische Inferenzproblem.

2. Aber gleichermaßen wird von Hypothesen gesprochen, wenn man sich gedanklich auf eine hypothetische oder fiktive Population bezieht, so wie dies von der Fisherschen Rhetorik nahegelegt wird. In der rein mathematischen Statistik ist diese Rhetorik unerheblich, weil sie den Gehalt der mathematischen Ableitungen nicht betrifft. Für die statistische Sozialforschung entsteht jedoch ein Problem, wenn sie potentiell sinnvolle Hypothesen formulieren möchte. Denn dies verlangt, daß der gemeinte Sinn expliziert wird. Z.B. heißt es in einem Lehrbuch von H. J. Loether und D. G. McTavish:

> "Sociologists are engaged in the business of developing generalizations about social behavior. They attempt to generate explanatory schemes that may predict the orderly operating processes of such social systems as total societies (e.g., the United States), communities (e.g., Solvang, California), organizations (e.g., the American Legion), and groups (e.g., the corner gang). These explanatory schemes are generally tested through the analysis of limited bodies of data, called *samples*. The samples are not inherently interesting in themselves, however. They are only of interest to the extent that they represent larger, more significant bodies of data known as *populations*." (Loether und McTavish 1974, S. 42)

Anhand dieser Ausführungen könnte man zunächst vermuten, daß sich die Hypothesen auf definite, real existierende Gesamtheiten beziehen sollen, denn die Autoren charakterisieren Populationen als „bodies of data", eine Formulierung, die nur dann einen Sinn machen könnte, wenn man sich auf eine real existierende Gesamtheit beziehen kann. Dann geht es jedoch so weiter:

> "As a scientist the sociologist is ultimately interested in developing generalizations about social behavior that are applicable to infinite populations, sometimes referred to as *general universes* (the terms *population* and *universe* may be used interchangeably). Because of the methodological problems involved, however, he generally does research on special populations or samples of special populations. In fact, sampling theory, as it is practically applied, focuses on *special populations* or *universes* rather than on general universes. By virtue of the fact that a simple random sample requires a list of population elements, the researcher is generally restricted to working with finite populations. He is faced, then, with the methodological problem of how to make the inferential leap from the special universe with which he must work to the general universe in which he is ultimately interested. [...]
> There are, then, two leaps typically involved in the process of drawing conclusions about general universes. One is the leap from sample to special universe from which the sample was drawn. The other is the leap from special to general universe. The theoretical leap from sample to special universe is the subject of this volume [Inferential Statistics for Sociologists], and the simple random sample is the technique designed to justify the leap. The leap from the special to the general universe is neither directly addressed nor justified by current theories and techniques of statistical inference." (Loether und McTavish 1974, S. 42)

Loether und McTavish folgen hier einer stichprobentheoretischen Betrachtungsweise, wie wir sie in Abschnitt 4.4.2 besprochen haben. Sie kommen demzufolge zu dem Ergebnis, daß die angestrebten „allgemeinen" Hypothesen mit stichprobentheoretischen Überlegungen nicht einschätzbar gemacht werden können.

3. Diese vorsichtige Position ist indessen eher eine Ausnahme. Meistens wird in der Methodenliteratur die Auffassung verbreitet, daß auch „allgemeine" Hypothesen mit statistischen Inferenzverfahren einschätzbar gemacht („getestet") werden können. Exemplarisch kann man an die in Abschnitt 1.2 zitierten Bemerkungen von D. Hochstädter denken, in denen der Inferenzstatistik die Aufgabe zugewiesen wird, „allgemeingültige Aussagen" und „allgemeine Regelmäßigkeiten" zu begründen. Seit sich das Schema „Stichprobe–Grundgesamtheit" als allgemeiner Denkrahmen für statistische Inferenz durchgesetzt hat, besteht der rhetorische Trick einfach darin, die Definition der Gesamtheit, auf die sich die Hypothesen beziehen sollen, unbestimmt zu lassen.[2]

[2] Dazu heißt es bei J. Galtung (1967, S. 316): „Science is generalizing, although it may elaborate special cases, as is done in the ideographic sciences. The hypothesis is about the units in the set S [in unserer Notation Ω], which we shall call the *intended field of tenability* (when it is confirmed, we drop the word 'intended'). To specify this set is to specify the conditions under which the hypothesis is assumed to be valid. A cardinal sin in social science research is precisely the failure to specify the conditions of assumed validity." — In welcher Weise ein Allgemeinheitsanspruch vertreten wird, ist tatsächlich oft schwer oder gar nicht erkennbar, insbesondere im Hinblick auf die zeitliche Eingrenzung der Gesamtheiten, auf die sich in der üblichen Rhetorik die Hypothesen beziehen sollen. Z.B. heißt es in einem Buch über „Statistik in den Sozialwissenschaften" von J. Kriz (1973, S. 106): „Ohne allzu weit darauf eingehen zu wollen, sei an dieser Stelle nur bemerkt, daß praktisch alle sozialwissenschaftlich interessanten Grundgesamtheiten unendlich sind, denn allgemeinere Aussagen sollen sich ja nicht nur auf den Zeitpunkt der Erhebung der Daten beziehen, sondern zumindest für einen kleineren Zeitraum Gültigkeit haben, d.h. repräsentativ sein. Selbst wenn man somit die Daten aller Menschen der BRD erheben würde, wäre dies nur eine Stichprobe aus allen möglichen Erhebungen innerhalb eines bestimmten Zeitraumes." Hier bleibt nicht nur der zeitliche Geltungsbereich unbestimmt; Kriz beschäftigt sich auch nicht mit der Frage, wie man heute schon Stichproben aus Gesamtheiten ziehen kann, die es erst in der Zukunft geben wird.

4. Die eher vorsichtige Position von Loether und McTavish entsteht daraus, daß sie zwei Arten von Populationen unterscheiden. Einerseits reale Gesamtheiten, auf die sich die Auswahlverfahren beziehen, durch die Stichproben definiert werden. Andererseits Superpopulationen, auf die sich die schließlich angestrebten Hypothesen beziehen. Die primär wichtige Frage besteht jedoch darin, ob durch eine gedankliche Bezugnahme auf hypothetische Superpopulationen überhaupt sinnvolle Hypothesen formuliert werden können. Daß das fragwürdig ist, hat man schon in Abschnitt 4.3.3 anhand der Ausführungen von M. J. Hagood sehen können, die versucht hat, diese Art der Hypothesenformulierung ernst zu nehmen. Ich glaube, daß sich ein Ausweg aus dieser Sackgasse, wenn überhaupt, nur dadurch finden läßt, daß man sich gedanklich nicht auf Populationen, sondern auf Prozesse bezieht, durch die soziale Sachverhalte und Ereignisse zustande kommen. Demzufolge sollten also zwei Arten von Hypothesen unterschieden werden.

a) Hypothesen, die sich auf Merkmalsverteilungen oder daraus ableitbare Größen in real existierenden Gesamtheiten beziehen, und

b) Hypothesen, die sich auf Prozesse beziehen, durch die soziale Sachverhalte und Ereignisse zustande kommen.

In Ermangelung besserer Bezeichnungen spreche ich im ersten Fall von *deskriptiven*, im zweiten Fall von *generativen* Hypothesen. Eine unserer Fragen sollte darin bestehen, ob und ggf. wie sich Hypothesen, die sich in ihrer sprachlichen Form auf indefinite, hypothetische, fiktive Populationen beziehen, auch als generative Hypothesen verstehen lassen.

5.1.2 Beziehen sich Hypothesen auf Variablen?

1. In der Literatur gibt es noch eine andere Rhetorik: Hypothesen werden nicht nur auf Populationen bezogen, sondern auch auf Variablen. Diese Rhetorik ist insbesondere in der Methodenliteratur zur empirischen Sozialforschung verbreitet. Zum Beispiel heißt es in einem neueren Methodenlehrbuch:

> „Allgemein bezeichnet man diejenigen Aussagen als *'Hypothesen'*, die einen Zusammenhang zwischen mindestens zwei Variablen postulieren. Unter einer *'Variablen'* versteht man einen Namen (z.B. 'Geschlecht', 'V1', 'Var 134') für die Menge von Merkmalsausprägungen, die Objekten (z.B. Personen) zugeschrieben werden.
> Die Variable 'Geschlecht' besitzt die Variablen- bzw. Merkmalsausprägungen

> 'männlich' und 'weiblich', die Variable 'Haushaltsgröße' z.B. die Ausprägungen '1', '2' usw., die Variable 'Mathematikangst' kann die Ausprägungen 'sehr groß', 'groß' usw. besitzen.
> Hypothesen können sowohl als *'Wenn-Dann-Aussagen'* (z.B. 'Wenn man Gase erhitzt, dann dehnen sie sich aus'), als auch als *'Je-Desto-Aussagen'* formuliert werden. Der von der Hypothese behauptete Zusammenhang kann prinzipiell jede beliebige mathematische Form annehmen; es sind sowohl lineare als auch nichtlineare Funktionen als Zusammenhangsform möglich. Wird der Verlauf dieser Funktionen nicht nur annäherungsweise charakterisiert, sondern auch die konkrete Funktionsgleichung mit den entsprechenden mathematischen Parametern angegeben, so spricht man von *'quantitativen Hypothesen'* bzw. quantitativen Gesetzen.
> *'Gesetze'* sind strukturell identisch mit Hypothesen. Man verwendet den Gesetzesbegriff jedoch vor allem dann, wenn sich die entsprechende Aussage bereits häufig an der Realität 'bewährt' hat.“ (Schnell, Hill und Esser 1999, S. 51f)

Ähnliche Erläuterungen zum Hypothesenbegriff findet man in vielen anderen Methodenlehrbüchern, die sich mehr oder weniger an Begriffsbildungen der Statistik orientieren.[3]

2. Versucht man, Erläuterungen dieser Art zu verstehen, stellt sich indessen ein Problem: in welcher Weise und mit welchem Anspruch beziehen sich Hypothesen auf die Realität? Das Problem zeigt sich zunächst darin, wie von Variablen gesprochen wird. Wie das angeführte Zitat deutlich macht, werden Variablen mit Merkmalsräumen identifiziert, also mit begrifflichen Konstruktionen, die – für sich genommen – keine Bezugnahme auf reale Sachverhalte oder Ereignisse beinhalten.[4] Das Problem wird deutlich, wenn man sich an unseren Begriff einer statistischen Variablen erinnert, der durch eine Abbildung

$$X : \Omega \longrightarrow \tilde{X}$$

definiert ist. Hier ist die Variable X eine Funktion, die jedem Mitglied aus einer Gesamtheit Ω einen bestimmten Wert in einem Merkmalsraum $\tilde{X}$ zuordnet. Durch diese Konzeption wird also deutlich, wie man sich mit statistischen Variablen auf Gegebenheiten in unserer Erfahrungswelt beziehen möchte: nämlich auf Eigenschaften von Mitgliedern einer angebbaren Menge von Dingen, Menschen oder Situationen. Wenn man

[3] Man vgl. etwa Sirkin 1995, S. 5ff.

[4] Daß hier ein Problem vorliegt, zeigt auch folgende Formulierung von Sirkin (1995, S. 8): „The concepts or ideas that we call variables are the phenomena of particular interest in our social science disciplines.“

dagegen Variablen als Merkmalsräume einführt, in unserer Notation $\tilde{X}$, gibt es keinerlei Anhaltspunkt, wie man sich auf empirisch explizierbare Sachverhalte oder Ereignisse beziehen möchte.

3. Das Problem ist durchaus folgenreich für die Frage, wie man sinnvolle Hypothesen formulieren kann. Werden nämlich Variablen mit Merkmalsräumen identifiziert, kann man der Vorstellung, daß sich Hypothesen auf „Zusammenhänge zwischen Variablen" beziehen, zunächst überhaupt keine sinnvolle Bedeutung geben. Man kann zwar gedanklich mathematische Funktionen zwischen Merkmalsräumen einführen; aber die Aussagen über „Zusammenhänge zwischen Variablen", wie sie mit statistischen Methoden konstruiert werden können, haben mit solchen Funktionen nichts zu tun. Aber auch wenn man sich gedanklich auf „Zusammenhänge zwischen Merkmalsräumen" beziehen könnte (oder möchte), bliebe vollständig offen, wie Aussagen über solche Zusammenhänge als Aussagen über empirisch explizierbare Aspekte unserer Erfahrungswelt verstanden werden könnten. Geht man dagegen von statistischen Variablen aus, ist der Realitätsbezug von vornherein gegeben. Die Idee, daß sich Hypothesen auf „Zusammenhänge zwischen Variablen" beziehen kann dann auf einfache Weise erläutert werden. Den Ausgangspunkt bildet eine zwei- oder mehrdimensionale statistische Variable

$$(X,Y) : \Omega \longrightarrow \tilde{X} \times \tilde{Y}$$

Man kann dann Hypothesen als Aussagen über die gemeinsame Häufigkeitsverteilung der Variablen formulieren, und natürlich auch bedingte Verteilungen betrachten usw. So kann man darstellen, in welcher Art von Verbundenheit die Werte der Variablen bei den Mitgliedern der Gesamtheit Ω auftreten.

4. Der Versuch, den Sinn von Hypothesen über „Zusammenhänge zwischen Variablen" dadurch verständlich zu machen, daß man sich auf *statistische* Variablen bezieht, hinterläßt jedoch ein Problem: man kann auf diese Weise nur deskriptive Hypothesen formulieren. Hypothesen sind dann Aussagen über Merkmalsverteilungen in realen, empirisch fixierbaren Gesamtheiten. Eine Bezugnahme auf statistische Variablen (im hier verwendeten Verständnis dieses Begriffs) liefert dagegen keinen Ausgangspunkt, um auch generative Hypothesen zu formulieren, also Hypothesen, die sich gedanklich auf Prozesse beziehen, durch die Dinge, Menschen und Situationen entstehen bzw. jeweils bestimmte Eigenschaften bekommen. Wenn man generative Hypothesen als Aussagen über Variablen formulieren möchte, benötigt man einen anderen Variablenbegriff.

5. In der theoretischen Statistik ist hierfür der Begriff einer Zufallsvariablen eingeführt worden. Dahinter verbirgt sich indessen nur eine andere sprachliche Form für das Reden über Zufallsgeneratoren. Denn Zufallsvariablen sind Variablen, bei denen man davon sprechen kann, daß sie ihre möglichen Werte mit bestimmten Wahrscheinlichkeiten annehmen.[5] Jeder Versuch, diesen Begriff zu explizieren, verlangt also eine gedankliche Bezugnahme auf Prozesse, durch die Zufallsvariablen ihre möglichen Werte annehmen können. Das ist indessen der Sinn der Vorstellung eines Zufallsgenerators: sie soll einen Prozeß vorstellbar und begrifflich explizierbar machen, durch den mögliche Sachverhalte oder Ereignisse mit bestimmten Wahrscheinlichkeiten entstehen können. Sobald eine solche Explikation vollzogen worden ist, sieht man, daß der Begriff 'Zufallsvariable' kein Grundbegriff der Wahrscheinlichkeitstheorie ist, sondern nur eine sprachliche Abkürzung für das Reden über Zufallsgeneratoren liefert. Erinnert man sich an die in Abschnitt 1.3 eingeführte Notation für Zufallsgeneratoren:

$$\mathcal{G} = (\mathcal{V}, (\mathcal{S}, \tilde{X}))$$

kann man Zufallsvariablen auch formal einführen, entweder als identische Abbildung des Merkmalsraums $\tilde{X}$ auf sich selbst oder als eine Abbildung, die von $\tilde{X}$ in einen neuen Merkmalsraum führt. Aber es ist klar, daß nur eine elliptische Redeweise vorliegt, wenn man eine identische Abbildung auf dem Merkmalsraum $\tilde{X}$ als eine Zufallsvariable anspricht.

6. Somit entstehen zwei Möglichkeiten für die Formulierung statistischer Hypothesen. Man kann einerseits durch einen Rückgriff auf statistische Variablen deskriptive Hypothesen formulieren; und man kann andererseits durch einen Rückgriff auf Zufallsvariablen generative Hypothesen formulieren. In beiden Fällen kann man Hypothesen auch auf „Zusammenhänge zwischen Variablen" beziehen. Damit werden wir uns in Abschnitt 5.4 beschäftigen. Hier soll zunächst nur festgehalten werden, daß es diese beiden unterschiedlichen Möglichkeiten gibt. Es sollte auch deutlich geworden sein, daß Hypothesen, für deren Formulierung

[5] Z.B. hat Tschuprow folgende Definition gegeben: „Der Begriff der zufälligen Variablen ordnet sich unter den allgemeinen mathematischen Begriff der variablen Größen als genus proximum ein, wobei als differentia specifica das Vorhandensein des Verteilungsgesetzes erscheint. Eine einziffrige Zahl ist eine diskontinuierliche Variable, welche zehn verschiedene Werte von 0 bis 9 annehmen kann. Sie wird zu einer zufälligen Variablen zehnter Ordnung, falls sie diese verschiedenen Werte mit bestimmten Wahrscheinlichkeiten annimmt." (Zit. bei Anderson 1929, S. 5)

Zufallsvariablen verwendet werden, als Hypothesen über Zufallsgeneratoren expliziert werden müssen. Es handelt sich nur um eine andere sprachliche Form für die generativen Hypothesen, wie sie im vorangegangenen Abschnitt erläutert worden sind.

5.1.3 Die Idee der Wiederholbarkeit

1. Verwendet man Begriffsbildungen und Vorstellungen der theoretischen Statistik, um generative Hypothesen als Hypothesen über Zufallsgeneratoren zu formulieren, entsteht ein Problem, auf das schon mehrfach hingewiesen worden ist: der Begriff eines Zufallsgenerators bezieht sich nicht unmittelbar auf Prozesse, die in unserer Erfahrungswelt ablaufen, sondern auf artifizielle oder rein gedanklich konstruierte *Verfahren*. Die Idee besteht natürlich darin, daß man Zufallsgeneratoren als Modelle verwenden möchte, um über reale Prozesse nachzudenken und ihren Ablauf einschätzbar zu machen. Gleichwohl sind Zufallsgeneratoren und reale Prozesse begrifflich zu unterscheiden. Wenn überhaupt ein Zusammenhang hergestellt werden kann, können reale Prozesse als Realisationen oder als Beispiele betrachtet werden, die durch die Aktivierung eines Zufallsgenerators zustande kommen.

2. Die Unterscheidung zwischen realen Prozessen und Zufallsgeneratoren ist im wesentlichen deshalb erforderlich, weil mit der Konzeption von Zufallsgeneratoren die Idee einer Wiederholbarkeit denkbar gemacht werden soll. Reale Prozesse (Ereignisse) sind nicht wiederholbar. Jeder reale Prozeß ist ein in Raum und Zeit einmaliges Vorkommnis. Dagegen ist ein Zufallsgenerator ein Verfahren, um beliebig viele Prozesse *in der gleichen Weise* erzeugen zu können. Zwar entsteht bei jeder Aktivierung eines Zufallsgenerators ein neuer, in seinem realen Ablauf einmaliger Prozeß. Aber die Konzeption eines Zufallsgenerators erlaubt es, diese Prozesse dadurch als gleich zu betrachten, daß sie auf die gleiche Weise zustandegekommen sind. Insofern liegt eine Abstraktion vor, die jedoch in spezifischer Weise vorgenommen wird. Man geht nicht von einer beliebigen Menge realer Prozesse aus und fragt, wie und wodurch sie vergleichbar gemacht werden könnten. Sondern umgekehrt wird die Abstraktion durch ein Verfahren definiert: Gleichheit von Prozessen wird dadurch definiert, daß sie durch das gleiche Verfahren erzeugt werden. Dies ist auch wichtig, um zu verstehen, was mit Wiederholbarkeit gemeint ist. Die Frage, ob sich reale Prozesse wiederholen können, hat zunächst keine klare Bedeutung. Kann man sich dagegen auf ein Verfahren beziehen, kann man dadurch erklären, was Wiederholbarkeit bedeuten soll: daß mehrere oder beliebig viele Prozesse mit dem Verfahren hervorgebracht werden können. Wiederholbarkeit wird also durch das Verfahren definiert.

3. Die zentrale Bedeutung der Idee der Wiederholbarkeit für die Wahrscheinlichkeitsrechnung wird zunächst dort sichtbar, wo ihre Grundbegriffe durch einen gedanklichen Rückgriff auf Experimente eingeführt werden (vgl. Abschnitt 4.4.3). Wesentlich ist, daß Experimente „unter den gleichen Bedingungen“ wiederholt werden können. Daß ihre Ergebnisse auch „vom Zufall abhängen“, ist eine sekundäre und eigentlich überflüssige Vorstellung. Der Anschluß an Begriffsbildungen der Wahrscheinlichkeitsrechnung erfolgt dadurch, daß man Experimente als Realisationen eines Zufallsgenerators konzipiert. Dadurch wird einerseits explizierbar, wie man bei Experimenten von einer Wiederholbarkeit sprechen möchte; insbesondere was damit gemeint sein soll, daß sie „unter den gleichen Bedingungen“ wiederholt werden können. Gemeint ist, daß man für die Durchführung von Experimenten ein Verfahren angeben kann. Weiterhin wird explizierbar, wie man sich auf die Ergebnisse eines Experiments beziehen möchte: durch die Konzeption einer Menge möglicher Ergebnisse, denen sich Wahrscheinlichkeiten zurechnen lassen. Die Frage, für welche Arten von Experimenten die Konzeption eines Zufallsgenerators sinnvoll sein kann, ist damit natürlich noch nicht entschieden. In der Wahrscheinlichkeitsrechnung stellt sich jedoch diese Frage nicht, da sie sich gar nicht mit Experimenten, sondern mit Zufallsgeneratoren beschäftigt.

4. Die Idee der Wiederholbarkeit wird auch oft dann in Anspruch genommen, wenn sich die Rhetorik auf Populationen bezieht. Beispiele wurden bereits in Abschnitt 4.3.2 angeführt. Hier zur Illustration noch eine Formulierung von A. Linder (1964, S. 36):

> „Mit Hilfe der statistischen Verfahren bearbeiten wir *Gesamtheiten von Einzelwerten*, die aus Versuchen, Beobachtungen oder statistischen Erhebungen gewonnen wurden. Den Versuch oder die Beobachtung können wir *wiederholen*; rein theoretisch betrachtet, lassen sich Versuche und Beobachtungen *unendlich oft* wiederholen. Die unendliche Gesamtheit der Ergebnisse dieser Beobachtungen oder Versuche, von denen wir annehmen, sie seien im wesentlichen unter gleichen Bedingungen zustande gekommen, nennen wir die *Grundgesamtheit*.“

Bemerkenswert ist, wie das Reden von Populationen mit einem Reden von Wiederholungen verbunden wird, denn im allgemeinen haben diese Vorstellungen unterschiedliche Konnotationen: mit dem Wort ‘Population’ wird auf eine (irgendwie) *gegebene* Menge von Dingen oder Sachverhalten verwiesen, mit der Idee der Wiederholung dagegen auf ein Verfahren, das wiederholt realisiert werden kann. Es erscheint mindestens erforderlich, zwischen zwei Arten der Wiederholbarkeit zu unterscheiden. Einerseits eine Wiederholbarkeit von Beobachtungen, die sich auf voraussetzbare Sachverhalte beziehen. Dies macht natürlich nur dann einen Sinn, wenn sich der vorausgesetzte Sachverhalt während seiner wiederholten Beobachtung nicht verändert. Das wird z.B. in der Theorie der Beobachtungsfehler vorausgesetzt. Im allgemeinen können jedoch Beobachtungen nicht wiederholt werden; oder genauer gesagt: man kann zwar beliebig viele Beobachtungen anstellen, aber man sieht dann auch unterschiedliche Dinge, und insofern handelt es sich nicht um Wiederholungen. Auch in der Stichprobentheorie ist ja die Wiederholbarkeit nur eine gedankliche Konstruktion, denn eine reale Stichprobenziehung kann nicht wiederholt werden. Die Wiederholbarkeit kommt nur als eine Denkmöglichkeit dadurch zustande, daß Auswahlverfahren als Zufallsgeneratoren konzipiert werden. — Andererseits kann man sich mit der Idee der Wiederholbarkeit auf reale Prozesse beziehen, z.B. auf Experimente. Dann muß jedoch erklärt werden, in welcher Weise man von einer Wiederholbarkeit sprechen möchte, und das verlangt eine gedankliche Bezugnahme auf ein Verfahren – oder eine Menge von Regeln –, um angeben zu können, in welcher Weise ein Prozeß als eine Wiederholung eines anderen Prozesses betrachtet werden soll.[6] Wird eine solche

[6]Andernfalls wird die Vorstellung, daß es genüge, etwas „gedanklich wiederholen“ zu können, leicht obskur. Z.B. heißt es bei W.K. Estes (1991, S. 27f): „*Population* is a more sophisticated concept. Occasionally the meaning is literatal, as in the case of the population of voters from which a sample is drawn for an opinion survey. However, more often the meaning is not literal, there being no identifiable population actually being sampled in a study. In general, the notion of a population always depends on the more basic concept of replication. Statistics are applicable to situations that are replicable, meaning repeatable in a particular sense. A replication of an experiment is a repetition with all identifiable causal factors that might affect scores either held constant or randomly varied.“ Aber dann heißt es kurz darauf: „It is important to be clear that it is not necessary that replication of a situation can be accomplished in practice, only that the concept is meaningful. For example, in the second illustration [a study involving only one subject (for example, an individual with a rare form of color blindness) with the experiment consisting of determining the accuracy of discrimination among a set of colored stimuli from a series of 100 observations], it might not be possible to obtain repeated sets of 100 observations on the subject under constant conditions because of variations in the internal state of the subject. Thus, in constructing a model for the situation, the term *population* would refer to the imaginary collection of observations that could be obtained on the subject, if an indefinite number of repetitions without changes in the experimental conditions were actually possible.“ Aber was wäre dann der Realitätsbezug eines solchen Modells?

Explikation vorgenommen, wird es dadurch auch vollständig überflüssig, von einer „Menge“ aller derjenigen Dinge oder Situationen zu sprechen, die durch Wiederholungen eines Verfahrens entstehen könnten.

5. Der Gedankengang sollte deutlich machen, daß man sich mit der Idee der Wiederholbarkeit nicht unmittelbar auf reale Prozesse (wozu auch Experimente und Beobachtungen gehören) beziehen kann, sondern daß die Idee die gedankliche Bezugnahme auf ein Verfahren (oder eine Regel oder Menge von Regeln) voraussetzt. Oder anders formuliert: Wiederholbarkeit gibt es nur im Kontext einer Theoriebildung, die reale Abläufe aus der gedanklichen Perspektive ihrer Wiederholbarkeit betrachten möchte. Zum Beispiel, wenn aus der Perspektive eines Bevölkerungsstatistikers Geburten als wiederholbare Prozesse aufgefaßt werden. Niemand wird sich ohne weiteres vorstellen können, daß sich die eigene Geburt wiederholt. Aber darum geht es auch nicht, sondern um die gedankliche Konstruktion eines Verfahrens, das es erlaubt, jede individuelle Geburt als Realisation dieses Verfahrens zu betrachten. So kann man sich klarmachen, daß Wiederholbarkeit keine Eigenschaft realer Prozesse ist – etwa so, daß sie einigen realen Prozessen zukommen, anderen nicht zukommen würde –, sondern daß Wiederholbarkeit durch eine Betrachtungsweise entsteht.

5.2 Zum Modellbegriff der Statistik

In der theoretischen Statistik wird der Hypothesenbegriff meistens im Kontext einer Konstruktion statistischer Modelle verwendet. Da der Modellbegriff oft sehr inflationär und unspezifisch verwendet wird,[7] soll zunächst erläutert werden, wie in der theoretischen (Sozial-) Statistik von Modellen gesprochen wird.

5.2.1 Zum Begriffsverständnis

1. Ausgangspunkt der Sozialstatistik sind Daten, die sich auf die Elemente einer statistischen Gesamtheit beziehen; es handelt sich um die Werte ein- oder mehrdimensionaler statistischer Variablen, die für eine bestimmte reale Gesamtheit definiert sind. Theoretische Sozialstatistik

[7]Zum Beispiel: „We all depend on models to interpret our everyday experiences. We interpret what we see in terms of mental models constructed on past experience and education.“ (Bartholomew 1995, S. 6) Dann mag es so scheinen, daß auch statistische Modelle nur eine Variante bilden (ebda.).

beginnt mit der Annahme, daß diese Daten bzw. die ihnen korrespondierenden Sachverhalte durch einen Zufallsgenerator zustande gekommen sind. In einem allgemeinen Sinn ist der Ausdruck *statistisches Modell* gleichbedeutend mit einer Annahme über einen Zufallsgenerator, der die jeweiligen Daten bzw. die ihnen korrespondierenden Sachverhalte erzeugt haben könnte.

2. Dies ist jedenfalls das dominierende Verständnis statistischer Modelle, das sich in der bisherigen Entwicklung der theoretischen (Sozial-) Statistik durchgesetzt hat. Exemplarisch sei auf folgende Ausführungen von D. R. Cox und N. Wermuth (1996) hingewiesen:

"The results of statistical analysis in one sense provide summaries of aspects of the data judged relevant for interpretation. While this important, essentially descriptive aspect should not be discarded, there are usually good reasons, for example the need to assess uncertainty, for basing analyses on explicitly probabilistic considerations. The basic assumptions of *probabilistic analyses* are as follows:

1. The data are observed values of random variables, i.e. of variables having probability distributions.
2. Reasonable working assumptions can be made about the nature of these distributions, usually that they are of a particular mathematical form involving, however, unknown constants, called parameters. We call this representation a model, or more fully a probability model, for the data.
3. Given the form of the model, we regard the objective of the analysis to be the summarization of evidence about either the unknown parameters in the model or, occasionally, about the values of further random variables connnected with the model, and, very importantly, the interpretation of that evidence.

Illustration. One of the simplest widely used models for n individuals with one explanatory variable, X, and one response variable, Y, is to suppose that observations for the n individuals in the study are such that, conditionally on all the values of X, the values of Y are independently normally distributed with mean varying linearly with each individual's value of X and with constant variance." (Cox und Wermuth 1996, S. 11f)

Die Autoren verwenden hier die Formulierung, daß die beobachteten Daten als Werte von Zufallsvariablen betrachtet werden. Das ist jedoch äquivalent zu der Annahme, daß sie durch Zufallsgeneratoren zustandegekommen sind.

3. Allerdings muß man an einer Stelle etwas aufpassen. Wenn in der statistischen Literatur von Modellen gesprochen wird, beziehen sich die Formulierungen meistens auf das Zustandekommen von *Daten*. Dieser Sprachgebrauch ist jedoch ambivalent, wie bereits im Hinblick auf die Formulierung „datenerzeugender Prozeß“ in Abschnitt 1.2 besprochen worden ist. Einerseits kann man an reale Prozesse denken, die die Sachverhalte hervorbringen, auf die sich die Daten beziehen; andererseits kann man sich auf die Beobachtungsprozesse beziehen, die die Daten liefern. Dementsprechend ist auch der Begriff eines statistischen Modells zweideutig. Einerseits kann es sich um eine Annahme über einen Zufallsgenerator handeln, der reale Prozesse repräsentieren soll, durch die soziale Sachverhalte und Ereignisse zustande kommen. Andererseits können sich Modelle auf die Beobachtungsverfahren beziehen, durch die Daten über als gegeben vorausgesetzte Sachverhalte erzeugt werden. Im letzteren Fall knüpft die Modellbildung an die Auswahlverfahren für die Datengewinnung an, die als Zufallsgeneratoren konzipiert werden, oder an Vorstellungen über das Zustandekommen von Beobachtungsfehlern. Wie schon in früheren Kapiteln mehrfach betont worden ist, gilt unser Interesse in diesem Skript ausschließlich einer theoretischen Sozialstatistik, die sich gedanklich auf Prozesse bezieht, durch die soziale Sachverhalte und Ereignisse zustande kommen, und die Zufallsgeneratoren verwendet, um solche Prozesse begrifflich zugänglich zu machen.

4. Die Vorgehensweise bei der Modellkonstruktion hängt davon ab, ob man sich auf nur eine oder auf mehrere statistische Variablen bezieht. Wenn es nur eine Variable gibt, ist die Konstruktion eines statistischen Modells im wesentlichen mit der Annahme eines Zufallsgenerators identisch, der die Werte dieser Variablen erzeugt haben könnte. Solche Modelle standen Ende des 19., Anfang des 20. Jahrhunderts im Mittelpunkt der theoretischen Sozialstatistik, haben jedoch in der späteren statistischen Sozialforschung weitgehend an Interesse verloren. Exemplarisch haben wir solche Modelle bereits bei der Besprechung der Kollektivmaßlehre Fechners kennen gelernt. Wegen ihrer historischen Bedeutung werden wir darauf in Abschnitt 5.3 noch einmal eingehen. In der zweiten Hälfte des 20. Jahrhunderts haben sich die Interessen der theoretischen Sozialstatistik zunehmend auf die Entwicklung von Modellen für zwei und mehr statistische Variablen verlagert. Exemplarisch kann man an Regressionsmodelle denken, wie sie auch in dem oben angeführten Zitat von Cox und Wermuth zur Illustration verwendet worden sind. Wenn sich die Modellbildung auf zwei oder mehr Variablen bezieht, kann man unterscheiden zwischen Modellen, die die gemeinsame Verteilung der Variablen vollständig spezifizieren, und Modellen für bedingte Verteilungen. Von der zweiten Art sind Regressionsmodelle, mit denen wir uns in Abschnitt 5.4 beschäftigen.

5.2.2 Bezugnahme auf Prozesse

1. Es sollte betont werden, daß sich Modelle der theoretischen Sozialstatistik in ihrem Modellbildungsanspruch auf Prozesse beziehen, durch die soziale Sachverhalte und Ereignisse zustande kommen. Eine wichtige Implikation ist, daß die in der statistischen Literatur vorwiegend behandelten Methoden stichprobentheoretischer Inferenz (vgl. Abschnitt 4.4.2) nicht sinnvoll anwendbar sind, um statistische Modelle zu beurteilen.

2. In der empirischen Sozialforschung, soweit sie sich mit der Konstruktion statistischer Modelle beschäftigt, wird ihr theoretischer Anspruch selten ernsthaft verfolgt. Nur wenige Autoren haben bisher versucht, den Modellbildungsanspruch wirklich auf Prozesse zu beziehen. Explizit wird ein solcher Anspruch jedoch z.B. von J. Coleman (1981) formuliert:

> "There is a distinct difference between the way I am proceeding and a more common way of proceeding in the analysis of data in social science. The more common way can be termed *statistical data analysis*, in which the method of statistical analysis is designed for a particular form of data, and, in effect, determined by that data form. [...]
>
> The approach I will outline in these chapters is somewhat different. The mathematical model is what may be termed a *model of the process*. The form of statistical analysis will be derived from, and subordinate to, the model of the process. As I will show in later chapters, a model of the same process can give rise to several different forms of statistical analysis, depending on the kind of data generated by the data-collecting method. Perhaps another way of putting the difference is to recognize that for analyzing problems of the sort that I have described earlier, there are two processes involved; one is the internal process within the unit of interest that results in its being in a given state. This is what I have called the substantive process. The second is the process by which data are generated, a process that is determined by the method of data collection. I am suggesting that the form of statistical analysis cannot depend only on the second, data-generating process, but must depend as well on the substantive process." (Coleman 1981, S. 5f)

Die Unterscheidung ist natürlich wichtig. Es sollte jedoch beachtet werden, daß sich die Modellvorstellungen, die in der Geschichte der theoretischen Sozialstatistik entwickelt worden sind, ihrer Intention nach immer schon auf die „substantive processes" gerichtet haben. Die Verwechslungen, die Coleman im Auge hat und die auch heute noch in Teilen der Literatur vorhanden sind, verdanken sich erst einer relativ späten Entwicklung, einer etwa seit Mitte der 1920er Jahre einsetzenden Entwicklung stichprobentheoretisch konzipierter Methoden statistischer Inferenz.

3. Weitere Mißverständnisse sind auch durch Fishers unglückliche Konzeption hypothetischer Populationen entstanden. Diese Rhetorik hat, worauf schon hingewiesen worden ist, sicherlich dazu beigetragen, statistische Modelle nicht systematisch auf *Prozesse* zu beziehen, durch die (soziale) Sachverhalte und Ereignisse zustande kommen. Infolgedessen kommt es auch in der statistischen Literatur immer wieder zu Formulierungen der folgenden Art:

> "*Statistical models.* Supposedly identical units made and used under the same conditions usually have different values of performance, dimensions, life, etc. Variability of such a performance variable is inherent in all products, and it is described by a statistical model or distribution.
>
> *Population and sample.* A statistical model describes some *population*. A manufacturer of fluorescent lamps is concerned with the future production of a certain lamp – an essentially infinite population. [...] To obtain information, we use a *sample* (a set of units) from the population. We analyse the sample data to get information on the underlying population distribution or to predict future data from the population." (Nelson 1982, S. 5f)

Vermeidet man die obskure Vorstellung einer hypothetischen Population, die aus bloß potentiell existierenden Dingen besteht, die es tatsächlich nicht gibt, wird deutlich, daß sich Modellbildung vielmehr auf einen Prozeß beziehen müßte, durch den Sachverhalte entstehen.[8] Dann wird allerdings auch klar, daß ein statistisches Modell nicht als eine Beschreibung

[8] Hier noch einmal ein Zitat, das zeigt, daß auch Fisher an Hypothesen über solche Prozesse interessiert war: „A very important feature of inductive inference, unknown in the field of deductive inference, is the framing of the hypothesis in terms of which the data are to be interpreted. This hypothesis must fulfil several requirements: (i) it must be in accordance with the facts of nature as so far known; (ii) it must specify the frequency distribution of all observational facts included in the data, so that the data as a whole may be taken as a typical sample; (iii) it must incorporate as parameters all constants of nature which it is intended to estimate, in addition possibly to special, or *ad hoc*, parameters; (iv) it must not be contradicted, *in any way judged relevant*, by the data in hand. If it satisfies these conditions it is therefore a scientific construct of a fairly elaborate type. It is by no means obvious that different persons should not put forward different successful hypotheses, among which the data can supply little or no discrimination. The hypothesis is sometimes called a model, but I should suggest that the word model should only be used for aspects of the hypothesis between which the data cannot discriminate." (Fisher 1955, S. 75) Ein statistisches Modell wird hier also als eine Hypothese über „die Natur" aufgefaßt, die die Daten – d.h. die Sachverhalte, auf die sich die Daten beziehen – hervorbringt. Fishers Vorschlag, den Modellbegriff auf solche Annahmen einzuschränken, die empirisch nicht geprüft

aufgefaßt werden kann, und zwar weder einer Population noch eines Prozesses. Stattdessen kann man davon sprechen, daß ein Modell eine begriffliche Vorstellung für potentiell realisierbare Prozesse verfügbar machen soll.[9]

5.3 Einfache Verteilungsmodelle

Historisch betrachtet begann Modellbildung in der theoretischen Sozialstatistik damit, theoretische Verteilungen für einzelne statistische Variablen zu konstruieren. Darauf soll in diesem Abschnitt etwas näher eingegangen werden.

5.3.1 Der theoretische Anspruch

1. Die Idee statistischer Modelle für Häufigkeitsverteilungen statistischer Variablen ist im Prinzip sehr einfach: gesucht ist ein Zufallsgenerator, der die Annahme erlaubt, daß die Werte einer statistischen Variablen durch ihn erzeugt sein könnten. Da der Zufallsgenerator durch ein Verteilungsgesetz charakterisiert werden kann (vgl. Abschnitt 2.1), kann man glauben, die beobachteten Erscheinungen würden durch dieses Verteilungsgesetz bestimmt. In den Grundzügen haben wir diesen Gedankengang bereits in Kap. 2 besprochen.

werden können, unterstreicht einerseits den hypothetischen Charakter der Modellbildung, ist jedoch andererseits fragwürdig. Denn es hängt vom vorausgesetzten Modell ab, welche Schlußfolgerungen aus Daten gezogen werden können.

[9] Da statistische Modelle nicht als Beschreibungen realer Vorgänge verstanden werden können, stellt sich natürlich in jedem Fall die Frage, welche empirisch explizierbaren Gesichtspunkte für die Modell*konstruktion* verwendet werden können, eine Frage, die *innerhalb* der statistischen Methodenlehre nicht behandelt wird. A. P. Dempster (1987, S. 2) bemerkte: „the problem of where models come from“ sei „the ‘missing half’ of statistics“; vgl. auch Mallows, 1998. Tatsächlich wird diese Frage in der Literatur zur theoretischen Statistik meistens nicht behandelt; z.B. schreiben Cox und Hinkley (1974, S. 2): „In the theoretical discussion, it will be usual to take a family of models as given.“ Die Frage „where models come from“ verweist indessen in der Sozialstatistik auf eine besondere Problematik, da in diesem Fall schon die Vorstellung eines „sozialen Prozesses“ so unklar ist, daß es für statistische Begriffsbildungen gewissermaßen keinen Ansatzpunkt gibt. Die Prozesse, auf die sich sozialstatistische Modelle scheinbar beziehen, sind infolgedessen *wesentlich* fiktiv. D. Freedman hat in einem Diskussionsbeitrag (1997a, S. 79) zunächst folgende Aussage von C. Achen (1987) zitiert: „No sensible social scientist believes any particular specification, coefficient estimate, or standard error. Social science theories ... imply that specifications and parameters constant over situations do not exist ... One searches for qualitative theory ... not for quantitative specifications.“ Dann macht er dazu folgende Bemerkung: „With Hooke's law [ein Beispiel aus der Physik, auf das sich Freedman zunächst bezogen hat] and the like, we are estimating parameters in specifications that are constant across time – at least to a very good degree of approximation. What are the social scientists doing when they estimate non-existing parameters, and put standard errors on the output? How can that help them search for qualitative theory? Those are among the first of my questions, and I never get answers.“ Man vgl. auch die Diskussion zwischen Freedman (1985) und Fienberg (1985) sowie die Diskussionsbeiträge in Shaffer, 1992.

2. Der Gedankengang hat allerdings eine fundamentale Schwäche: die Zufallsgeneratoren sind nicht nur rein hypothetische Konstrukte, mehr noch, sie werden aus den je vorliegenden Daten konstruiert. Sie sind also nur Deutungsschema für die Prozesse, die die Sachverhalte, auf die sich die jeweils vorliegenden Daten beziehen, hervorgebracht haben könnten. Diese Tatsache hat auch zu einer Reihe von Diskussionen unter den Statistikern geführt, die sich Ende des 19., Anfang des 20. Jahrhunderts hauptsächlich mit der Erfindung von Verteilungsgesetzen beschäftigt haben. Zum Beispiel hat C. V. L. Charlier ein System von Verteilungsfunktionen, das von Karl Pearson entwickelt worden ist, in seinen „Grundzügen der mathematischen Statistik“ (1920) folgendermaßen kritisiert:

„Die Darstellung der mathematischen Statistik, die hier gegeben wird, basiert nicht direkt auf den Pearsonschen Untersuchungen [...]. Ohne eine eingehende Kritik dieser Untersuchungen, die ich übrigens auf höchste bewundere, vornehmen zu wollen, glaube ich hier doch hervorheben zu müssen, daß die Pearsonschen Methoden einen wesentlichen Fehler besitzen, der dadurch charakterisiert werden kann, daß sie keine genügende Allgemeinheit in der Wahl des Ausgangspunktes wie auch in der praktischen Anwendung besitzen. Seine 7 Fehlergesetztypen sind ohne Zweifel vortreffliche Interpolationsformeln, sie sind aber ohne Rücksicht darauf abgeleitet, wie solche Fehlergesetze genetisch entstehen können.“ (Charlier 1920, S. 6)

Im letzten Satz kommt recht gut die theoretische Idee zum Ausdruck, die mit der Entwicklung von Verteilungsgesetzen verfolgt worden ist: daß solche Gesetze zeigen sollen, wie beobachtbare Häufigkeitsverteilungen entstehen. Die Idee ist, daß man durch eine Konstruktion von Verteilungsgesetzen etwas über den Entstehungsprozeß von Erscheinungen lernen könne. Meistens verbindet sich diese Betrachtungsweise mit Vorstellungen über Verursachungszusammenhänge. Zum Beispiel hat J. C. Kapteyn (1903) ein System von Verteilungsgesetzen entwickelt (im wesentlichen ein System log-normaler Verteilungsfunktionen), das er u.a. mit folgender Feststellung zu begründen versucht:

“It assigns the connection between the form of the curves and the action of the causes to which the form is due. The knowledge of the connection may lead us, at least in some cases, to precious indications about the nature of true causes."[10]

Dieser Betrachtungsweise folgen auch noch J. Aitchison und J. A. C. Brown in ihrer Monographie über die Lognormal-Verteilung (1957):

"It has occasionally been argued that the success with which any particular frequency curve graduates empirical data is a sufficient criterion of its worth. There are, indeed, many occasions on which such an argument may be satisfactory. Yet there are at least two important reasons for seeking a more fundamental basis in the theory of probability for any system of frequency curves to which we attach a more than transient importance. First by providing such a basis we may obtain a clearer insight into underlying natural or sociological processes; this in its turn will often suggest a wider application of the system. Secondly, a knowledge of the elementary assumptions from which the law of frequency may be derived will enable us more easily to modify the law to meet new circumstances." (Aitchison und Brown 1957, S. 20)

Mehr oder weniger deutlich ausgesprochen findet man bei allen Autoren, die in der Konstruktion von Verteilungsgesetzen für empirische Häufigkeitsverteilungen eine sinnvolle statistische Aufgabe erblickt haben, die Hoffnung oder den Anspruch, dadurch etwas über den Entstehungsprozeß der Erscheinungen, auf die sich die empirischen Häufigkeitsverteilungen beziehen, lernen zu können.

5.3.2 Hypothetische Elementarursachen

1. Was ist von diesem theoretischen Anspruch zu halten? Der Hauptgrund, der ihn fragwürdig erscheinen läßt, liegt aus meiner Sicht darin, daß hypothetische Zufallsgeneratoren keine Einsichten in Verursachungszusammenhänge vermitteln können. Bevor dieses Argument genauer besprochen wird, soll dargestellt werden, wie die Konstrukteure von Verteilungsgesetzen in gewisser Weise nur Gedankengänge der Theorie der Beobachtungsfehler übernommen und mit einer neuen Rhetorik überzogen haben. Dies gilt jedenfalls für die Mehrzahl derjenigen Statistiker, die Verteilungsfunktionen nicht nur als ein Darstellungsmittel für empirische Häufigkeitsverteilungen aufgefaßt, sondern sie als Verteilungsgesetze mit einem Erklärungsanspruch konzipiert haben.[11]

[10] Zitiert nach Aitchison und Brown 1957, S. 20.

[11] Es sei angemerkt, daß es auch Bemühungen um eine Begründung von Verteilungsgesetzen gegeben hat, bei denen nicht von Vorstellungen der Theorie der Beobachtungsfehler ausgegangen worden ist. Als Beispiel sei auf B. Gompertz (1779–1865) hingewiesen, der ein Verteilungsgesetz für menschliche Lebensdauern finden wollte. Den Gedankengang kann man sich anhand folgender Bemerkungen von E. Blaschke (1906, S. 160f) verdeutlichen:

„Nach *Gompertz* kann man sich den Tod durch eine bestimmte Ursache 'die *Todeskraft* (force of mortality)' herbeigeführt denken und diese *letztere* als Funktion des Alters zu bestimmen suchen. Werden die Todeskraft für das $x$te Lebensjahr durch $\varphi(x)$, die Personen des Alters x durch l_x bezeichnet, so drückt $l_x\varphi(x)dx$ die Anzahl der zwischen dem Alter x und $x + dx$ Sterbenden $= -dl_x$ aus. Sohin besteht

$$\frac{-dl_x}{l_x} = \varphi(x)dx.$$

Behufs Bestimmung von $\varphi(x)$ nimmt Gompertz an, daß die Todeskraft der Abnahme des *Widerstandes* des menschlichen Organismus gegen die Zerstörung gleich sei und dieser Widerstand am Ende eines jeden gleich großen, unendlich kleinen Altersintervalles stets den gleich vielten Teil des jedesmaligen Restes verliere."

Je nachdem wie $\varphi(x)$ spezifiziert wird, gelangt man zu unterschiedlichen Verteilungsfunktionen. Gompertz' Vorschlag war, $\varphi(x) = ab^x$ anzunehmen. Die Details sind hier jedoch unwichtig. Bemerkenswert ist vielmehr der rein rhetorische Gebrauch des Wortes 'Ursache'. Es ist evident, daß man mit Gompertz' Formel keine Einsichten in Todesursachen gewinnen kann.

2. Bemerkenswert ist zunächst eine Unterscheidung zwischen qualitativen und quantitativen Merkmalsräumen statistischer Variablen. Wenn es sich um Variablen mit qualitativen Merkmalsräumen handelt, kann man unmittelbar von Häufigkeitsverteilungen sprechen; und man kann weiterhin Zufallsgeneratoren, die die Werte der Variablen hervorbringen könnten, auf einfache Weise durch ein Urnenschema konzipieren. Wenn man dagegen bei quantitativen Merkmalsräumen von einer Vorstellung beliebig teilbarer Größen ausgeht und für ihre numerische Repräsentation die reellen Zahlen verwendet, kann man nicht mehr unmittelbar von der relativen Häufigkeit oder Wahrscheinlichkeit numerisch genau bestimmter Merkmalswerte sprechen. In vielen Lehrbüchern wurden deshalb zwei Teilgebiete der Statistik unterschieden. Charlier (1920) hat eine Terminologie eingeführt, die später auch von einigen anderen Autoren übernommen worden ist.[12] Er spricht einerseits von einer *homograden Statistik*, die sich mit „homograden Merkmalen" beschäftigt, d.h. mit statistischen Variablen, deren Merkmalsräume eine endliche Anzahl qualitativ unterschiedlicher Attribute zusammenfassen; als Spezialfall treten Indikatorvariablen auf, die das Eintreten eines Ereignisses erfassen. Davon wird eine *heterograde Statistik* unterschieden, die sich mit „hetero-

[12] Z.B. von Anderson 1935, S. 25; Anderson 1957; Kellerer 1960, S. 116; Most 1962, S. 47ff.

graden Merkmalen“ beschäftigt, d.h. mit statistischen Variablen, deren Merkmalsräume metrische und im Prinzip beliebig teilbare Größen repräsentieren.

3. Ein Grund für diese Unterscheidung liegt darin, daß nur im Bereich qualitativer Merkmale – indem man sich auf ihre relativen Häufigkeiten bezieht – einfache Modellvorstellungen für Zufallsgeneratoren, z.B. in Form des allgegenwärtigen Urnenschemas, gebildet werden können. Charlier hat das so gesagt:

„Während in der homograden Statistik die Gruppierung der Elemente der statistischen Reihe um das Medium in völlig befriedigender Weise durch die Theoreme von Bernoulli, Poisson und Lexis erklärt werden kann, liegen die Dinge vollkommen anders, wenn es sich um heterograde statistische Reihen handelt. Keine *mathematischen* Argumente gestatten beispielsweise die mittleren Temperaturabweichungen in Lund für den Monat Mai vorauszusagen, und keine apriorischen Schlüsse sind hinreichend, um auch nur näherungsweise die Dispersion einer statistischen Reihe, betreffend die Länge erwachsener Männer (einer gewissen Rasse), zu schätzen.“ (Charlier 1920, S. 58f)

In dieser Formulierung ist die Gegenüberstellung zwar etwas merkwürdig; denn auch die Erscheinungen, mit denen sich die homograde Statistik beschäftigt, können natürlich nicht auf mathematische Weise erklärt werden, etwa durch die Unterstellung eines Urnenmodells. Die Frage läuft vielmehr darauf hinaus, wie in der heterograden Statistik Modellvorstellungen für Zufallsgeneratoren gebildet werden können, die die beobachteten Werte einer statistischen Variablen erzeugt haben könnten.

4. Der Weg zur Beantwortung dieser Frage war indessen schon durch die Anwendungen der Wahrscheinlichkeitsrechnung in der Theorie der Beobachtungsfehler vorgezeichnet. Es wurde schon erwähnt, wie Quetelet und dann auch andere Autoren, wie z.B. Fechner, an das „Fehlergesetz“ angeknüpft haben. Auch Charlier knüpft an diese Vorstellungen an. Im Anschluß an seine eben zitierten Hinweise auf Schwierigkeiten bei heterograden Merkmalen heißt es:

„Man wird hier gezwungen, zu verwickelteren Erklärungsgründen zu greifen. Der erste Schritt ist der, eine hinreichend allgemeine Hypothese aufzusuchen, die zu erklären imstande wäre, wie die Abweichung, die ein Individuum der Population mit Hinsicht auf das zu untersuchende Attribut zeigt, entstanden sein kann. Die Hypothese, die ich meinen Untersuchungen zugrunde gelegt habe,[13] ist die, daß *jede individuelle Abweichung als Summe einer Menge zwar unbekannter, aber kleiner Größen, die Elementarfehler genannt werden, betrachtet werden kann.*“ (Charlier 1920, S. 59)

[13][Charlier macht hier folgende Anmerkung:] Die Hypothese als solche ist nicht neu, sondern wurde, wenn auch in etwas speziellerer Form, schon im Jahre 1837 von Hagen aufgestellt, der hieraus das sog. Gaußsche Fehlergesetz ableitete.

Dieses Zitat macht nicht nur deutlich, wie Charlier an Gedankengänge der Theorie der Beobachtungsfehler anknüpft, sondern zeigt auch, daß er Zufallsgeneratoren konstruieren möchte, die sich auf den Entstehungsprozeß quantitativer Größen beziehen. Die Hypothese wird mit folgendem Beispiel erläutert:

„Man betrachte beispielsweise die Körperlänge in einer Population von erwachsenen Männern. Wenn für alle Männer der Population angenommen werden kann, daß sie vollkommen gleichartige Vorfahren besitzen, genau dieselbe Erziehung, Ernährung erhalten haben, denselben klimatischen Einflüssen ausgesetzt worden sind, daß alle anderen Umstände, welche einen Einfluß auf die Körperlänge des Menschen haben können, identisch dieselben für alle Männer der Population gewesen sind, so müßte man den Schluß ziehen, daß die Körperlänge aller dieser Männer dieselbe sein würde, ebenso sicher, wie die Wirkung aus ihrer Ursache bestimmt ist. Die Verschiedenheiten in der Abstammung,[14] Erziehung, Ernährung usw. können hier als ebenso viele *Fehlerquellen* hinsichtlich der Körperlänge dieser Männer betrachtet werden. Jede Fehlerquelle verursacht einen positiven oder negativen *Elementarfehler* in der Länge. Durch eine Addition aller dieser kleinen Größen entsteht die resultierende Abweichung der Länge eines Individuums von der idealen Länge, die alle haben würden, wenn sie denselben Einflüssen ausgesetzt gewesen wären. Offenbar muß die Anzahl der Fehlerquellen als sehr groß oder unendlich betrachtet werden.“ (Charlier 1920, S. 59f)

Schließlich verwendet Charlier seine Hypothese der Elementarfehler, um folgende Aufgabe zu formulieren:

„Aus der Hypothese der Elementarfehler die Gesetze abzuleiten, die für die Verteilung der Elemente in einer heterograden statistischen Reihe gelten, ist eine Aufgabe, welche in das Gebiet der mathematischen Wahrscheinlichkeitstheorie fällt.“ (Charlier 1920, S. 60)

5. Charliers Erläuterungen zeigen, daß man anstelle von „Elementarfehlern“ auch von „Elementarursachen“ sprechen kann. Daß von *Fehlern* gesprochen wird, ist zunächst ein Tribut an die Theorie der Beobachtungsfehler, in der die Begriffsbildungen und Vorstellungen entwickelt worden sind. Es kommt darin aber auch zum Ausdruck, daß sich die

[14][Anmerkung von Charlier:] Was hier kurz Abstammung genannt wird, umfaßt natürlich eine unendliche Menge von Fehlerquellen: ein Versuch, diese hier näher zu gruppieren, ist überflüssig.

Theoriebildung nicht wirklich auf Ursachen bezieht. Die rhetorische Bezugnahme auf Ursachen bleibt rein metaphorisch. Insofern kann das Ziel auch nicht darin bestehen, etwas über Verursachungszusammenhänge herauszufinden. Bestenfalls kann man davon sprechen, daß die Hypothese der Elementarfehler dazu dienen soll, beobachtbare *Unterschiede* zwischen Merkmalswerten spekulativ erklärbar zu machen. Hier wird indessen eine weitere Konnotation aus der Theorie der Beobachtungsfehler wirksam. In dieser Theorie erscheint es sinnvoll, vorauszusetzen, daß die einzelnen Messungen Abweichungen von einem in der Realität bestimmten Wert darstellen. In der Sozialstatistik macht jedoch diese Vorstellung keinen Sinn. Nicht nur kommen zum Beispiel unterschiedliche Körpergrößen nicht durch Meßfehler zustande; ihre Unterschiede können überhaupt nicht als Abweichungen von einem „Idealwert“ oder einer „Norm“ interpretiert werden. Indem die Sozialstatistiker den Vorstellungen und Begriffsbildungen der Theorie der Beobachtungsfehler gefolgt sind, ist es kaum vermeidbar, daß sie *Unterschiede* zwischen beobachteten Merkmalswerten als *Abweichungen* von einem „Idealwert“ konzipieren. So z.B. in dem oben angeführten Zitat, in dem Charlier von einer „idealen Länge“ spricht. Das Problem besteht indessen nicht in erster Linie darin, daß Formulierungen dieser Art in einer offensichtlich fragwürdgen Weise „Normen“ und „Normalitätsvorstellungen“ suggerieren;[15] daß dies nicht unbedingt gemeint sein muß, zeigen z.B. folgende Ausführungen von Arne Fisher:

> “According to the Laplacean–Hagen–Charlier theory every variate or individual deviation from a certain norm is generated as the sum of a mass of small and unknown quantities – generally infinite in number – which are known as elementary errors (deviations). The word error must of course be taken in a different sense than that we usually associate with the word. In precision measurements we are actually dealing with true or natural errors arising from imperfections of the instruments and the observer, but it would of course not be right to regard a deviation of say 5 centimeters from the mean stature of a population group as an error in the usual sense of the word. Used in its wider sense as an expression for deviations the term will, however, be readily understood, and it is in this sense we shall use it in the following pages.” (A. Fisher 1928, S. 172)

Als Frage bleibt aber, mit welcher Berechtigung empirisch nur als unterschiedliche Werte gegebene Beobachtungen als *Abweichungen von* …

[15] Man vgl. die Besprechung der Ausführungen von Lexis in Abschnitt 2.5.1.

konzipiert werden können.[16]

6. Das führt jedoch zu einer grundsätzlicheren Frage: Welchem Erkenntniszweck soll die Hypothese der Elementarfehler dienen? Zwar kann sie sicherlich verwendet werden, um *fiktive* Zufallsgeneratoren zu konstruieren, durch die man sich das Zustandekommen der jeweils beobachteten Werte einer Variablen vorstellbar machen kann. Aber warum könnte es sinnvoll erscheinen, sich zu diesem Zweck der Hypothese der Elementarfehler zu bedienen? Man könnte es ja viel einfacher haben, indem man für jede statistische Variable unmittelbar einen Zufallsgenerator annimmt, der ihre Werte erzeugt, und zwar mit der jeweils beobachteten oder daraus durch eine „Idealisierung“ gebildeten Verteilung. Sei zum Beispiel eine statistische Variable

$$X : \Omega \longrightarrow \tilde{X}$$

gegeben. Man könnte annehmen, daß es einen Zufallsgenerator gibt, der sich durch eine Zufallsvariable X^* repräsentieren läßt, so daß die empirisch ermittelten Werte von X als Realisationen der Zufallsvariablen X^* aufgefaßt werden können. Warum könnte es sinnvoll erscheinen, sich außerdem eine Reihe fiktiver Zufallsvariablen

$$X_1^*, X_2^*, X_3^*, \ldots, X_n^*, \ldots$$

auszudenken und dann X^* als eine Funktion dieser fiktiven Zufallsvariablen darzustellen?

[16] Wie in Abschnitt 2.5.1 schon besprochen worden ist, gab es mannigfache Spekulationen über den Charakter der „Mitte“, die benötigt wird, um sinnvoll von „Abweichungen“ sprechen zu können. Hier zur Illustration noch eine Ausführung von Edgeworth (1908, S. 653): „In the classical treatment of errors-of-observation the ‘probable error’ is ancillary to the measurement of a real quantity. The relation is not very different in the problems which occupied Quetelet – the search for *l'homme moyen* and like entities – though here, as Mr. Galton has recently reminded us, the object is not, as before, get rid of errors. ‘These errors of deviation’, he says, ‘were the very things I wanted to preserve and know about’. The relation of the constituents to their mean has been compared to the relation, which has so much exercised philosophers, between the members of a class and the essence or definition thereof. The typical value of a set or ‘series’ of objects varying in quantity is not indeed like the ‘first substances’ of mediæval metaphysics, stuck into all the individual members of the group, not certainly into those which are below the mean – sovereigns, for instance, which are under 123.274 grains, if that is the mean weight, as it is the standard, of newly-coined sovereigns. Yet that typical weight might perhaps be regarded as part of that unanalysable general resemblance which, according to the modern logician, goes to make up the definition of a general term.“

7. Die Hypothese der Elementarfehler suggeriert, daß hierdurch ein Erkenntnisfortschritt erzielt werden könnte. Aber eine bloß metaphorische Interpretation der fiktiven Zufallsvariablen $X_1^*, X_2^*, X_3^*, \dots$ als Repräsentanten für in der Realität wirksame Elementarursachen ändert am fiktiven Charakter des Gedankengangs nichts. Nicht nur erhält man keine Einsichten in die Ursachen, die möglicherweise die beobachteten Erscheinungen hervorgerufen haben. Selbst wenn es gelänge, den Zufallsvariablen $X_1^*, X_2^*, X_3^*, \dots$ eine empirisch explizierbare Bedeutung zu geben, bliebe eine wichtige Frage immer noch offen: ob und wann die Elementarursachen überhaupt eine empirisch relevante Aktivität entfalten. Das Problem wurde bereits in Abschnitt 4.2.2 besprochen. Ein Zufallsgenerator, gleichgültig ob artifiziell oder hypothetisch, erklärt nicht das tatsächliche Auftreten der Sachverhalte oder Ereignisse, die er hervorbringen *könnte*. Zwar kann die Kenntnis eines Zufallsgenerators unter Umständen gleichwohl eine Einsicht in die Beschaffenheit des Prozesses vermitteln, der gewisse Erscheinungen hervorbringt (*wenn* er sie hervorbringt, also z.B. *wenn* mit einem Würfel tatsächlich geworfen wird). Aber liefert diese Überlegung einen Ansatzpunkt, um auch die Hypothese der Elementarfehler als Ausgangspunkt für einen potentiellen Erkenntnisgewinn auffassen zu können?

8. Man kann Beispiele konstruieren, für die sich eine positive Antwort geben läßt. Exemplarisch kann man an das sog. Galton-Brett denken, ein von Francis Galton ausgedachter Mechanismus zur Illustration eines Zufallsgenerators für eine symmetrische Binomialverteilung.[17] Dadurch kann die Darstellung einer binomialverteilten Zufallsvariablen durch eine Summe einfacher Zufallsvariablen veranschaulicht werden. Aber das Argument verläuft in diesem Beispiel in umgekehrter Richtung: es wird deutlich, wie man Begriffsbildungen der Wahrscheinlichkeitstheorie verwenden kann, um ein Verständnis für einen empirisch demonstrierbaren Mechanismus zu gewinnen. Für die in der theoretischen Sozialstatistik intendierten Anwendungen der Wahrscheinlichkeitstheorie liegt das Problem indessen darin, daß es schwerfällt oder sogar unmöglich ist, in empirisch explizierbarer Weise auf Prozesse hinzuweisen, deren Ablauf verständlich gemacht werden soll. Der hypothetisch unterstellte Zufallsgenerator wird dann nur als ein Deutungsschema verwendet.

9. Dann aber wird fragwürdig, ob die Hypothese der Elementarfehler in irgendeiner Weise zu einer Einsicht führen kann. Denn für sich

[17] Vgl. Galton 1889, S. 63ff. Durch einfache Modifikationen kann man auch nichtsymmetrische Verteilungen erzeugen.

genommen folgt aus dieser Hypothese überhaupt nichts. Die Aufgabe, wie sie z.B. von Charlier konzipiert worden ist, würde darin bestehen, eine Funktion g zu finden, so daß man eine Darstellung

$$X^* = g(X_1^*, X_2^*, X_3^*, \dots, X_n^*, \dots)$$

bekommt. Man könnte dann sagen, daß diese Darstellung zeigt, wie der Zufallsgenerator funktioniert, der durch die Zufallsvariable X^* repräsentiert wird. Wenn jedoch die Zufallsvariablen $X_1^*, X_2^*, X_3^*, \dots$ rein fiktiver Natur sind, können daraus auch keine Anhaltspunkte für die Konstruktion der Funktion g gewonnen werden. Der einzige empirisch verfügbare Anhaltspunkt kommt vielmehr durch die Annahme zustande, daß es sich bei den Werten der *statistischen* Variablen X zugleich um Realisationen der Zufallsvariablen X^* handelt. Schon diese, aber dann erst recht alle weiteren Annahmen über die interne Beschaffenheit des Zufallsgenerators für die Zufallsvariable X^* sind rein spekulativer Natur.[18] Infolgedessen erschließt die Hypothese der Elementarfehler nur ein Betätigungsfeld für die Mathematik. Man kann der Phantasie freien Lauf lassen, um beliebige Verteilungsgesetze zu konstruieren; und man kann

[18] Einen besonders eklatanten Ausdruck findet dies in der Konstruktion von Mischverteilungsmodellen, etwa in der Form von „Varianzkomponentenmodellen“, wie sie in der gegenwärtigen statistischen Sozialforschung eine gewisse Beliebtheit gewonnen haben. Die Idee ist indessen keineswegs neu und findet sich auch schon bei Charlier (1920, S. 77f): „Eine Frage, über welche die Diskussion kaum als eröffnet angesehen werden kann, ist die, wie man die Abweichungen einer Frequenzkurve von der normalen Form deuten [!] soll. Daß die Deutung letzten Endes mit der Beschaffenheit der elementaren Fehlerquellen zusammenhängt, ist offenbar. Die Schwierigkeit liegt aber darin, daß die Lösung des Problems *vieldeutig* ist, ja, daß sogar unendlich viele Lösungen vorhanden sind. [...] Wenn man, die Terminologie Quetelets beibehaltend, ein von einer *normalen* Frequenzkurve charakterisiertes statistisches Objekt einen *Typus* nennt, so kann man im allgemeinen sagen, daß ein statistisches Objekt mit *nichtnormaler* Frequenzkurve (also mit bestimmter Schiefheit, Exzeß usw.) als *aus mehreren Typen zusammengesetzt* angesehen werden kann. Man nehme zum Beispiel an, daß die Körperlänge der 14jährigen Schulknaben in Schweden einen gewissen *Typus* darstellt und daß dies auch mit der Länge der 15jährigen Schulknaben der Fall ist. Würde man nun ein statistisches Objekt in der Weise bilden, daß man aufs Geratewohl 1000 14jährige und 1000 15jährige Schulknaben nähme, so würde die Frequenzkurve für diese 2000 Individuen nicht länger normal sein. [...] Jedes statistische Objekt, das eine Mischung einer größeren Anzahl von Typen ist, wird in dieser Weise im allgemeinen eine Frequenzkurve von *nichtnormaler* Form besitzen. Man kann hieraus verstehen, daß die Normalkurve (die Gaußsche Kurve) nicht Regel, sondern Ausnahme in der Statistik sein dürfte.“ Aber selbst der letzte Satz ist fragwürdig, weil dabei schon vorausgesetzt wird, daß natürliche oder soziale Prozesse so funktionieren, wie es durch die Hypothese der Elementarfehler vorstellbar gemacht werden soll.

sich beliebig viele „Mechanismen“ ausdenken, die Zufallsvariablen als Funktionen anderer Zufallsvariablen darstellbar machen.

5.3.3 Darstellung oder Erklärung?

1. Mit den begrifflichen Hilfsmitteln der Mathematik können allerdings beliebig viele und beliebig unterschiedliche Verteilungsgesetze konstruiert werden. In der theoretischen Statistik geschieht dies typischerweise so, daß zusammengehörige Familien von Verteilungsfunktionen konzipiert werden, die sich nur durch einige wenige Parameter unterscheiden. Als Beispiel kann man an die Familie der Normalverteilungsfunktionen denken, deren Mitglieder sich durch zwei Parameter unterscheiden, die sich auf den Mittelwert und die Varianz der Verteilung beziehen. Soweit man die Suche nach Verteilungsgesetzen rein formal betrachtet, besteht sie einfach darin, eine möglichst große Vielfalt von Familien parametrischer Verteilungsfunktionen auszudenken. Wie schon ausgeführt worden ist, spielten solche Bemühungen in der Frühzeit der theoretischen Statistik eine große Rolle, gelegentlich wurde darin sogar eine Hauptaufgabe erblickt:

> “The fundamental problem of mathematical statistics is to develop an analytical expression that is capable of representing any distribution of magnitude-determinations, *i.e., variates*. The Gaussian or Normal Curve of Error, Pearson’s Generalized Probability Curves, and Charlier’s Type A and Type B curves are all the results of investigations of this fundamental problem.” (Carver 1921, S. 721)

2. Sich hieran anschließend besteht die statistische Modellkonstruktion – oder genauer: die Konstruktion parametrischer Modelle – zunächst aus zwei Schritten:

a) Es wird eine Familie parametrischer Verteilungsfunktionen ausgedacht oder aus schon vorhandenen Angeboten ausgewählt; dieser Schritt wird oft als „Modellspezifikation“ bezeichnet.[19]

[19] Rao (1995, S. 128f) sagt dazu: „Spezifikation ist die Wahl eines stochastischen Modells, bezüglich dessen die beobachteten Daten analysiert werden, um bestimmte Fragen zu beantworten oder um ein bestimmtes Element des Modells auszuwählen zwecks Vorhersage zukünftiger Beobachtungen.“ Die Bedeutung einer richtigen Spezifikation wird von vielen Autoren betont; z.B. von G. Arminger (1990, S. 253) auf folgende Weise: „A crucial issue in the interpretation of results from fitting statistical models to empirical data is the underlying but rarely voiced assumption that the class of models fitted to the data contains the true model of the random mechanism generating the data. If this is not the case, the parameters estimated from a wrong model can easily be heavily biased and lead to wrong conclusions about the nature of relationships and causal mechanism.“ Aber wie kann ein „true model“ bestimmt werden, wenn es sich auf einen rein fiktiven Zufallsgenerator bezieht?

b) Dann wird mit irgendeinem Rechenverfahren dasjenige Mitglied der vorgegebenen Menge möglicher Modelle ausgesucht, das „am besten zu den jeweils verfügbaren Daten paßt“. Bei parametrisch vorgegebenen Verteilungsfamilien besteht diese Aufgabe darin, bestimmte Werte für die Parameter zu finden, und man spricht insofern auch von einer *Schätzung* der Parameter mithilfe der verfügbaren Daten.[20]

Dann stellt sich die Frage, wie die Schätzung vorgenommen werden soll, d.h., was damit gemeint sein soll, daß ein bestimmtes Mitglied der Verteilungsfamilie „am besten zu den Daten paßt“.

3. Eine Antwort auf diese Frage hängt davon ab, welchen Sinn man mit der Modellkonstruktion verbindet. Zwei Betrachtungsweisen können unterschieden werden. Eine Betrachtungsweise verzichtet auf einen theoretischen Anspruch. In diesem Fall dient eine parametrische Verteilungsfunktion nur dem Zweck, eine vorgegebene Häufigkeitsverteilung darzustellen. Das kann z.B. sinnvoll sein, wenn man die Unterschiede in einer Reihe empirischer Häufigkeitsverteilungen durch einige wenige Parameter charakterisieren möchte. Parametrische Verteilungsfunktionen dienen dann als ein Hilfsmittel der Datenanalyse. Aber es gibt keinen weitergehenden theoretischen Anspruch; insbesondere wird die zur Darstellung einer empirischen Häufigkeitsverteilung verwendete Verteilungsfunktion nicht als ein Verteilungsgesetz interpretiert, das in irgendeiner Weise Einsichten in das Zustandekommen der Daten vermitteln könnte. Infol-

[20] Dies entspricht der Formulierung der Aufgaben einer theoretischen Statistik, wie sie von Fisher in seiner bereits erwähnten Abhandlung „On the Mathematical Foundations of Theoretical Statistics“ (1922) gegeben wurde. Abschnitt 3 der Abhandlung trägt den Titel „The Problems of Statistics“ und beginnt mit folgenden Worten: „The problems which arise in reduction of data may be conveniently divided into three types: (1) Problems of specification. These arise in the choice of the mathematical form of the population. (2) Problems of estimation. These involve the choice of methods of calculating from a sample statistical derivates, or as we shall call them statistics, which are designed to estimate the values of the parameters of the hypothetical population. (3) Problems of distribution. These include discussions of the distribution of statistics derived from samples, or in general any functions of quantities whose distribution is known.“ (Fisher 1922, S. 313) Diese Gliederung der Aufgaben einer theoretischen Statistik ist in der statistischen Literatur oft wiederholt worden, z.B. in der folgenden Formulierung von A. J. Dobson (1990, S. 49): „Statistical modelling involves three steps: (1) specifying models; (2) estimating parameters; (3) making inferences – that is, testing hypotheses, obtaining confidence intervals, and assessing the goodness of fit of models.“

gedessen kann in diesem Fall auch die Frage, ob sich zur Darstellung einer Häufigkeitsverteilung eine Verteilungsfunktion besser als eine andere eignet, nur pragmatisch unter Verweis auf den jeweils verfolgten Darstellungszweck beantwortet werden.

4. In der theoretischen Statistik ist jedoch eine andere Betrachtungsweise vorherrschend. Sie interpretiert die Mitglieder einer Familie parametrischer Verteilungsfunktionen als eine Menge möglicher statistischer Modelle zur Formulierung generativer Hypothesen. Jedes Mitglied der Verteilungsfamilie verweist dann auf einen hypothetischen Zufallsgenerator, der die Daten – bzw. die Sachverhalte, auf die sich die Daten beziehen – erzeugt haben könnte. Verteilungsfunktionen werden also als Verteilungsgesetze aufgefaßt, die dem Entstehungsprozeß der Daten zugrunde liegen. Dem entspricht dann die Aufgabe der statistischen Modellbildung: es gilt, einen Zufallsgenerator zu konstruieren, der die Daten erzeugt haben könnte. Daraus wird auch verständlich, was in der theoretischen Statistik unter einer Schätzung von Modellparametern verstanden wird. Man geht von der Annahme aus, daß die Daten bzw. die durch sie repräsentierten Sachverhalte tatsächlich durch einen zunächst unbekannten Zufallsgenerator entstanden sind; und die Aufgabe besteht darin, (parametrisch vorgebbare) Eigenschaften dieses Zufallsgenerators zu ermitteln (zu schätzen).[21]

[21]Die Rhetorik knüpft meistens daran an, daß man sich Verteilungsgesetze in Analogie zu Naturgesetzen denken möchte. Dem entspricht, daß die Parameter einer Verteilungsfunktion nicht als im Prinzip beliebig wählbare mathematische Größen, sondern als „Konstanten“ betrachtet werden, die in gewisser Weise Eigenschaften der natürlichen oder sozialen Prozesse ausdrücken, die die Sachverhalte, auf die sich die jeweiligen Daten beziehen, hervorgebracht haben. Z.B. spricht Fisher des öfteren von „Naturkonstanten“ (vgl. das in Abschnitt 5.2.2 angeführte Zitat sowie auch die Hinweise auf Pearson und Edgeworth in Abschnitt 2.3). Die Idee, daß es solche „Naturkonstanten“ gibt, ist eng verbunden mit der Vorstellung nomologischer Naturgesetze, wie sie im 19. Jahrhundert entwickelt worden ist. Zum Beispiel schrieb W. Stanley Jevons (1835–1882) in seinen „Principles of Science“ (1877):

“Our principal object must be to discover general quantitative laws of nature; but a very large amount of preliminary labour is employed in the accurate determination of the dimensions of existing objects, and the numerical relations between diverse forces and phenomena.” (Jevons 1877, S. 328)

Als Naturkonstanten werden die in der Formulierung solcher Naturgesetze verwendeten Zahlenwerte bezeichnet. Jevons hat dann versucht, diesen Sprachgebrauch etwas allgemeiner zu fassen, indem er folgende Arten von Konstanten unterschied: mathematical constants, physical constants, astronomical constants, terrestrial numbers, organic numbers, social numbers (Jevons 1877, S. 329). Damit wird zugleich nahegelegt, daß man auf allen diesen Gebieten in einer im Prinzip vergleichbaren Weise von quantitativ formulierbaren Gesetzen sprechen könne. Über „social numbers“ sagt Jevons:

“*Social numbers.* Little allusion need be made in this work to the fact that man in his economic, sanitary, intellectual, æstetic, or moral relations may become the subject of sciences, the highest and most useful of all sciences. Every one who is engaged in statistical enquiry must acknowledge the possibility of natural laws governing such statistical facts. Hence we must allot a distinct place to numerical information relating to the numbers, ages, physical and sanitary condition, mortality, &c., of different people, in short, to vital statistics. Economic statistics, comprehending the quantities of commodities produced, existing, exchanged and consumed, constitute another extensive body of science. In the progress of time exact investigation may possibly subdue regions of phenomena which at present defy all scientific treatment.” (Jevons 1877, S. 334)

Die Idee ist also, daß sich auch im Bereich sozialer Erscheinungen quantitative Gesetze finden lassen, die Konstanten enthalten, die mithilfe von Daten geschätzt werden können.

5. Man versteht dann auch, wie in diesem Kontext nicht nur von „Schätzen“, sondern auch von „Hypothesentests“ gesprochen wird. Die Hypothesen beziehen sich auf die hypothetischen Zufallsgeneratoren, die für den Entstehungsprozeß der Daten unterstellt werden. Um zu prüfen, ob für eine gegebene Menge von Daten ein Verteilungsgesetz angenommen werden kann, sind vor allem sog. Goodness-of-Fit-Tests entwickelt worden.[22] Die Rhetorik orientiert sich dabei oft an Fishers Idee hypothetischer unendlicher Grundgesamtheiten. Dann lautet die Frage, ob die Daten aus einer hypothetisch angenommenen Grundgesamtheit stammen könnten. Z.B. leiten D'Agostino und Stephens ihr Buch über solche Tests mit der Bemerkung ein:

“This book is devoted to the presentation and discussion of goodness-of-fit techniques. By these we mean methods of examining how well a sample of data agrees with a given distribution as its population.” (D'Agostino und Stephens 1986, S. 1)

Es ist indessen nicht erforderlich, die in früheren Abschnitten vorgetragene Kritik hier zu wiederholen.

6. Wichtiger ist die Frage, welchen potentiellen Sinn die Konstruktion und Schätzung eines statistischen Modells haben kann. Oder anders formuliert: Auf welche Aspekte unserer Erfahrungswelt beziehen sich die in der Form eines statistischen Modells geschätzen Verteilungsfunktionen bzw. ihre Parameter? Man kann grob drei Kontexte unterscheiden. Zunächst einen stichprobentheoretischen Kontext, von dem wir zwar eingangs die statistische Modellbildung abgegrenzt haben, der hier jedoch

[22]Eine ausführliche Darstellung findet man bei D'Agostino und Stephens, 1986.

der Vollständigkeit halber noch einmal zu erwähnen ist. In diesem Kontext beziehen sich die zu ermittelnden Verteilungen bzw. ihre Parameter auf eine real existierende statistische Gesamtheit. In diesem Fall kann man auch annehmen, daß den Verteilungsparametern in gewisser Weise real existierende Größen entsprechen. Wenn man z.B. das Durchschnittseinkommen in einer bestimmten realen Gesamtheit von Einkommensbeziehern mithilfe der Daten einer Stichprobe schätzen möchte, kann man annehmen, daß es das Durchschnittseinkommen, das man schätzen möchte, tatsächlich gibt, im Sinne einer Größe, die man unabhängig von dem verwendeten Schätzverfahren definieren kann.

7. Es ist allerdings klar, daß man im Kontext stichprobentheoretischer Inferenz nicht zu Verteilungsgesetzen gelangt. Ein unterschiedlicher Kontext entsteht erst durch eine theoretische Bezugnahme auf Prozesse, die die Sachverhalte, auf die sich Daten beziehen, hervorgebracht haben könnten, also bei der Formulierung generativer Hypothesen. Der Ansatz der theoretischen Statistik besteht dann darin, solche Prozesse als Realisationen von Zufallsgeneratoren zu deuten, woraus sich die korrespondierende Aufgabe ergibt, mit Hilfe von Daten, die der Zufallsgenerator tatsächlich hervorgebracht hat oder die er im Sinne einer hypothetischen Konstruktion hervorgebracht haben könnte, Einsichten in seine Beschaffenheit zu gewinnen. In diesem Fall läuft die Frage nach der Bedeutung der zu schätzenden (oder durch einen Hypothesentest zu prüfenden) Verteilungsgesetze darauf hinaus, ob der unterstellte Zufallsgenerator empirisch expliziert werden kann. Das ist möglich, wenn es sich um artifizielle Zufallsgeneratoren handelt. Dann kann man sich auf ein Verfahren beziehen, das auf demonstrierbare Weise verwendet werden kann, und man kann die Aufgabe der Modellbildung darin sehen, Eigenschaften dieses Verfahrens zu ermitteln und darstellbar zu machen. Ein analoges Verständnis erscheint möglich, wenn man sich auf empirisch identifizierbare Prozesse beziehen kann, die sinnvoll als Realisationen eines Zufallsgenerators verstanden werden können. Man kann in diesen Fällen von einem *verfahrenstechnischen Kontext* für Anwendungen der Statistik sprechen. Statistische Modelle dienen in diesem Kontext zur Ermittlung und Darstellung von Eigenschaften von Verfahren.

8. Eine solche verfahrenstechnische Deutung statistischer Modelle erscheint auch dann möglich, wenn sich das Erkenntnisinteresse, wie z.B. in der experimentellen Psychologie, darauf richtet, das Verhalten von Menschen berechenbar zu machen. Anders verhält es sich jedoch, wenn man sich in der theoretischen Sozialstatistik mit Zufallsgeneratoren auf soziale Prozesse beziehen möchte. Die besonderen Schwierigkeiten resultieren daraus, daß man mit sozialstatistischen Daten – in der Form, in der sie üblicherweise erhoben werden und zur Verfügung stehen – keinen empirisch explizierbaren Bezug auf soziale Prozesse vornehmen kann. Infolgedessen kann man in diesem Fall gar nicht davon sprechen, daß man sich mit einem Zufallsgenerator auf einen Prozeß bezieht. Sondern es gibt nur einen fiktiven Zufallsgenerator, an dessen Konzeption ein metaphorisches Reden von sozialen Prozessen angehängt wird. Dadurch wird schließlich auch das Reden vom Schätzen der Modellparameter fragwürdig. Denn das Wort 'schätzen' verweist im gewöhnlichen Sprachgebrauch – ähnlich wie das Wort 'messen' – darauf, daß es etwas gibt, das man schätzen kann. Aber die Größen, die durch sozialstatistische Modelle geschätzt werden, gibt es nicht in irgendeiner betimmbaren Weise als Tatsachen in unserer Erfahrungswelt; sie werden vielmehr durch die Modellkonstruktion und die dann angewendeten Schätzverfahren erzeugt.

9. Es ist deshalb verständlich, daß die Ermittlung sozialstatistischer Verteilungsgesetze ein von Anfang an umstrittenes Unternehmen gewesen ist. Insbesondere in diesem Kontext gab es auch immer ein unklares Neben- und Gegeneinander der beiden eingangs unterschiedenen Betrachtungsweisen: daß man Verteilungsfunktionen einerseits nur zur Darstellung von empirischen Häufigkeitsverteilungen, die aus jeweils gegebenen Daten ermittelbar sind, verwenden kann; und daß man sie andererseits auch als Verteilungsgesetze interpretieren kann, die – in einer allerdings nie aufgeklärten Weise – den empirischen Häufigkeitsverteilungen „zugrunde liegen“ und sie „beherrschen“. Verteter einer theoretischen Sozialstatistik haben sich natürlich immer wieder bemüht, die zweite Betrachtungsweise zu verteidigen.

10. Um die Diskussion zu illustrieren, sei exemplarisch auf eine Abhandlung von L. v. Bortkiewicz über „Realismus und Formalismus in der mathematischen Statistik“ (1915) hingewiesen. Die erste der eben unterschiedenen Betrachtungsweisen nennt Bortkiewicz „Formalismus“, die zweite „Realismus“. In seiner Abhandlung wendet sich Bortkiewicz am Beispiel einer Arbeit von L. Whitaker (1914), einer Autorin aus der Pearson-Schule, gegen den „Formalismus“:

> „In den Kreisen, die der mathematischen Statistik ferner stehen, begegnet man nicht selten der Auffassung, daß das ganze Streben der mathematischen Statistiker auf die Auffindung solcher 'deskriptiven' Formeln gerichtet sei [gemeint ist eine Verwendung von Verteilungsfunktionen bloß zur Darstellung einer gegebenen Häufigkeitsverteilung], wobei dann die üblichen Erwägungen über die Variabilität und Mutabilität der Objekte, mit denen sich die Statistik befaßt,

dazu herhalten müssen, um das Illusorische dieses Strebens darzutun. Demgegenüber dürfte es nicht unangebracht sein, die Worte eines der Meister der mathematischen Statistik in Erinnerung zu bringen, der für die in Frage stehende Art der Behandlung statistischer Daten nichts übrig hat als Hohn und Verachtung.

'Wenn ... Formeln aufgestellt werden', sagt *G. F. Knapp*, 'die ohne jede Beziehung auf die Natur des Vorgangs nichts weiter leisten als die beobachteten Zahlenwerte gleichsam in nuce zu umschließen, was bei Annahme einer willkürlichen Menge von sinnlosen Konstanten gar nicht einmal schwer ist, so hat das nichts mehr mit der angewandten Mathematik gemein als den Gebrauch von Buchstaben.'

Dieses spöttisch-wegwerfende Urteil ist allerdings etwas zu hart und muß, sofern das der Fall ist, auf die Rechnung jenes wissenschaftlichen Radikalismus gesetzt werden, von welchem *Knapp* so manche Probe gegeben hat. So hieße es z.B. sicherlich zu weit gehen, wollte man den analytischen Formeln, welche die Abhängigkeit der Sterbenswahrscheinlichkeit oder der Sterbensintensität von dem Lebensalter ausdrücken, jede Bedeutung absprechen. Darin hat aber *Knapp* unbedingt Recht, daß die Zurückführung eines statistischen Ergebnisses auf eine rein empirische Formel, vom Standpunkte der wahren Aufgaben der forschenden Statistik aus gesehen, etwas durchaus subalternes darstellt. Zugleich ist folgendes nicht außer Acht zu lassen: Deskriptive Formeln sind, wenn überhaupt, so doch nur insofern existenzberechtigt, als der Bereich ihrer näherungsweisen Gültigkeit in nicht allzu enge Grenzen eingeschlossen ist. Gilt aber eine derartige Formel nur in dem besonderen Fall, für welchen sie aufgestellt wird, so ist ihr theoretischer und praktischer Wert gleich Null. Gerade um solche Formeln von rein singulärer Gültigkeit handelt es sich aber bei *Lucy Whitaker*, und obschon ihre Konstanten nicht zahlreich sind – kommt sie doch jeweils mit zweien aus –, so sind sie darum nicht weniger 'sinnlos'.

Bezeichnet man als *'Formalismus'* diejenige Richtung in der mathematischen Statistik, welche in der Auffindung deskriptiver Formeln die Hauptaufgabe der Wissenschaft erblickt, so wird man von *Lucy Whitaker* wohl sagen dürfen, daß sie sich ganz und gar dem Formalismus verschrieben hat. Wer von der mathematischen Statistik verlangt und erwartet, daß sie das Verständnis für die eigenartigen Regelmäßigkeiten, die die sozialen Massenerscheinungen aufweisen, vertieft, die Methoden der Stoffverarbeitung verfeinert, die Aufdeckung der Wirkungsweise verschiedener Faktoren erleichtert, dem wird eine Arbeit wie die *Whitaker*sche nur Enttäuschung bringen, denn von alledem leistet sie nichts." (v. Bortkiewicz 1915, S. 244f)

Wenn ich es richtig verstehe, plädiert v. Bortkiewicz für eine theoretische Sozialstatistik, die nicht nur empirische Häufigkeitsverteilungen durch (ad hoc, wie auch immer konzipierte) Verteilungsfunktionen darstellt, sondern die sich um eine Ermittlung von Verteilungsgesetzen bemüht,

die den „sozialen Massenerscheinungen" zugrunde liegen, also um eine Konstruktion statistischer Modelle, so wie dies in Abschnitt 5.2.1 erläutert worden ist.

11. Dennoch bleibt die Frage, ob solche Modelle für soziale Sachverhalte und Ereignisse sinnvoll formuliert werden können. Tatsächlich bestand auch in v. Bortkiewicz' eigenen Untersuchungen das Argument schließlich nur darin, daß sich in einigen Fällen empirische Häufigkeitsverteilungen durch einfache Verteilungsfunktionen näherungsweise gut *darstellen* lassen. Exemplarisch sei noch einmal auf sein „Gesetz der kleinen Zahlen" (1898) hingewiesen, in dem er die Familie der Poisson-Verteilungen verwendet, um Häufigkeitsverteilungen darzustellen, bei denen ein Ereignis sehr selten eintrifft (vgl. Abschnitt 3.1.1). Sein Fazit lautet:

„Jene fast vollständige Übereinstimmung der Theorie mit der Erfahrung, welche sich hierbei herausstellt, gestattet kaum noch einen Zweifel über die objektive Bedeutung des Wahrscheinlichkeitsbegriffs für die untersuchten Gebiete des statistischen Geschehens. Wenn auch das Beobachtungsfeld, auf welches sich die Untersuchung im 2. Kapitel bezieht, ein örtlich und zeitlich beschränktes ist, so erscheint die erwähnte Schlußfolgerung auf den objektiven Charakter des Wahrscheinlichkeitsbegriffs an ähnliche Schranken nicht gebunden. Was in dieser Beziehung von Selbstmorden oder Unfällen in einzelnen Territorien bezw. für bestimmte Personenkreise und für bestimmte Zeiträume gilt, muß offenbar eine allgemeine Geltung haben. Wo und wann immer in einer menschlichen Gesellschaft Selbstmorde begangen werden und sich Unfälle ereignen, dürfte die Art des Zustandekommens dieser Geschehnisse eine solche sein, welche die Anwendung des Wahrscheinlichkeitsbegriffs zuläßt." (v. Bortkiewicz 1898, S. 38f)

Aber die von v. Bortkiewicz demonstrierte Tatsache, daß man für eine gewisse Anzahl empirischer Häufigkeitsverteilungen eine Familie mathematisch ähnlicher Verteilungsfunktionen konstruieren kann, liefert natürlich keine Begründung für seine Schlußfolgerungen.

12. Einfache Verteilungsgesetze, wie sie Ende des 19., Anfang des 20. Jahrhunderts in der theoretischen Sozialstatistik beliebt waren, sind in der weiteren Entwicklung weitgehend in Vergessenheit geraten. Es ist auch kaum noch strittig, daß eine Darstellung empirischer Häufigkeitsverteilungen durch theoretisch ausgedachte Verteilungsfunktionen nicht zu einer Begründung für generative Hypothesen führen kann.[23] Stattdes-

[23] Zum Beispiel hat R. von Mises (1957, S. 166f) die Verwendung theoretischer Verteilungsfunktionen zur Darstellung empirischer Häufigkeitsverteilungen pauschal der *deskriptiven* Statistik zugerechnet.

sen ist die Regressionsrechnung zum hauptsächlich verwendeten Werkzeug der Statistik ausgearbeitet worden. Damit beschäftigen wir uns im nächsten Abschnitt.

5.4 Konditionale Verteilungsmodelle

Die Regressionsrechnung wurde zuerst Ende des 19. Jahrhunderts von Francis Galton und Karl Pearson zur Untersuchung der Vererbung von Merkmalen entwickelt.[24] In der weiteren Entwicklung der Statistik sind in einer kaum noch überschaubaren Vielfalt unterschiedliche Varianten der Regressionsrechnung entwickelt und als theoretische Modelle gedeutet worden. Wir folgen hier nicht der historischen Entwicklung, sondern beginnen mit einigen Erläuterungen zum allgemeinen Begriff einer Regressionsfunktion. Im Anschluß daran wird besprochen, wie Regressionsfunktionen interpretiert werden können. Erst im nächsten Kapitel knüpfen wir wieder an die historische Entwicklung der theoretischen Sozialstatistik an und besprechen, wie sich Vorstellungen über „statistische Ursachenforschung“ entwickelt haben.

[24] Der motivierende Grundgedanke verläuft sehr ähnlich zu einer Überlegung von Lexis, die in Abschnitt 3.2 besprochen worden ist. Z.B. heißt es bei Pearson (1896, S. 255): „In the first place, we must definitely free our minds, in the present state of our knowledge of the mechanism of inheritance and reproduction, of any hope of reaching a mathematical relation expressing the degree of correlation between individual parent and individual offspring. The causes in any individual case of inheritance are far too complex to admit of exact treatment; [...]. We must proceed from inheritance in the mass to inheritance in narrower and narrower classes, rather than attempt to build up general rules on the observation of individual instances. Shortly, we must proceed by the method of statistics, rather than by the consideration of typical cases. It may seem discouraging to the medical practitioner, with the problem before him of inheritance in a particular family, to be told that nothing but averages, means, and probabilities with regard to large classes can as yet scientifically dealt with; but the very nature of the distribution of variation, whether healthy or morbid, seems to indicate that we are dealing with that sphere of indefinitely numerous small causes, which in so many other instances has shown itself only amenable to the calculus of chance, and not to any analysis of the individual instance. On the other hand, the mathematical theory will be of assistance to the medical man by answering, *inter alia*, in its discussion of regression the problem as to the average effect upon the offspring of given degrees of morbid variation in the parents. It may enable the physician, in many cases, to state a belief based on a high degree of probability, if it offers no ground for dogma in individual cases.“

5.4.1 Regressionsfunktionen

1. Bei der Regressionsrechnung wird vorausgesetzt, daß man sich auf mindestens zwei statistische Variablen beziehen kann. Um die folgenden Erläuterungen einfach formulieren zu können, beziehen wir uns auf eine zweidimensionale statistische Variable

$$(X,Y) : \Omega \longrightarrow \tilde{X} \times \tilde{Y}$$

Wenn man möchte, kann man sich natürlich vorstellen, daß sowohl X als auch Y mehrdimensionale Variablen sind. Ist nun dieser Ausgangspunkt gegeben, kann man zunächst – ganz analog zur Situation mit nur einer Variablen – eine zweidimensionale Häufigkeitsverteilung bilden, d.h. eine Funktion

$$h : \mathcal{P}(\tilde{X} \times \tilde{Y}) \longrightarrow \mathbf{R}$$

die jeder zweidimensionalen Merkmalsmenge $\tilde{x} \times \tilde{y}$, wobei $\tilde{x} \subseteq \tilde{X}$ und $\tilde{y} \subseteq \tilde{Y}$ ist, ihre relative Häufigkeit $h(\tilde{x} \times \tilde{y})$ zuordnet, d.h. den Anteil der Mitglieder von Ω, bei denen die Variable X einen Wert in $\tilde{x}$ und die Variable Y einen Wert in $\tilde{y}$ aufweist. Als eine äquivalente Notation wird auch

$$h(X \in \tilde{x}, Y \in \tilde{y})$$

verwendet. Gemeint sind jeweils relative Häufigkeiten, die durch eine Bezugnahme auf die Merkmalswerte der Elemente einer statistischen Gesamtheit Ω definiert sind.

2. Die Regressionsrechnung beginnt mit der Idee, daß man nicht nur die gemeinsame Verteilung von zwei Variablen, sondern auch bedingte Verteilungen betrachten kann. Am einfachsten verständlich ist das, wenn die Merkmalsräume der Variablen, also $\tilde{X}$ und $\tilde{Y}$, diskret sind, d.h. nur abzählbar viele Merkmalswerte umfassen. Es gibt dann für jeweils zwei Werte $x \in \tilde{X}$ und $y \in \tilde{Y}$ eine relative Häufigkeit $h(Y = y, X = x)$. Diese zweidimensionale Häufigkeit kann mithilfe des Begriffs einer bedingten Häufigkeitsverteilung z.B. folgendermaßen dargestellt werden:

$$h(Y = y, X = x) \;=\; h(Y = y \,|\, X = x)\, h(X = x)$$

Auf der linken Seite steht die 2-dimensionale Häufigkeitsverteilung von (X,Y). Auf der rechten Seite steht zunächst die bedingte Verteilung von Y für einen bestimmten Wert von X, dann folgt die Häufigkeit, mit der

X den Wert x annimmt. Dabei ist die Wahl des Wertes x beliebig; es muß nur vorausgesetzt werden, daß $h(X = x) > 0$ ist. Die gemeinsame Verteilung von (X, Y) kann also aus einer bedingten Verteilung und der zugehörigen Randverteilung rekonstruiert werden.

3. Die Regressionsrechnung konzentiert sich auf einen Teil der eben angegebenen Formel, nämlich auf die bedingte Verteilung; bei der Regression von Y auf X also auf den Ausdruck

$$h(Y = y \,|\, X = x)$$

Dieser Ausdruck kann unterschiedlich interpretiert werden.

a) Zunächst kann man sich vorstellen, daß x irgendein bestimmter Wert aus dem Merkmalsraum von X und y irgendein bestimmter Wert aus dem Merkmalsraum von Y ist. Dann ist $h(Y = y \,|\, X = x)$ eine bestimmte Zahl, nämlich die relative Häufigkeit dafür, daß Y den Wert y annimmt, unter der Bedingung, daß X den Wert x annimmt. Oder anders formuliert: Man betrachtet in diesem Fall die Teilgesamtheit derjenigen Elemente von Ω, bei denen X den Wert x annimmt; und dann berechnet man die relative Häufigkeit von „$Y = y$“ in dieser Teilgesamtheit.

b) Man kann sich auch vorstellen, daß zwar x irgendein bestimmter Wert aus dem Merkmalsraum von X ist, jedoch y kein bestimmter Wert, sondern eine Variable (genauer gesagt: eine logische im Unterschied zu einer statistischen Variablen). Man betrachtet dann nicht mehr eine bestimmte Zahl, sondern eine Funktion:

$$y \longrightarrow h(Y = y \,|\, X = x)$$

Sie ordnet, bei gegebenem Wert x, jedem $y \in \tilde{Y}$ die relative Häufigkeit $h(Y = y \,|\, X = x)$ zu. Wenn in der Statistik von einer bedingten Verteilung gesprochen wird, ist stets *diese* Interpretation gemeint. Das macht auch deutlich, daß Verteilungen keine Zahlen sind, sondern Funktionen, durch die Merkmalswerten, oder Mengen von Merkmalswerten, relative Häufigkeiten zugeordnet werden.

c) Schließlich gibt es noch eine dritte Interpretation, die für die Regressionsrechnung charakteristisch ist. Hierbei werden sowohl x als auch y als logische Variablen interpretiert. Es resultiert dann eine zweistufige Funktion:

$$\underbrace{x \longrightarrow \overbrace{(y \longrightarrow h(Y = y \,|\, X = x))}^{\text{bedingte Verteilung}}}_{\text{Regressionsfunktion}} \tag{5.4.1}$$

Bei der Regression von Y auf X wird also jedem Merkmalswert $x \in \tilde{X}$ eine Funktion zugeordnet, und zwar die bedingte Verteilung von Y bei dem jeweiligen x.

Somit ist also der Begriff einer allgemeinen Regressionsfunktion recht einfach verständlich. Alle weiteren Fragen betreffen nicht den Begriff, sondern beziehen sich darauf, wie man Regressionsfunktionen *darstellen* kann.

4. Das für die Praxis der Regressionsrechnung vornehmlich wichtige Problem besteht darin, daß eine Regressionsfunktion keine gewöhnliche Funktion ist. In der allgemeinen Form, in der wir sie eingeführt haben, ordnet eine Regressionsfunktion jedem $x \in \tilde{X}$ nicht eine bestimmte Zahl, sondern eine Funktion zu. Das ist praktisch nur schwer darstellbar. Bei der praktischen Regressionsrechnung charakterisiert man deshalb die bedingte Verteilung, also $h(Y = y \,|\, X = x)$, durch einfache Attribute und betrachtet Funktionen der Art

$$x \longrightarrow \text{Charakterisierung von}\,\{y \longrightarrow h(Y = y \,|\, X = x)\}$$

Je nachdem, wie die bedingte Verteilung $h(Y = y \,|\, X = x)$ charakterisiert wird, entsteht unterschiedliche Arten spezifischer Regressionsfunktionen.

5. Welche Eigenschaften der durch X bedingten Verteilung von Y für ihre Charakterisierung verwendet werden, ist im Prinzip gleichgültig; es ist jedoch sinnvoll, dabei auf den Merkmalsraum von Y Rücksicht zu nehmen und insbesondere zu unterscheiden, ob Y eine qualitative oder quantitative Variable ist. In praktischen Anwendungen der Regressionsrechnung werden hauptsächlich drei unterschiedliche Charakterisierungen verwendet.

a) Man verwendet oft Mittelwerte, wenn Y eine metrische Variable ist. Die Regressionsfunktion nimmt dann die Form

$$x \longrightarrow \mathrm{M}(Y \,|\, X = x)$$

an, wobei $\mathrm{M}(Y \mid X = x)$ der durch $X = x$ bedingte Mittelwert von Y ist. Man kann in diesem Fall von einer *Mittelwertsregression* sprechen, und sie bildet die in Anwendungen am häufigsten verwendete Variante der Regressionsrechnung.

b) Wenn Y einen quantitativen (ordinalen oder metrischen) Merkmalsraum hat, kann man ihre bedingte Verteilung durch Quantile charakterisieren und gelangt zu einer *Quantilsregression*:

$$x \longrightarrow Q_p(Y \mid X = x)$$

In diesem Fall ordnet die Regressionsfunktion jedem Wert $x \in \tilde{X}$ ein bestimmtes p-Quantil der durch $X = x$ bedingten Verteilung von Y zu. Ein mit der Mittelwertsregression vergleichbarer Spezialfall entsteht, wenn man bedingte Mediane verwendet.

c) Wenn schließlich Y einen qualitativen Merkmalsraum hat, etwa $\tilde{Y} = \{1, \ldots, m\}$, kann es auch sinnvoll sein, direkt die bedingten Häufigkeiten zu verwenden. Eine allgemeine Regressionsfunktion für die bedingte Verteilung von Y in Abhängigkeit von den Werten $x \in \tilde{X}$ besteht dann aus m Funktionen

$$x \longrightarrow h(Y = j \mid X = x) \quad \text{für } j = 1, \ldots, m$$

Die $j$te Funktion zeigt, wie die relative Häufigkeit dafür, daß Y den Wert j annimmt, von den Werten von X abhängt; zusammengenommen zeigen die Funktionen, wie die Verteilung von Y von den Werten von X abhängt. Man kann in diesem Fall von einer *Häufigkeitsregression* sprechen; als parametrische Varianten werden z.B. Logit- und Probitmodelle verwendet.

6. An die vorstehend skizzierten Begriffsbildungen schließt sich eine große Vielfalt praktischer Ansätze zur Konstruktion von Regressionsfunktionen an. Darauf einzugehen ist im Rahmen dieses Skripts jedoch weder sinnvoll noch erforderlich. Hier gilt unser Interesse vielmehr der Frage, wie Regressionsfunktionen in der theoretischen (Sozial-) Statistik konzipiert und interpretiert worden sind, d.h. wie sie mit der Idee verknüpft worden sind, daß die Sachverhalte, auf die sich statistische Daten beziehen, durch Zufallsgeneratoren erzeugt werden.

5.4.2 Deutung durch Zufallsgeneratoren

1. Wie im vorangegangenen Abschnitt gezeigt worden ist, können die Begriffsbildungen der Regressionsrechnung zunächst vollständig ohne eine Bezugnahme auf Zufallsgeneratoren eingeführt werden. Es ist ausreichend, daß man sich auf statistische Variablen beziehen kann, die für definite statistische Gesamtheiten definiert sind. Die jeweils verfügbaren Werte solcher Variablen bilden in jedem Fall den Ausgangspunkt für die Konstruktion von Regressionsfunktionen.

2. Theoretische Statistik beginnt dagegen mit der Annahme, daß die Sachverhalte, auf die sich Daten beziehen, entweder tatsächlich durch Zufallsgeneratoren zustande gekommen sind oder durch fiktiv unterstellbare Zufallsgeneratoren hervorgebracht worden sein könnten. Hat man es mit einer zweidimensionalen statistischen Variablen (X, Y) zu tun, kann man zunächst an einen Zufallsgenerator denken, der ihre Werte erzeugt, also jedesmal ein Wertepaar (x, y). Als Beispiel kann man daran denken, daß mit zwei Würfeln gleichzeitig gespielt wird. Also können sich auch alle Gedankengänge anschließen, die bisher über Verteilungsgesetze für einfache Variablen besprochen worden sind. Hat man es mit zwei- oder mehrdimensionalen Variablen zu tun, beziehen sich die Annahmen darauf, wie Zufallsgeneratoren beschaffen sind oder beschaffen sein könnten, von denen man weiß oder glauben kann, daß sie die jeweils gegebenen Daten erzeugt haben könnten.

3. Mit der Regressionsrechnung entsteht eine Variante dieser Vorgehensweise. Man kann sich den Unterschied anhand von zwei Bildern verdeutlichen. Bezieht man sich direkt auf eine zweidimensionale Variable (X, Y), besteht der Ausgangspunkt in der Annahme eines Zufallsgenerators der folgenden Art:

$$\boxed{\mathcal{G}} \longrightarrow (x, y)$$

Bei der Regressionsrechnung geht es jedoch darum, Einsichten in die durch die Werte einer Variablen (X) bedingte Verteilung einer anderen Variablen (Y) zu gewinnen. Das Bild sieht dann so aus:

$$x \longrightarrow \boxed{\mathcal{G}^*} \longrightarrow y$$

Der Zufallsgenerator $\mathcal{G}^*$ bekommt in diesem Fall einen Wert der Variablen X gewissermaßen als Input; und je nachdem erzeugt er auf eine dadurch bedingte Weise einen Wert für die Variable Y. Bei dieser Betrachtungsweise sollte allerdings betont werden, daß die Werte von X *als*

Bedingungen für die Funktionsweise von $\mathcal{G}^*$ verstanden werden sollen, nicht jedoch als Ereignisse, die den Zufallsgenerator aktivieren könnten. $\mathcal{G}^*$ ist wie jeder andere Zufallsgenerator nur ein Verfahren, das nicht von sich aus tätig werden kann. Man kann aber den Wert der Variablen X als eine Bedingung verstehen, von dem die Funktionsweise des Zufallsgenerators $\mathcal{G}^*$ abhängt. Faßt man die Werte von X als Parameter auf, kann man die Notation $\mathcal{G}^*(x)$ verwenden, um formal die Abhängigkeit der Funktionsweise des Zufallsgenerators von den Werten von X zum Ausdruck zu bringen.

4. Wichtig ist nun, daß man durch diese Konzeption die Fragestellung auf eine neue Weise formulieren kann. Geht man von dem ersten der beiden Bilder aus, kann man nur direkt und unmittelbar nach einem Verteilungsgesetz für den Zufallsgenerator $\mathcal{G}$ fragen. Verwendet man dagegen das zweite Bild, kann man die Frage in zwei Teilfragen zerlegen. Die erste Teilfrage betrifft die Beschaffenheit der Funktion

$$x \longrightarrow \mathcal{G}^*(x)$$

Dies ist das theoretische Pendent zu den Regressionsfunktionen, wie sie im vorangegangenen Abschnitt eingeführt worden sind. Dort wurde jedem Wert von X eine Häufigkeitsverteilung zugeordnet; hier wird jedem Wert von X ein Zufallsgenerator bzw. eine Wahrscheinlichkeitsverteilung zugeordnet. Die zweite Teilfrage richtet sich dann darauf, wie die Zufallsgeneratoren $\mathcal{G}^*(x)$ beschaffen sind.

5. Daß man die Frage nach der Funktionsweise von $\mathcal{G}$ in die beiden Teilfragen aufspalten kann, ist deshalb wichtig, weil sich dadurch neue rhetorische Möglichkeiten erschließen. Man kann sich nämlich $\mathcal{G}$ als eine Menge von Zufallsgeneratoren vorstellen, aus denen, je nachdem welcher Wert von X als Input auftritt, jeweils ein bestimmtes Element ausgewählt wird. Die Regressionsfunktion kann somit als eine Auswahlfunktion betrachtet werden, die je nachdem, welche Bedingungen vorliegen, einen bestimmten Zufallsgenerator auswählt, mit dem dann ein Wert der Variablen Y erzeugt wird. Die Regressionsfunktion wird dadurch zu einem separat konzipierbaren Aspekt des Verteilungsgesetzes, das den Zufallsgenerator $\mathcal{G}$ charakterisiert, und teilt mithin mit diesem den Charakter eines Gesetzes im Sinne einer Beschreibung eines Verfahrens.

6. Betrachten wir zum besseren Verständnis ein Beispiel, etwa einen Toaster. Die Variable Y bezieht sich auf die Dauer der Glühzeit, die Variable X erfaßt die Stellung des Hebels, mit dem die Glühdauer eingestellt werden kann. Je nachdem, wie der Hebel eingestellt ist, resultiert

eine unterschiedliche Glühdauer. Auch wenn der Hebel nicht verändert wird, werden aber bei einer Wiederholung die Dauern der Glühzeiten variieren. Ein Aspekt der Funktionsweise des Toasters kann also durch eine Regressionsfunktion

$$x \longrightarrow \mathcal{G}^*(x)$$

dargestellt werden. Je nach Stellung des Hebels gibt es einen bestimmten Wert der Variablen X; davon abhängig wird ein Zufallsgenerator $\mathcal{G}^*(x)$ ausgewählt, dessen Aktivierung die resultierende Glühdauer erzeugt. Dieses Beispiel illustriert auch, in welcher Weise die Werte der Variablen X als Bedingungen wirksam werden, nicht jedoch als Ursachen, die einen Prozeß ingangsetzen. Ein Toaster fängt nicht dadurch an zu glühen, daß man seine Glühdauer einstellt, er muß vielmehr ausdrücklich, üblicherweise mit einem separaten Schalter, aktiviert werden. Das Beispiel zeigt weiterhin, in welcher Weise man hier von einem Regressions*modell* sprechen kann. Das Modell liefert nicht eine Beschreibung im üblichen Sinne des Wortes, sondern soll dazu dienen, einen Aspekt der Funktionsweise des Toasters kenntlich zu machen. Zu diesem Zweck wird der Toaster als ein *Verfahren* zur Erzeugung von Glühdauern in Abhängigkeit von der Einstellung eines Hebels aufgefaßt. Das Modell setzt also bei einer spezifischen begrifflichen Repräsentation des empirischen Toasters an, auf den sich das Beispiel bezieht. Die meisten Eigenschaften des Toasters, als eines bestimmten Gegenstandes in unserer Erfahrungswelt, bleiben unberücksichtigt. Insbesondere bleibt unberücksichtigt, daß der Toaster, wie jeder andere Gegenstand, einen zeitlichen Entwicklungsprozeß durchmacht.[25]

7. Betont werden sollte noch einmal, daß die Formulierung eines Regressionsmodells spezifische Vorstellungen über den Ablauf des Prozesses impliziert, der die Werte der Variablen (X, Y) hervorbringt. Zwar kann der Übergang zu einer konditionalen Verteilung rein formal immer durchgeführt werden. Wenn man sich jedoch auf realisierbare Zufallsgeneratoren beziehen möchte, stellt sich die Frage, wie ein Verfahren bzw. ihm korrespondierende Prozesse vorstellbar gemacht werden können, die tatsächlich so funktionieren, daß Werte von X als Bedingungen für das Zustandekommen von Werten von Y aufgefaßt werden können. Bei dem

[25]Eine dementsprechend beliebte Betrachtungsweise der Modellkonstrukteure haben H. Freudenthal und H.-G. Steiner (1966, S. 191) so formuliert: „Man ist sich heutzutage darüber klar geworden, daß die Mathematik nicht auf die Wirklichkeit, sondern auf ein Modell der Wirklichkeit angewandt wird.“

eben verwendeten Beispiel eines Toasters ist dies ersichtlich der Fall. Ein weniger einfaches Beispiel liegt dann vor, wenn gleichzeitig mit zwei Würfeln geworfen wird. In diesem Fall kann der Prozeß, der die Werte von (X, Y) erzeugt, nicht so konzipiert werden, daß Werte von X als Bedingungen dafür auftreten, wie Y ihre Werte annimmt. Zwar kann man immer noch eine konditionale Verteilung bilden; und ein Wissen darüber mag auch nützlich sein, wenn die Aufgabe darin besteht, Werte von Y vorauszusagen, wenn zunächst nur der für X realisierte Wert bekannt ist. Dann dient jedoch das Regressionsmodell als ein Hilfsmittel für ein Prognoseproblem, nicht zur Charakterisierung eines Verfahrens bzw. der ihm korrespondiernden Prozesse.

5.4.3 Generative Regressionsmodelle

1. Regressionsmodelle werden in vielen unterschiedlichen Kontexten verwendet. Eine erste Unterscheidung ergibt sich daraus, ob man sich mit dem Modell auf eine statistische Gesamtheit oder auf einen Prozeß bezieht. Im ersten Fall wird das Regressionsmodell für eine statistische Gesamtheit Ω formuliert und man kommt vollständig mit den in Abschnitt 5.4.1 erläuterten Begriffsbildungen aus; insbesondere ist es dann weder erforderlich noch sinnvoll, einen Zufallsgenerator zu unterstellen, der die Sachverhalte, auf die sich die Daten beziehen, hervorgebracht haben könnte. Ein statistisches Inferenzproblem entsteht dann nur daraus, daß die verfügbaren Daten unvollständig und/oder ungenau sind. Eine Standardsituation besteht darin, daß sich die Daten auf eine Zufallsstichprobe aus Ω beziehen und daß dann diese Daten verwendet werden sollen, um eine Regressionsfunktion für die Gesamtheit Ω zu schätzen. Das ist ein Spezialfall stichprobentheoretischer Inferenz. Der Charakter der Regressionsfunktion ändert sich dadurch nicht; sie liefert eine Charakterisierung der Häufigkeitsverteilung der für die Gesamtheit Ω definierten statistischen Variablen. Die Regressionsfunktion liefert in diesem Fall eine statistische Beschreibung eines voraussetzbaren Sachverhalts.

2. Ein vollständig anderer Kontext entsteht, wenn man sich mit der Formulierung eines Regressionsmodells auf einen Prozeß bezieht, der die Sachverhalte, auf die sich die Daten beziehen, hervorbringt oder hervorbringen könnte. Mit einer solchen Betrachtungsweise beginnt die Konzeption von Regressionsmodellen in der theoretischen Statistik. Zur Unterscheidung kann man in diesem Fall von *generativen Regressionsmodellen* sprechen. Hat man als Ausgangspunkt eine statistische Variable

(X, Y), besteht die Idee darin, daß man Werte der Variablen X als Bedingungen für Prozesse betrachten kann, die die Sachverhalte hervorbringen, auf die sich die Variable Y bezieht. Für die Formulierung generativer Regressionsmodelle sind also Annahmen über diese Prozesse erforderlich. Die Basisannahme besteht darin, daß man diese Prozesse als Realisationen von Zufallsgeneratoren konzipieren kann, deren Funktionsweise von den als Bedingungen vorausgesetzten Werten von X abhängt. Diese Basisannahme charakterisiert generative Regressionsmodelle als statistische Modelle, so wie dies Wort in Abschnitt 5.2.1 eingeführt worden ist. Wiederum gibt es unterschiedliche Anwendungskontexte. Einer dieser Kontexte wird durch das oben verwendete Beispiel der Glühzeiten eines Toasters illustriert. Man kann sich viele ähnliche Beispiele ausdenken. Als Fragestellung liegt in jedem Fall zugrunde, wie ein empirisch explizierbarer Prozeß von gewissen identifizierbaren Bedingungen abhängt. Unterschiedliche Kontexte entstehen daraus, wie der Prozeß beschaffen ist und ob man sich überhaupt auf einen empirisch identifizierbaren Prozeß beziehen kann.[26]

3. Um den ersten dieser beiden Gesichtspunkte zu verdeutlichen, kann man an die Verwendung von Regressionsmodellen in der experimentellen Psychologie denken. Es ist üblich, in diesem Fall von der Vorstellung eines Reiz-Reaktion-Schemas auszugehen, etwa in folgender

[26] In der statistischen Literatur wird meistens nur pauschal von „Experimenten“ gesprochen, und dieses Wort wird meistens so unspezifisch verwendet, daß es auf beliebige Prozesse anwendbar erscheint, notfalls dadurch, daß man nur eine „gedankliche Wiederholbarkeit“ unterstellt. Zum Beispiel schreiben Cox und Hinkley (1974, S. 5): „For the more elaborate parts of statistical theory we start from the representation of observations in terms of random variables and the idea that normally the parameters of the underlying probability distributions are the quantities of real interest. Yet this view needs to be treated with proper caution and should not be taken for granted. Where the data are obtained by the random sampling of a physically existing population, the parameters have a reasonably clear meaning as properties of the population. In other cases the probability distributions refer to what would happen if the experiment were repeated a large number of times under identical conditions; this is always to some extent hypothetical and with ‘unique’ data, such as economic time series, repetition under identical conditions is entirely hypothetical. [Man erinnere sich an die Ausführungen von Hagood, die in Abschnitt 4.3.4 besprochen worden sind.] Nevertheless the introduction of probability distributions does seem a fruitful way of trying to separate the meaningful from the accidential features of data. Parameters are to be regarded sometimes as representing underlying properties of a random system and sometimes as giving concise descriptions of observations that would be obtained by repetition under the same conditions.“ Dagegen möchte ich betonen, daß der potentielle Sinn eines statistischen Modells davon abhängt, ob der Prozeß, auf den es sich der Intention nach beziehen soll, empirisch expliziert werden kann.

Form:

$$x \longrightarrow \boxed{A} \longrightarrow y$$

Eine Person A wird mit einem Reiz x konfrontiert, dann ermittelt man die Reaktion y. Offenbar gibt es eine Analogie zu dem oben zur Illustration von Regressionsmodellen verwendeten Bild. Auf dieser Analogie beruht auch die Formulierung theoretischer Regressionsmodelle in der experimentellen Psychologie. Die Person A wird als ein Zufallsgenerator aufgefaßt, der in Abhängigkeit von den durch x vorausgesetzten Bedingungen zufällig gewisse Reaktionen erzeugt.[27] Allerdings ist auch klar, wo die Analogie an Sinngrenzen stößt. Menschen funktionieren nicht auf die gleiche Weise wie Maschinen, z.B. Toaster. Menschen sind Akteure, die sich auf die Bedingungen ihres Verhaltens reflexiv beziehen können und dadurch auch lernfähig sind. In der Formulierung von Regressionsmodellen wird von diesen Besonderheiten abstrahiert; Menschen werden mit Maschinen gleichgesetzt. Das mag zweckmäßig sein, wenn man daran interessiert ist, Verhaltensweisen von Menschen berechenbar zu machen.[28] Aber es ist auch klar, daß man mit einem solchen Modellansatz keine Einsichten in solche Aspekte menschlichen Verhaltens gewinnen kann, durch die sie sich von Maschinen unterscheiden.

4. Schließlich kann man an Regressionsmodelle in der theoretischen Sozialstatistik denken. Die hauptsächliche Schwierigkeit resultiert in diesem Fall daraus, daß es keinen empirisch explizierbaren Prozeß gibt, auf den sich die Modellbildung beziehen könnte. Denken wir als Beispiel an Einkommensfunktionen. Die Variable X erfaßt gewisse Merkmale der Einkommensbezieher, z.B. ihr Geschlecht und ihr Alter, die Variable Y

[27]In dieser Formulierung bezieht sich das Regressionsmodell auf eine bestimmte Person A. Bei der Formulierung psychologischer Regressionsmodelle wird im allgemeinen davon ausgegangen, daß es sich nicht auf irgendeine bestimmte Person, sondern auf eine indefinite Gesamtheit von Personen bezieht, also auf einen Kollektivgegenstand im Sinne Fechners.

[28]Dies wird am deutlichsten bei denjenigen Ansätzen, die statistische Modelle unmittelbar mit praktischen Entscheidungen verknüpfen. Exemplarisch sei auf J. Neyman (1971, S. 12f) hingewiesen: „In many instances human behavior has to be determined on the basis of experiments or of observations, the results of which are subject to variation that is impossible to ignore. An effort is made to substitute for real experiments and real observations certain hypothetical stochastic mechanisms with probabilities idealizing the relative frequencies. Naturally, some of these stochastic models are grossly inadequate, but some others do reproduce the observable frequencies to a satisfactory degree. The available models are then used to deduce rules of behavior (or statistical decision functions) that, in a sense, minimize the frequency of wrong decisions.“

erfaßt die Höhe ihres Einkommens. Dann kann man für den Zusammenhang der Variablen eine Regressionsfunktion konstruieren.[29] Führt man diese Konstruktion mit einer Menge von Daten, die man für eine bestimmte Menge von Einkommensbeziehern gewonnen hat, durch, liefert die Regressionsfunktion eine statistische Beschreibung des Zusammenhangs zwischen Geschlecht, Alter und Einkommen in dieser statistischen Gesamtheit. Ein generatives Regressionsmodell entsteht jedoch erst dadurch, daß man sich auf einen Prozeß bezieht, der in Abhängigkeit vom Geschlecht und vom Alter die Höhe der Einkommen hervorbringt. Es ist zwar leicht, sich anstelle eines solchen Prozesses einen Zufallsgenerator vorzustellen. Das Problem besteht jedoch darin, daß es sich um eine vollständig fiktive Vorstellung handelt. Und zwar fiktiv nicht deshalb, weil man die Vorstellung eines Zufallsgenerators verwendet, sondern weil es in diesem Fall keinen empirisch explizierbaren Prozeß gibt, für dessen begriffliche Repräsentation die Vorstellung eines Zufallsgenerators verwendet werden könnte. Es gibt keinen Prozeß, der in Abhängigkeit vom Geschlecht und vom Alter Einkommen hervorbringt.

5.4.4 Modelle für Situationen

1. Allerdings kann man versuchen, sich auf die einzelnen Prozesse zu beziehen, durch die es in der sozialen Realität dazu kommt, daß jeweils bestimmte Menschen jeweils bestimmte Einkommen erzielen. Das kann man empirisch untersuchen. Man kann an beliebig vielen Beispielen untersuchen, wie Arbeitsverträge zustande kommen, durch die u.a. die Höhe der Arbeitseinkommen fixiert wird. Und man kann auch untersuchen, ob und ggf. wie dabei auf das Geschlecht und Alter der jeweiligen Personen bezug genommen wird. Solche Untersuchungen müssen sich zunächst auf jeweils bestimmte einzelne Situationen beziehen. Aber unter Umständen kann es gelingen, die Idee eines abstrakten Situationstyps zu konzipieren, so daß die vielen einzelnen Situationen, in denen Arbeitsverträge abgeschlossen werden, als Beispiele dieses Situationstyps angesehen werden können. In unserem Beispiel könnte man vielleicht einen Situationstyp S folgendermaßen konzipieren:

$$S = (\omega, X_1, X_2, \ldots, Y)$$

Der Situationstyp bezieht sich auf Situationen, in denen Arbeitsverträge abgeschlossen werden. Jedes Beispiel für einen solchen Situationstyp

[29]Bei der Konstruktion von Einkommensfunktionen werden meistens Verfahren der linearen Mittelwertsregression verwendet.

kann dann unter anderem durch eine Person ω, auf die sich der Vertrag bezieht, und durch eine Reihe von Situationsvariablen charakterisiert werden. In unserem Beispiel könnten X_1 und X_2 das Geschlecht und das Alter der Person erfassen, und Y die Höhe des für sie durch den Arbeitsvertrag fixierten Einkommens. Man kann sich weiterhin vorstellen, daß in jeder Situation, die als ein Beispiel für den Situationstyp S gelten kann, ein Prozeß stattfindet, durch den darüber entschieden wird, welcher Wert für die Situationsvariable Y fixiert wird, und bei dessen Ablauf u.U. darauf bezug genommen wird, welche Werte die Variablen $X_1, X_2, \ldots$ in der gegebenen Situation aufweisen. Schließlich kann man ein Regressionsmodell formulieren, das durch eine Bezugnahme auf Situationen des Typs S explizierbar wird.

2. Es ist auch klar, wie man ein solches Regressionsmodell berechenbar machen kann. Man beobachtet eine Reihe von Situationen, die als Beispiele für den Situationstyp S gelten können, etwa

$$s_i = (\omega_i, x_{i1}, x_{i2}, \ldots, y_i)$$

für $i = 1, \ldots, n$. Man hat dann Daten für n beobachtete Situationen des Typs S und kann sie verwenden, um ein für den Situationstyp formuliertes Regressionsmodell zu berechnen. Schließlich kann man als einen diskutierbaren Anspruch formulieren, daß das Regressionsmodell etwas darüber aussagt, wie die Prozesse, die in Situationen des Typs S ablaufen, funktionieren.

3. Der Versuch, den Sinn von Regressionsmodellen in der Sozialstatistik so zu verstehen, daß sie sich auf Situationstypen beziehen – bzw. auf Prozesse, die man mit Beispielen für Situationstypen assoziieren kann –, erscheint durchaus sinnvoll. Dieser Ansatz macht nicht nur deutlich, daß der potentielle Sinn eines Regressionsmodells davon abhängt, daß man sich in seiner Formulierung auf empirisch explizierbare Prozesse beziehen kann;[30] er zeigt auch, wie man das vielleicht erreichen kann. Es bleibt jedoch die Frage, ob es irgendeinen Erkenntnisgewinn vermitteln könnte, die Prozesse, auf die man sich dann beziehen kann, als Zufallsgeneratoren zu konzipieren.

4. Folgt man der Idee, sozialstatistische Regressionsmodelle auf Situationstypen zu beziehen, besteht zunächst keinerlei Grund dafür, die Prozesse, die in Situationen stattfinden, die als Beispiele des theoretisch konzipierten Situationstyps gelten können, als Zufallsgeneratoren

[30] Etwa so wie dies in dem in Abschnitt 5.2.2 angeführten Zitat von Coleman angedeutet worden ist.

aufzufassen. Jede empirische Bezugnahme auf bestimmte Beispiele des Situationstyps zeigt auch, daß die in ihnen stattfindenden Prozesse keinerlei Ähnlichkeit mit der Funktionsweise von Zufallsgeneratoren aufweisen. Vielmehr handelt es sich um Interaktionsprozesse der jeweils beteiligten Akteure. Daraus resultiert jedoch ein Widerspruch zu den gedanklichen Voraussetzungen eines Zufallsgenerators. Der Begriff eines Zufallsgenerators impliziert, daß es sich um ein Verfahren handelt, das unter einer Reihe angebbarer Bedingungen in der gleichen Weise wiederholt angewendet werden kann. Wenn man sich bei der Formulierung von Regressionsmodellen auf Situationstypen bezieht, kann man zwar *Bedingungen der Situation* durch Situationsvariablen fixieren, nicht jedoch das Verhalten und die Interaktionsprozesse der Akteure, die in den jeweils konkreten Situationen tätig werden. Wenn man gleichwohl Zufallsgeneratoren verwendet, um sich ein theoretisches Bild für den Ablauf von Prozessen in sozialen Situationen zu machen, wird nicht nur davon abstrahiert, daß es sich um Prozesse handelt, die durch Akteure zustande kommen; es wird vielmehr eine theoretische Konzeption verwendet, die dem Verständnis sozialer Interaktionsprozesse widerspricht. Das mag nicht wichtig sein, wenn man nur daran interessiert ist, den Ablauf sozialer Situationen voraussagbar zu machen. Wenn man jedoch ein soziologisches Erkenntnisinteresse verfolgt, d.h. Einsichten in die Beschaffenheit und Funktionsweise sozialer Prozesse gewinnen möchte, ist die Vorstellung, daß sie als Zufallsgeneratoren konzipiert werden können, nicht nur nicht hilfreich, sondern irreführend.

5. Die Idee, sozialstatistische Regressionsmodelle als Modelle für Situationstypen zu konzipieren, soll hier nicht näher verfolgt werden, denn sie ist in der bisherigen Entwicklung der theoretischen Sozialstatistik nicht ausgearbeitet worden. Sie sollte hier nur angedeutet werden, um zu zeigen, daß man sich Varianten generativer Regressionsmodelle vorstellen kann, die ohne eine Unterstellung fiktiver Zufallsgeneratoren auskommen. Wollte man die Idee näher verfolgen, käme es vor allem darauf an, den Begriff eines Situationstyps zu explizieren, auf den sich die Modellbildung beziehen soll. Er kann ja nicht als Menge oder Klasse aller Situationen konzipiert werden, die als Beispiele des Situationstyps vorstellbar sind; das wäre nicht weniger obskur als eine hypothetische Population. Allerdings wäre es wichtig, daß man zeigen kann, daß und wie ein Situationstyp durch Beispiele exemplifiziert werden kann und in welcher Weise Überlegungen, die anhand des Modells gebildet werden können, zur Reflexion von Situationen genutzt werden können, die als Beispiele des Situationstyps vorstellbar sind.

Kapitel 6

Statistische Ursachenforschung

In den vorangegangenen Kapiteln wurde bereits gelegentlich angedeutet, daß sich Fragestellungen der theoretischen Sozialstatistik meistens auch direkt oder indirekt auf „Verursachungszusammenhänge“ beziehen. In diesem Kapitel sollen einige der Vorstellungen besprochen werden, die in der Entwicklung der Sozialstatistik zum Reden von Ursachen gebildet worden sind.

6.1 Ursachen und Bedingungen

G. U. Yule und M. G. Kendall schreiben in ihrem Lehrbuch:

> “When the statistician has arranged and compressed his data into a suitable form, or decided on the functions and evaluated the quantities which he has chosen to describe them, the first stage of his inquiry is finished. It may be that he would wish to take it no further; for instance, if he is preparing an index number for the economist he may wish to hand over the number to that person without comment, for him to make such use of it as he thinks fit. More frequently, however, he has prepared the data for his own use as a statistician. He then proceeds to the next stage, that of analysis and elucidation of the causal system which gave rise to them.” (Yule und Kendall 1950, S. xxii)

Ähnlich wurden in vielen Lehrbüchern Aufgaben statistischer Ursachenforschung betont; hier noch eine Formulierung aus der 5. Auflage der „Elements of Statistics“ von A. L. Bowley (1926, S. 17, 1. Aufl. 1901):

> “A most important function of statistics is to produce evidence showing the relation of one group of phenomena to another; for the information obtained is presumably intended as a guide for action, the guidance is generally needed to show what actions are likely to produce certain desired effects, and this is best investigated by finding how such effects have been produced in the past. We have then to determine whether changes in one measurable quantity have produced changes in another; a problem very often insoluble, but one on which most light can be obtained by the study of the relevant statistics in the light of mathematics, the mathematics of probability, and it is in this particular branch of mathematics that recent statistical progress has been chiefly made.”

Hier zeigt sich auch deutlich, daß die Überlegungen auf *generative* Hypothesen zielen, so wie dies Wort in Abschnitt 5.1.1 eingeführt worden ist. Aber können und ggf. in welcher Weise können sozialstatistische Untersuchungen zu Einsichten in Ursachen für soziale Sachverhalte und Ereignisse beitragen? Daß eine Antwort nicht offen zutage liegt, hat man schon in Abschnitt 3.2.2 sehen können, wo wir uns mit Überlegungen von Lexis beschäftigt haben.

6.1.1 Entstehung einer fragwürdigen Rhetorik

1. Der Ausdruck „statistische Ursachenforschung" taucht seit etwa Anfang des 20. Jahrhunderts in den Lehrbüchern zur Statistik auf.[1] Damit verbunden ist die Entwicklung einer neuen Rhetorik. Um sie zu verstehen, kann zunächst vom alltagspraktischen Reden von Ursachen und Bedingungen ausgegangen werden. In diesem Kontext gibt es eine deutliche Unterscheidung. Als Ursachen werden „Kräfte" oder „Akteure" angesprochen, die etwas bewirken können; dagegen wird mit Bedingungen auf Eigenschaften einer Situation verwiesen, auf einen Kontext, in dem sich Ursachen als „Kräfte" oder „Akteure" betätigen. Demgegenüber besteht die neue Rhetorik, die sich in der Sozialstatistik herausgebildelt hat, zunächst darin, daß die Bedingungen eines Geschehens als seine Ursachen bezeichnet werden.

2. Karl Pearson hat das in seiner „Grammar of Science" (1900) damit zu begründen versucht, daß im umgangssprachlichen Reden von „Kräften" und „Akteuren" ein animistisches Weltbild zum Ausdruck kommt, von dem sich eine wissenschaftliche Natur- und Gesellschaftsbetrachtung lossagen müsse.[2] Tatsächlich ist jedoch durch die neue Rhetorik, die Bedingungen als Ursachen erscheinen läßt, das animistische Weltbild nicht überwunden worden, sondern hat nur eine neue und wesentlich obskurere Form angenommen. Als Beispiel sei auf folgende Ausführung im „Grundriß der Statistik" von F. Zizek (1921) hingewiesen:

[1] Einen guten Einblick in die Diskussion, wie sie in den 1920er Jahren in Deutschland geführt wurde, liefert eine Abhandlung von F. Zizek über „Ursachenbegriffe und Ursachenforschung in der Statistik" (1928a). Eine instruktive Darstellung, wie statistische Ursachenforschung in der amerikanischen Soziologie rezipiert wurde, findet sich bei Bernert, 1983.

[2] Vgl. Pearson 1900, Kap. IV. Ich erwähne hier Pearson, weil er für die Entwicklung der Statistik sehr einflußreich geworden ist. Die Idee, „Kräfte" und „Akteure" aus wissenschaftlichen Kausalitätsvorstellungen zu eliminieren, ist indessen ein Grundgedanke des gesamten „mechanistischen" Weltbildes, wie es in der Wissenschaft seit Descartes verfolgt wurde. Eine sehr explizite Ausarbeitung hat der Gedanke vor allem in der „Logik" von John Stuart Mill (1877, Kap. V) bekommen.

> „Die Statistik unternimmt es, den Faktoren, welche auf die verschiedenen Erscheinungen einwirken, nachzuforschen, ihre Wirksamkeit festzustellen und die Stärke ihres Einflusses zu messen. Es handelt sich da um Faktoren, welche auf Massen von Einzelfällen einwirken und auf die Gestaltung von Massenerscheinungen Einfluß nehmen, nicht etwa um die manchmal bei Erhebungen erfragten konkreten Ursachen der einzelnen Ereignisse (Todesursache, Ursache des Betriebsunfalls, Brandursache usw.)." (Zizek 1921, S. 164)

Zur Erläuterung werden folgende Beispiele angegeben:

> „Ein Faktor, der nach den verschiedensten Richtungen Wirkungen ausstrahlt, ist zunächst das Geschlecht; Frauen und Männer haben verschiedene Sterbenshäufigkeit, Lebensdauer, Verbrechens-, Selbstmordhäufigkeit, eine verschiedene Quote der Erwerbstätigen usw. Ebenso mannigfach sind die Wirkungen des Alters: mit dem Alter schwankt die Geschlechtsgliederung der Lebenden, die Sterbenshäufigkeit und fernere Lebensdauer, die Erkrankungshäufigkeit, die Quote der Erwerbstätigen, die Lohnhöhe, die Selbstmordziffer und noch manches andere. Auch der Familienstand übt gewisse Einflüsse aus, z.B. auf die Sterbenswahrscheinlichkeit. Große Tragweite besitzt ferner der Unterschied von Stadt und Land; die städtische und die ländliche Bevölkerung unterscheiden sich hinsichtlich des Altersaufbaues, der Geburten- und Sterbenshäufigkeit usw.; das Leben in der Stadt löst eben nach verschiedenen Richtungen andere Wirkungen aus als das Leben auf dem platten Lande." (Zizek 1921, S. 164)

Man erkennt, wie hier nicht nur jede Unterscheidung zwischen Akteuren und Bedingungen ihrer Tätigkeiten preisgegeben wird, sondern daß eine Rhetorik verwendet wird, die äußere Bedingungen eines Geschehens als die eigentlich wirksamen Kräfte (Faktoren) erscheinen lassen soll.[3]

3. Hier noch ein weiteres Beispiel aus einem Lehrbuch von H. Westergaard und H. C. Nybølle (1928):

> „Im Grade der Klarstellung der Kausalitätsverhältnisse erweitert sich auch die menschliche Erkenntnis, und in entsprechendem Umfange dürfen wir Schlüsse

[3] Hier noch eine Illustration aus einem neueren Lehrbuch der Statistik von R. M. Sirkin (1995, S. 21): „When we do research, we are ultimately seeking to infer the *cause* of the phenomenon we study. What variables when changing in value cause the variable we are studying to change? What factors account for the amount of rainfall? If levels of cloudiness are associated with levels of rainfall, and we assume for a moment that the actions of other variables such as temperature and humidity play no role in the direct relationship between cloudiness and rainfall, we are tempted to suggest a *causal relationship* between the two. We might assume that changes in levels of cloudiness bring about changes in the amount of rainfall. In effect, we are saying that it is the clouds that cause the rain." Verallgemeinert man dieses Beispiel, erscheint es so, als ob wir in einer Welt leben, deren eigentliche Akteure Variablen sind.

auf die nächste Zukunft wagen. [...] Die Aufgabe, die somit in erster Linie der Statistik gestellt wird, erweist sich nun allerdings oft als unlösbar. Wenn auch theoretisch genommen die Möglichkeit vorhanden ist, die wirkenden Ursachen ausfindig zu machen, so ist dieses doch bei der tatsächlichen Sachlage häufig äußerst schwierig, weil nicht eine einzelne, sondern viele, oft eine ungemein große Zahl von wirkenden Ursachen Berücksichtigung finden muß. Man erhöht z.B. die Getränkesteuer und beobachtet ein Abnehmen der Trunksucht; jene Maßregel ist aber wahrscheinlich nicht die einzige Kraft, welches dieses Resultat geschaffen hat; man hat es vielleicht gleichzeitig mit der Wirkung eines strengeren Strafgesetzes, mit einer Reduktion der Anzahl von Schankstellen oder mit einer Enthaltsamkeitsbewegung zu tun, mit Faktoren, die sämtlich aus einer allmählich entstandenen allgemeinen Mißstimmung gegen die alkoholischen Getränke hervorgegangen sein können.
Nun kann man meistens nicht eine einzelne Ursache aussondern, um die Wirkung derselben zu erforschen. Je gründlicher man die Sache untersucht, desto mehr Ursachen kommen zum Vorschein, bis sich unter dem Mikroskop des Forschers eine fast unübersehbare Zahl gegen- und miteinander wirkender Kräfte enthüllt.“ (Westergaard und Nybølle 1928, S. 3f)

Hier wird auch eine weitere Vorstellung sichtbar, die die Entwicklung der Sozialstatistik bis heute durchzieht: daß sie es mit Sachverhalten und Ereignissen zu tun habe, die durch eine sehr große Anzahl einzelner Ursachen hervorgerufen werden.[4] Wie sich diese Vorstellung insbesondere unter Soziologen verbreitet hat, zeigt folgende Bemerkung von F. A.

[4]Ganz explizit hat insbesondere Tschuprow (1905, 1906) versucht, theoretische Statistik auf die Idee einer „Vielzahl von Ursachen“ zu beziehen. Yule und Kendall (1950, S. xivff) haben sogar vorgeschlagen, diese Vorstellung definitorisch zu fixieren: „By *Statistics* we mean quantitative data affected to a marked extent by a multiplicity of causes. By *Statistical Methods* we mean methods specially adapted to the elucidation of quantitative data affected by a multiplicity of causes.“ Zur Erläuterung schreiben sie:

“One of the principal aims of Science is to trace, amidst the tangled complex of the external world, the operation of what are called ‘laws’ – to interpret a multiplicity of natural phenomena in terms of a few fundamental principles. A knowledge of the operation of these laws enables us to talk of ‘cause’ and ‘effect’. The metaphysical problems associated with these words need not detain us, but since in the sequel we shall often use them, it is proper to explain that we adopt them as a convenient way of expressing serviceable and familiar ideas. We shall be dealing with the everyday world, where ‘law’ and ‘cause’ have significant and important connotations.
With this convention, we may say that any physical event, and in particular that described by quantitative data, is produced by the operation of one or more causes. The number of causes which produce any particular effect may be, and usually is, extremely large. For instance, the height of a man is causally linked with his race, his ancestry, his habitation, his diet during youth, his age, his occupation, and at any given moment even with his position and the time of day.

Ross (1934, S. 464):

“It is platitudinous to say that social phenomena are results of complexities of causes. Almost every beginning student in the field seizes upon this as one of the few certainties in sociology.”

Man kann vermuten, daß diese Vorstellung mit der neuen Rhetorik zusammenhängt. Denn sicherlich kann man sich bei jedem realen Geschehen beliebig viele Bedingungen vorstellbar machen, durch die der jeweilige Kontext charakterisiert werden kann. Werden dann die Bedingungen eines Geschehens als die eigentlich wirksamen Ursachen, Kräfte, Faktoren dargestellt, entsteht die Vorstellung einer Vielzahl von Ursachen, die die einzelnen Geschehnisse hervorbringen.[5] — Und schließlich erscheint

Experiment, the great weapon of scientific inquiry, derives its power from the ability of the experimenter to replace such complex systems of causation by simple systems in which only one causal circumstance is allowed to vary at a time. This is perhaps an ideal, but it is one which is closely approached with the technique of modern laboratory practice.
Let us, however, turn for a moment to social science, as the parent of the methods termed ‘statistical’, and consider its characteristics as compared, say, with physics or chemistry. One characteristic stands out so markedly that attention has been repeatedly directed to it by ‘statistical’ writers as the source of the peculiar difficulties of their science – the observer of social facts cannot experiment, but must deal with circumstances as they occur, apart from his control. The simplification open to the experimenter being impossible, the observer has, in general, to deal with highly complicated cases of multiple causation – cases in which a given result may be due to any one of a number of alternative causes or to a number of different causes acting jointly. [...]
It is with data afffected by numerous causes that Statistics is mainly concerned. Experiment seeks to disentangle a complex of causes by removing all but one of them, or rather by concentrating on the study of one and reducing the others, as far as circumstances permit, to a comparatively small residuum. Statistics, denied this resource, must accept for analysis data subject to the influence of a host of causes, and must try to discover from the data themselves which causes are the important ones and how much of the observed effect is due to the operation of each.” (ebda.)

[5]Sehr deutlich wird dies bei J. S. Mill, der sich maßgeblich dafür eingesetzt hat, daß Ursachen als Bedingungen aufgefaßt werden sollten. Mill geht zunächst von der Annahme eines „Kausalgesetzes“ aus: „Das Causalgesetz, dessen Erkenntnis der Grundpfeiler der inductiven Philosophie ist, besteht bloss in der allbekannten Wahrheit, dass, unabhängig von einer jeden Betrachtung bezüglich der letzten Erzeugungsweise von Naturerscheinungen und von jeder Frage nach den ‘Dingen an sich’, die Beobachtung eine Unveränderlichkeit der Succession zwischen einer Thatsache in der Natur und einer andern, die ihr vorhergegangen ist, nachweist.“ (Mill 1877, S. 406f) Kurz darauf heißt es dann: „Wenn überhaupt je, so besteht diese unveränderliche Folge nur selten zwischen einer folgenden und einer einzigen vorhergehenden Naturerscheinung, zwischen einem einzelnen Antecedens und einem Consequens, aber gewöhnlich zwischen einer folgenden und einer Summe von verschiedenen vorhergehenden Er-

es als eine Hauptaufgabe der Sozialstatistik, möglichst viele dieser Ursachen zu erfassen, d.h. statistische Gesamtheiten nach möglichst vielen Merkmalen zu zergliedern.[6]

4. Die neue Rhetorik ist vergleichsweise harmlos, wenn man sich auf artifizielle Zufallsgeneratoren beziehen kann. Man kann sich an folgendem Bild orientieren:

$$\begin{array}{ccccc} E & \longrightarrow & \boxed{\mathcal{G}} & \longrightarrow & Y \\ & & \uparrow & & \\ & & X & & \end{array} \tag{6.1.1}$$

Der Zufallsgenerator $\mathcal{G}$ wird durch Ereignisse des Typs E aktiviert und kann dadurch Sachverhalte hervorbringen, die durch Werte einer Zufallsvariablen Y charakterisiert werden können. Weiterhin wird angenommen, daß die Funktionsweise des Zufallsgenerators von gewissen Bedingungen abhängt, die durch Werte einer Variablen X erfaßt werden können.[7] Es ist dann offensichtlich, welcher Art die Wahrscheinlichkeitsaussagen sind, die durch ein Regressionsmodell gebildet werden können;

---

scheinungen, deren aller Zusammenwirken nöthig ist, um die folgenden Erscheinungen hervorzubringen, d.h. damit sie ihnen gewiss folgen. [. . .] Das Ganze dieser Umstände machte in diesem besonderen Falle die *Bedingungen* des Phänomens, oder mit anderen Worten, die Reihe von Antecedentien aus, welche dasselbe hervorriefen und ohne welche es nicht stattgefunden hätte. Die wahre Ursache ist das Ganze dieser Antecedentien und philosophisch gesprochen, haben wir kein Recht, den Namen Ursache einer einzigen von ihnen ausschliesslich der andern zu geben.“ (ebda., S. 407f) Die Formulierungen zeigen im übrigen, wie schwer es selbst Mill fällt, den Begriff einer Bedingung ernst zu nehmen; stattdessen werden Bedingungen durchgängig mit Worten wie „bewirken“ und „erzeugen“ verknüpft. — Dagegen heißt es in der „Logik“ von C. Sigwart (1911, S. 137): „Die Analyse der *Causalitätsvorstellung* muss, die Frage nach einem allgemeinen Causalprincip zunächst beiseite lassend, von der *Vorstellung des Wirkens eines Dings auf ein anderes* ausgehen, welche in der gewöhnlichen Sprache, vor allem in den transitiven Verben, überall vorausgesetzt und in den einzelnen Fällen als unzweifelhaft gegeben betrachtet wird.“

[6] Z.B. in einer Formulierung von H. Forcher (1913, S. 238f): „. . . ein ungemein wichtiges, fundamentales Postulat jeder statistischen Forschung: Größtmögliche Differenzierung der zu untersuchenden Massen. Je weiter man mit der Differenzierung der Masse geht, desto eher ist man in der Lage, den tatsächlichen Erscheinungen des Lebens auf den Grund zu sehen, desto eher sind Vermutungen über das Wesen der untersuchten Masse aus den Beobachtungen gerechtfertigt.“

[7] X ist in diesem Bild keine Zufallsvariable, sondern formal ein Parameter für den Zufallsgenerator.

sie haben die Form:

$$\mathrm{P}(Y = y \,|\, X = x) \;\approx\; f(x)$$

Aussagen dieser Form zeigen, wie die Funktionsweise des Zufallsgenerators von Bedingungen abhängt.

5. Das Bild macht weiterhin deutlich, daß sich die konditionalen Wahrscheinlichkeitsaussagen über die Funktionsweise des Zufallsgenerators nicht unmittelbar auf empirische Prozesse beziehen. Denn zu solchen Prozessen kommt es nur dadurch, daß der Zufallsgenerator durch Ereignisse des Typs E aktiviert wird. Erst dann kommt es zu einem Prozeß, der unter jeweils gegebenen, durch Werte von X repräsentierbaren Bedingungen abläuft und einen Sachverhalt hervorbringt, der durch einen Wert von Y gekennzeichnet werden kann. Somit wird auch deutlich, daß Ereignisse des Typs E und Bedingungen, die durch X erfaßt werden, eine wesentlich unterschiedliche Rolle spielen. Durch Ereignisse des Typs E kommt ein Geschehen zustande, das dann – wenn es zustande kommt – in einem Kontext abläuft, der durch Bedingungen charakterisiert werden kann. Im übrigen ist bemerkenswert, daß Ereignisse des Typs E in der statistischen Modellbildung gar nicht vorkommen. Das Regressionsmodell bezieht sich allein darauf, die Abhängigkeit der Funktionsweise eines Zufallsgenerators (mit dem auf potentiell realisierbare Prozesse verwiesen wird) von Bedingungen zum Ausdruck zu bringen.

6. Das Bild (6.1.1) setzt voraus, daß der Zufallsgenerator als ein Verfahren konzipiert werden kann. Wie in Abschnitt 3.1.1 besprochen worden ist, stößt diese Vorstellung jedoch an Sinngrenzen, wenn man sich auf Entstehungsprozesse sozialer Sachverhalte und Ereignisse beziehen möchte. Man benötigt dann ein anderes Bild, das man in folgender Weise andeuten kann:

$$\begin{array}{ccc} \boxed{\mathcal{G}\ [\text{Akteure}]} & \longrightarrow & Y \\ \uparrow & & \\ X & & \end{array} \tag{6.1.2}$$

Man hat es jetzt mit Akteuren zu tun, die unter gewissen durch X charakterisierbaren Bedingungen Sachverhalte hervorbringen, die durch Werte der Variablen Y gekennzeichnet werden können. Der Unterschied zum ersten Bild besteht darin, daß Ereignisse des Typs E unsichtbar werden; an ihre Stelle treten Akteure, die jedoch in der Konzeption des

Zufallsgenerators verschwinden. Für beide Bilder gilt zwar, daß das statistische Modell nur Bedindungen eines Prozesses repräsentieren kann, nicht jedoch den Prozeß selbst; dafür dient ja die Konzeption eines Zufallsgenerators: daß er die explizite Repräsentation eines Prozesses überflüssig machen soll. Aber indem das zweite Bild Akteure (bzw. Ereignisse des Typs E) unsichtbar werden läßt, kann leicht die Vorstellung entstehen, daß jetzt die Bedingungen (X) als Faktoren wirksam werden, die das Geschehen hervorbringen.

7. Für das statistische Modell ist die Unterscheidung zwischen den beiden Bildern gleichgültig. Es abstrahiert in jedem Fall von Akteuren (oder Ereignissen des Typs E), durch die Prozesse inganggesetzt werden. Die Aussagen des Modells beziehen sich immer nur darauf, wie Prozeßabläufe von Bedingungen abhängig sind oder sein könnten. Das ist unproblematisch, solange man sich in einem verfahrenstechnischen Kontext bewegt, in dem eine Unterscheidung zwischen Verfahren und Akteuren, die die Verfahren anwenden könnten, als eine selbstverständliche Sinnvoraussetzung gegeben ist. Irreführende Vorstellungen entstehen erst dann, wenn Akteure in das Verfahren eingebaut und dadurch unsichtbar gemacht werden.[8] Orientiert man sich am Bild (6.1.2) kann zwar immer noch der Gedanke formuliert werden, daß es Akteure sind, die Sachverhalte hervorbringen, und daß sie dies in einer Situation tun, die durch Bedingungen charakterisierbar ist. Aber dieser Gedanke kann nicht mit den begrifflichen Hilfsmitteln eines statistischen Modells formuliert werden. Im Gegenteil, das statistische Modell läßt nicht nur jede Erinnerung an Akteure verschwinden, sondern abstrahiert explizit von Akteuren, indem es ihre Tätigkeiten in der Vorstellung eines Zufallsgenerators verschwinden läßt. Nimmt man das Modell ernst, ist es „zufällig“, was die Akteure tun, und wichtig wird allein, daß ihre Betätigungen von Bedingungen abhängig sind. Mithin ist es nicht nur eine Gedankenlosigkeit, wenn die Bedingungen als (potentiell wirksam werdende) *Faktoren* bezeichnet werden. Vielmehr kommt darin zum Ausdruck, wie man sich soziale Prozesse vorstellen möchte.

[8] Hier setze ich voraus, daß man an Einsichten in Prozesse interessiert ist, an denen Akteure mitwirken. Wie schon besprochen worden ist, kann man – wie etwa bei der Verwendung statistischer Modelle in der Experimentalpsychologie – Zufallsgeneratoren auch deshalb verwenden, weil man sich nicht für Akteure, sondern nur für ihre dispositionalen Eigenschaften, soweit sie durch Verfahren erfaßbar sind, interessiert.

6.1.2 Gibt es Hinweise auf Ursachen?

1. Wir werden auf diese Überlegung später noch einmal zurückkommen. Zuvor soll auf ein anderes Problem aufmerksam gemacht werden, das durch die Rhetorik sozialstatistischer Ursachenforschung erzeugt wird: daß man über Ursachen tatsächlich gar nichts herausfindet. So verlief ja bereits das Argument bei Lexis, mit dem wir uns in Abschnitt 3.2 beschäftigt haben. Über die Ursachen, die sozialen Sachverhalten und Ereignissen zugrunde liegen, kann der Sozialstatistiker keine Auskunft geben. Lexis' Idee bestand darin, *stattdessen* soziale Situationen als Chancensysteme zu interpretieren, die mit gewissen Wahrscheinlichkeiten Sachverhalte und Ereignisse hervorbringen. Lexis hat auch schon angedeutet, wie sozialstatistische Ursachenforschung an diese Idee anknüpfen kann: indem sie sich um eine möglichst weitgehende Differenzierung sozialer Situationen bemüht. Den Sinn des Vorschlags kann man sich anhand von Bild (6.1.2) verdeutlichen. Lexis räumt ein, daß der Sozialstatistiker keine Aussagen darüber machen kann, wie soziale Sachverhalte und Ereignisse entstehen, insbesondere dann nicht, wenn sie durch Akteure hervorgebracht werden. Davon abstrahiert der Sozialstatistiker, indem er sich den Entstehungsprozeß als einen Zufallsgenerator (ein Chancensystem) vorstellbar macht. Andererseits kann jedoch der Sozialstatistiker untersuchen, wie unterschiedliche Bedingungen mit unterschiedlichen Chancensystemen korrespondieren. Z.B. kann der Sozialstatistiker herausfinden, daß es für Männer und Frauen einer gewissen Altersgruppe unterschiedliche Sterbenswahrscheinlichkeiten gibt.

2. Erinnert man sich an die Konzeption sozialstatistischer Ursachenforschung, wie sie im vorangegangenen Abschnitt anhand eines Zitats von F. Zizek erläutert wurde, sieht man, daß sie genau den Gedankengängen von Lexis entspricht. Wie schon bei Lexis stellt sich somit die Frage, ob und ggf. was man über Verursachungszusammenhänge lernen kann. Die Frage stellt sich in jedem Fall, wenn ein sozialstatistisches Regressionsmodell berechnet worden ist: Warum ist ein Chancensystem (eine so interpretierte empirische Häufigkeitsverteilung) unter Bedingungen x_1 ein anderes als unter Bedingungen x_2? Warum treten z.B. tödliche Arbeitsunfälle in einem Berufszweig x_1 mit einer größeren Häufigkeit auf als in einem Berufszweig x_2?

3. Fragen dieser Art müssen nicht unbedingt gestellt werden. Wenn es nur darum geht, Informationen zu gewinnen, um bessere Voraussagen machen zu können, sind Fragen nach Ursachen überflüssig und es genügt vollständig die Information, wie sie durch ein statistisches Regressionsmodell verfügbar gemacht wird. Z.B. ist es für ein Versicherungsunternehmen nicht wichtig, warum die Sterblichkeit in einem Berufszweig x_1 größer ist als in einem Berufszweig x_2; wichtig ist vielmehr, das Ausmaß der Unterschiede einschätzen und bei den Versicherungsnehmern unterscheiden zu können, welchem Berufszweig sie angehören. — Die Frage nach Ursachen stellt sich dagegen, wenn Menschen auf das Zustandekommen von Sachverhalten und Ereignissen Einfluß nehmen wollen. Dann aber genügt es nicht, um bei diesem Beispiel zu bleiben, daß man das Ausmaß der Sterblichkeit in unterschiedlichen Berufszweigen ermittelt. Vielmehr liefert dies bestenfalls einen ersten Hinweis auf Warum-Fragen. Wenn man weiterkommen will, muß man vielmehr die Arbeitsbedingungen in den verschiedenen Berufszweigen untersuchen und versuchen, anhand der tatsächlich eingetretenen Todesfälle herauszufinden, wie sie zustandegekommen sind und in welcher Weise ihr Zustandekommen mit den Arbeitsbedingungen zusammenhängt.

4. Indem sozialstatistische Untersuchungen Hinweise auf mögliche Verursachungszusammenhänge liefern können, sind sie sicherlich nicht nutzlos. Allerdings sollte explizit festgestellt werden, daß sie keinen Ersatz für Ursachenforschung liefern können. Insofern ist schon der Ausdruck „statistische Ursachenforschung“ irreführend. Gelegentlich wird das auch von Autoren eingeräumt, die in der Ursachenforschung eine der Hauptaufgaben der Sozialstatistik sehen. Zum Beispiel sagte F. Zizek (1928a, S. 388):

> „Zu beachten ist ferner, daß die kausalen Faktoren meist *nicht* die eigentlichen, tieferen, *letzten Ursachen* der in den Zahlen in Erscheinung tretenden Wirkungen sind; der kausale Faktor dient oft nur zur Zusammenfassung von Untersuchungseinheiten nach einem der statistischen Erhebung zugänglichen äußeren Moment. Nähere Untersuchungen, eventuell mithilfe anderer Wissenschaften, können die eigentliche Erklärung liefern; hinsichtlich des Einflusses z.B. der verschiedenen Berufe auf die Sterblichkeit können nähere Untersuchungen – auf medizinischer Grundlage – über die konkreten gesundheitlichen Gefahren der einzelnen Berufe (Staubentwicklung, schädliche Dämpfe usw.) in Betracht kommen.“

Ähnlich hat sich auch Flaskämper (1959, S. 162) geäußert:

> „Die Statistik kann uns immer nur einen Hinweis auf die eigentlichen Ursachen geben. Sie kann z.B. feststellen, daß die Sterblichkeit in einzelnen Berufen verschieden hoch ist, der Beruf also von Einfluß auf die Höhe der Sterblichkeit ist. Warum er das ist, kann uns die Statistik aber nicht sagen.“

Wenn dies eingeräumt wird, macht es allerdings keinen Sinn, die Bedingungen für soziales Geschehen, die in statistischen Modellen durch

Variablen erfaßt werden können, als „Ursachen“ oder „Faktoren“ zu bezeichnen.[9] In der Theoriebildung müßte man vielmehr von einer begrifflichen Unterscheidung zwischen einerseits Prozessen und ihren Akteuren und andererseits Bedingungen, unter denen sich die Prozesse entwickeln, ausgehen. Dadurch ließe sich auch genauer sagen, was mit statistischen Begriffsbildungen erreicht werden kann: sie können verwendet werden, um Bedingungen von Prozessen mit ihren Ergebnissen zu verknüpfen. Aber sie liefern keine begriffliche Repräsentation der Prozesse selbst.

5. Das ist unproblematisch, wenn man Prozesse als Realisationen von Verfahren (Zufallsgeneratoren) nicht nur vorstellbar machen, sondern auch empirisch explizieren kann. Ein statistisches Modell kann dann so verstanden werden, daß es Einsichten in die Funktionsweise eines Verfahrens liefern kann. In der Sozialstatistik tritt indessen ein Problem auf, auf das schon mehrfach hingewiesen wurde: daß die Prozesse, auf die sich sozialstatistische Modelle beziehen, durch Akteure zustandekommen, deren Betätigungen nur ausnahmsweise als Realisationen von Verfahren verstanden werden können. Dies schließt es nicht aus, daß auch sozialstatistische Modelle Hinweise auf Verursachungszusammenhänge liefern können. Aber es gibt in den meisten Fällen keine klaren Vorstellungen darüber, wie sich eine Untersuchung der Prozesse anschließen könnte. In der statistisch orientierten empirischen Sozialforschung ist es infolgedessen üblich geworden, die eigentliche Forschungsaufgabe nur in der Konstruktion eines statistischen Modells zu sehen und dieses Modell dann nachträglich mit einer „Interpretation“ zu deuten.[10] Aber die Nützlichkeit solcher „Interpretationen“ ist zweifelhaft. Denn wenn man ein statistisches Modell nur für Voraussagen verwenden möchte, sind solche Interpretationen überflüssig. Aber auch wenn das Erkenntnisinteresse stattdessen auf ein Verständnis sozialer Prozesse zielt, ist fragwürdig,

[9] Wann und wo die begriffliche Gleichsetzung von ‘Variablen’ und ‘Ursachen’ entstanden ist, ist unklar; aber bereits 1906 bemerkte E. Blaschke (1906, S. 10): „Mit Beziehung auf die Gesamtheiten werden die Merkmale vielfach als *Ursachen* bezeichnet. Selbstverständlich haben dieselben mit denen der Naturwissenschaften nichts gemein; aus einem Ursachenschema der Statistik kann a priori kein Resultat, nicht einmal die Erkenntnis abgeleitet werden, daß dasselbe Ursachenschema immer zum gleichen Resultate führen müßte.“

[10] Vgl. dazu bereits Stouffer 1934, S. 485: „It is repeating the obvious so say that, even if the sociologist has successfully estimated the probability of an observed relationship arising from a suitably defined population of possibilities, his task usually will not be ended. He seeks to supply a reasonable explanation of the relationship found, unless, as is seldom the case, the explanation is self-evident from the manner in which the problem was set up.“

was man lernen kann. Denn die typischerweise angebotenen „Interpretationen" beruhen nicht auf einer Untersuchung sozialer Prozesse, sondern bestehen in einer im Prinzip beliebigen Deutung der statistisch ermittelten Zusammenhänge.[11]

6.2 Gleichartigkeit und Homogenität

In einem engen Zusammenhang mit sozialstatistischer Ursachenforschung steht eine Idee, die bereits bei der Besprechnung der Auffassungen von Lexis in Abschnitt 3.2.3 erwähnt worden ist: daß der Sozialstatistiker die Dinge, Menschen oder Situationen, für die er Merkmalsverteilungen (oder Chancensysteme) ermitteln will, in möglichst „gleichartige" bzw. „homogene" Gruppen einteilen sollte. In diesem Abschnitt soll darauf etwas näher eingegangen werden.

6.2.1 Die Gleichartigkeitsforderung

1. Eine in der Literatur zur Sozialstatistik regelmäßig wiederkehrende Forderung ist, daß statistische Gesamtheiten (die durch Merkmalsverteilungen und daraus ableitbare Größen charakterisiert werden sollen) aus möglichst „gleichartigen" oder „homogenen" Elementen bestehen sollten. Z.B. sagt E. Czuber gleich zu Beginn seiner „Statistischen Forschungsmethoden" (1921):

„Die Gegenstände, von welchen hier die Rede ist, sind stets von gleicher Art. Doch kann die Gleichartigkeit verschiedenen Grades sein, je nachdem die Übereinstimmung in einer geringeren oder größeren Zahl von Merkmalen gefordert wird. Das richtet sich nach dem Zwecke der Untersuchung: der Grad der Gleichförmigkeit muß der Fragestellung angepaßt sein, soll die Antwort einen Wert haben. Zu geringe Gleichartigkeit kann dem Ergebnis jede wissenschaftliche Bedeutung benehmen.

Wenn es sich z.B. um die menschliche Körpergröße handelt, so hätte es mit Rücksicht auf die ungleiche Körperentwicklung in verschiedenen Altern und bei den beiden Geschlechtern kaum einen Sinn, wenn man ein Gemisch jugendlicher und erwachsener, männlicher und weiblicher Personen zur Grundlage nähme: mit dem gefundenen Resultate wäre kaum etwas anzufangen. Man

[11] Man kann sich auch fragen, ob nicht bereits die Idee, daß konditionale Häufigkeitsverteilungen einer Interpretation bedürfen, im Widerspruch dazu steht, sie zugleich als Verteilungsgesetze aufzufassen.

wird, um ein wissenschaftlich brauchbares Resultat zu gewinnen, einen Personenkreis verwenden müssen, der in gewissen Merkmalen: Geschlecht, Altersstufe, Abstammung und anderes einheitlich ist." (Czuber 1921, S. 1f)

Aber die Begründung für diese Gleichartigkeitsforderung bleibt auf merkwürdige Weise unbestimmt.

2. Man kann sich jedenfalls leicht klarmachen, daß die Forderung nach Gleichartigkeit für die Verwendbarkeit statistischer Begriffsbildungen gar keine Rolle spielt. Um eine statistische Variable

$$X : \Omega \longrightarrow \tilde{X}$$

definieren zu können, ist es nur erforderlich, daß jedem Element aus der Gesamtheit Ω ein Wert im Merkmalsraum $\tilde{X}$ zugeordnet werden kann. Bezieht sich z.B. $\tilde{X}$ auf die Eigenschaft Gewicht, könnte man beliebige Gegenstände zu einer Gesamtheit zusammenfassen, wobei nur vorausgesetzt werden muß, daß jedem Gegenstand ein Gewicht zugeordnet werden kann. Dann kann man eine statistische Variable bilden und alle weiteren Begriffsbildungen daran anschließen; also Häufigkeitsverteilungen bilden, Mittelwerte berechnen, usw.

3. Es ist auch bemerkenswert, daß Czuber den Begriff einer „Gleichartigkeit" nur komparativ bestimmt: eine Gesamtheit Ω_1 ist „gleichartiger" als eine Gesamtheit Ω_2, wenn die Elemente von Ω_1 in mehr Eigenschaften übereinstimmen als die Elemente von Ω_2. Sieht man genauer hin, findet man zwei Überlegungen. Die erste Überlegung ist leicht verständlich: welche Elemente zu einer statistischen Gesamtheit zusammengefaßt werden sollten, muß sich nach der Fragestellung richten. Wenn man sich z.B. für die Gewichtsverteilung bei neugeborenen Kindern interessiert, sollte man eben nur neugeborene Kinder zu einer Gesamtheit zusammenfassen. Die zweite Überlegung ist indessen schwerer verständlich: daß statistische Aussagen über eine Gesamtheit nur dann eine „wissenschaftliche Bedeutung" haben können, wenn es sich bei der Gesamtheit nicht um ein „Gemisch" handelt. Denn jede statistische Gesamtheit ist ein „Gemisch", auch z.B. eine Menge neugeborener Kinder. Nicht nur unterscheiden sie sich in ihrem Gewicht, man kann sie auch durch beliebig viele weitere Eigenschaften unterscheiden. Die Frage ist, ob und ggf. aus welchen Gründen man das tun sollte.

4. Um den potentiellen Sinn der Gleichartigkeitsforderung in der Sozialstatistik zu verstehen, können vielleicht die folgenden Ausführungen aus einem Lehrbuch der Statistik von A. Kaufmann (1913) helfen:

„*Zergliederung in homogene Teilmassen.* Unter Beobachtung des Gesetzes der großen Zahl gehört die *Zergliederung der zu untersuchenden Massen zu möglichst homogenen Teilmassen* zu den wichtigsten Postulaten der statistischen Methodologie, – sie führt zu höchst wichtigen Ergebnissen im Gebiete der Erkenntnis der Erscheinungen des sozialen Lebens und deren ursächlicher Zusammenhänge, – während die Mißachtung oder ungenügende Beachtung dieser Forderung zu den unerfreulichsten Fehlschlüssen und Mißverständnissen führen kann.

Ich möchte zunächst einige Zahlenbeispiele anführen, die ich der Westergaardschen Monographie über 'Mortalität und Morbidität' entnehme. Das eine Beispiel bezieht sich auf die Sterblichkeitsverhältnisse der verschiedenen Berufsklassen, und zwar greife ich zwei davon heraus – die Geistlichen und die Eisenbahnbeamten. Wenn man diese zwei Berufsgruppen unzergliedert nimmt, so ergeben sie, nach englischen Daten, die folgenden, auf das Tausend von Lebenden berechneten Sterblichkeitszahlen: Geistliche 19, Eisenbahnbeamte 17. Die Sterblichkeit der Geistlichen ist also *höher* als die der Eisenbahnbeamten. Dies widerspricht aber offenbar allem, was wir über die Lebensverhältnisse dieser Berufsklassen wissen; es widerspricht auch den Daten über das Durchschnittsalter der Angehörigen beider Gruppen beim Tode, welches sich für die Geistlichen mit 62,6, für die Eisenbahnbeamten nur mit 43,2 Jahren berechnet. Dieser letztere Widerspruch gibt uns aber auch einen Fingerzeig darüber, wo der Weg zur Beseitigung der beiden Widersprüche zu suchen ist. Wir zergliedern die Sterblichkeit der beiden Berufsgruppen *nach dem Alter* sowohl der Verstorbenen als der Lebenden, und gelangen zu den folgenden zwei Zahlenreihen:

Alter	Geistliche	Eisenbahner
25 – 35	5	12
35 – 45	6	15
45 – 55	13	22
55 – 65	23	41
65 – 75	52	71
75 u. darüber	150	205

Die nach dem maßgebenden Merkmal zergliederten Daten ergeben also das *direkte Gegenteil* von dem, was man aus den unzergliederten Gesamtzahlen schließen konnte: die Sterblichkeit der Eisenbahnbeamten ist in allen Altersgruppen nicht nur nicht schwächer als die der Geistlichen, sondern *viel stärker*, was auch mit den viel ungünstigeren Lebens- und Arbeitsverhältnissen der ersteren in vollem Einklang steht. Wenn die unzergliederten Gesamtzahlen eine größere Sterblichkeit der Geistlichen aufweisen, so kommt dies nur von dem viel höheren Durchschnittsalter dieser Berufsgruppe; die höhere Sterblichkeit war nicht die *der Geistlichen als solchen*, sondern die der in dieser Berufsklasse vorherrschenden *höheren Altersgruppen.*“ (Kaufmann 1913, S. 99f)

Der wesentliche Gesichtspunkt besteht darin, daß man zwei oder mehr Gesamtheiten im Hinblick auf die Werte eines Merkmalsraums vergleichen möchte. Man hat dann folgende Situation:

$$X_j : \Omega_j \longrightarrow \tilde{X}$$

wobei $j = 1, \ldots, m$ eine Reihe von Gesamtheiten indiziert. Das Ziel besteht darin, die Verteilungen der Variablen X_j zu vergleichen, wofür am einfachsten z.B. die Mittelwerte $\mathrm{M}(X_j)$ verwendet werden können.

5. Kaufmanns Argument verläuft nun so: Um zu einem sinnvollen Vergleich der Verteilungen oder auch nur ihrer Mittelwerte zu gelangen, sollten die Gesamtheiten Ω_j möglichst gleichartig sein. Versuchen wir, die Überlegung anhand des von Kaufmann angeführten Beispiels etwas genauer zu verstehen. Es gibt in diesem Beispiel drei statistische Variablen: das Alter A, die Berufsgruppe B und die Sterblichkeit S (die als eine Indikatorvariable für Todesfälle definiert werden kann). Die Daten können dann durch eine dreidimensionale Variable

$$(A, B, S) : \Omega \longrightarrow \tilde{A} \times \tilde{B} \times \tilde{S}$$

repräsentiert werden. Zunächst vergleicht Kaufmann die bedingten Mittelwerte $\mathrm{M}(S \,|\, B = 1)$ für Geistliche und $\mathrm{M}(S \,|\, B = 2)$ für Eisenbahner und stellt fest:

$$\mathrm{M}(S \,|\, B = 1) > \mathrm{M}(S \,|\, B = 2)$$

womit festgestellt wird, daß die Sterblichkeit der Geistlichen größer ist als die der Eisenbahner. Aber ist dies, wie die Bemerkungen von Kaufmann nahelegen, eine falsche Feststellung? Das kann man so nicht sagen, denn – die Daten vorausgesetzt – ist es eine durchaus richtige Feststellung über die Sterblichkeit der Personen in Ω, wenn man sie nach ihrer Berufsgruppe unterscheidet. Aber natürlich kann man die Überlegung weiterführen, indem man die Personen zusätzlich nach ihrem Alter unterscheidet. So kommt man zu der Feststellung:

$$\text{Für alle } \tilde{a} \in \tilde{A}: \quad \mathrm{M}(S \,|\, B = 1, A = \tilde{a}) < \mathrm{M}(S \,|\, B = 2, A = \tilde{a})$$

so daß man erkennt, daß in jeder Altersgruppe die Sterblichkeit der Geistlichen niedriger ist als die der Eisenbahner.

6. Die Überlegung erscheint durchaus sinnvoll. Indem man die Personen zusätzlich nach ihrem Alter unterscheidet, lernt man etwas neues über Bedingungen der Sterblichkeit. Zwar lernt man nicht unmittelbar etwas über Ursachen der Sterblichkeit; denn weder das Alter noch die Berufszugehörigkeit sind „Todesursachen". Aber man bekommt Hinweise darauf, daß Eisenbahner im Vergleich zu Geistlichen früher sterben; und man könnte diese statistische Feststellung zum Anlaß nehmen, die Ursachen dieses Sachverhalts zu untersuchen, d.h. die konkreten Prozesse, die bei Eisenbahnern und Geistlichen zu Todesfällen führen. Dann kann man vielleicht feststellen, daß es bei Eisenbahnern häufiger zu tödlichen Arbeitsunfällen kommt, und sich auch überlegen, wie dies geändert werden könnte.

6.2.2 Grenzen der Homogenisierbarkeit

1. Indem man die Personen in Ω nach ihrer Berufsgruppe unterscheidet, bekommt man zwei relativ homogenere Teilgruppen. Unterscheidet man zusätzlich nach dem Alter, bekommt man relativ noch homogenere Teilgruppen. Um auf diesen Kontext für die Verwendung des Homogenitätsbegriff zu verweisen, spreche ich von *statistischer* Homogenität. Die Definition ist einfach folgende: die Mitglieder einer Gesamtheit Ω sind statistisch homogen bezüglich eines Merkmalsraums $\tilde{Y}$, wenn alle Mitglieder von Ω den gleichen Merkmalswert in $\tilde{Y}$ haben. Indem man mehr als einen Merkmalsraum verwendet, kann man dementsprechend von größerer oder geringerer statistischer Homogenität sprechen.[12]

2. Das im vorangegangenen Abschnitt besprochene Beispiel Kaufmanns zeigt, daß statistische Gruppenvergleiche immer nur bezüglich einer vorausgesetzten Gruppenbildung gültig sind. Auch die von ihm gefundene Feststellung, daß die Eisenbahner in allen Altersgruppen eine höhere Sterblichkeit aufweisen als die Geistlichen, könnte relativiert werden. Man könnte sich eine Vielzahl weiterer Variablen ausdenken, um die Personen in Ω zu unterscheiden. Es ist auch nicht ausgeschlossen, daß man eine weitere Variable finden kann, etwa eine Indikatorvariable G mit einem Merkmalsraum $\tilde{G} = \{1, 2\}$, der sich auf „genetische Ausstattung" bezieht, so daß man zu folgendem Sterblichkeitsvergleich gelangt:

$$\text{Für alle } \tilde{a} \in \tilde{A},\ \tilde{g} \in \tilde{G}:\quad \mathrm{M}(S \mid B = 1, A = \tilde{a}, G = \tilde{g}) \ > \ \mathrm{M}(S \mid B = 2, A = \tilde{a}, G = \tilde{g})$$

Aber dadurch wird die von Kaufmann gefundene Aussage nicht falsch. Es zeigt sich zunächst nur, daß statistische Gruppenvergleiche davon

[12] Sobald man Kovariablen einführt, können natürliche unterschiedliche Varianten statistischer Homogenität definiert werden.

abhängig sind, wie die Gruppen gebildet werden.

3. Es zeigt sich indessen noch etwas: daß Werte sozialstatistischer Variablen nicht unmittelbar als Ursachen aufgefaßt werden können. Bestenfalls kann mit ihnen auf unterschiedliche *Bedingungen* verwiesen werden, also auf Eigenschaften einer Situation (eines Kontextes). Die schließlich resultierenden Sachverhalte oder Ereignisse werden dagegen durch Prozesse hervorgebracht, die sich in der Situation, im Rahmen der durch den Kontext vorgegebenen Bedingungen abspielen. Wenn man gleichwohl von Ursachen sprechen möchte, kann man natürlich Formulierungen wie „statistische Ursachen" oder kurz „s-Ursachen" verwenden. Wichtig ist jedoch, daß es sich nicht um Ursachen im gewöhnlichen Verständnis dieses Wortes, sondern um Bedingungen handelt.

4. Dieser Hinweis ist wichtig, denn durch die Vorstellung, daß es sich bei Bedingungen um Ursachen handelt, hat sich in der Sozialstatistik eine durchaus fragwürdige Auffassung verbreitet: daß jedes soziale Geschehen durch eine unübersehbar große Anzahl von Bedingungen determiniert wird.[13] Dadurch gerät nicht nur die begriffliche Unterscheidung zwischen Prozessen und Bedingungen aus dem Blickfeld, sondern sie wird durch eine neue Problemstellung substituiert: wie man möglichst viele der Bedingungen erfassen kann, von denen das Zustandekommen sozialer Sachverhalte und Ereignisse abhängt. Als zwar praktisch nicht erreichbares, jedoch theoretisch konzipierbares Ideal erscheint dann ein statistisches Modell, das alle Bedingungen erfaßt; etwa in der Form einer Mittelwertsregression

$$\mathrm{M}(Y \mid X_1 = x_1, X_2 = x_2, X_3 = x_3, \ldots) \ = \ f(x_1, x_2, x_3, \ldots)$$

so daß die Varianz der durch $X_1, X_2, X_3, \ldots$ konditionierten Variablen Y Null wird.

5. In der Frühzeit der Sozialstatistik, in der die begrifflichen Hilfsmittel der Regressionsrechnung erst rudimentär verfügbar waren, ist dieses

[13] Hinweise wurden bereits in Abschnitt 6.1.1 gegeben. Hier noch eine Formulierung aus „Statistical Reasoning in Sociology" von J. H. Mueller, K. F. Schuessler und H. L. Costner (1970, S. 239): „It is an axiom of science, as well as of common sense, that no event in nature 'just happens', but events always occur under very specific circumstances, either known or unknown. Therefore an event never should be viewed in isolation. It must be considered as a product of the joint operation of many forces, each of which contributes a variable element to the observed outcome. Thus, family size may depend upon such factors as age at marriage, level of income, extent of the mother's employment outside the home, and the religious ideology entertained by the parents."

theoretische Ideal durch die Vorstellung einer „absoluten Homogenität“ zum Ausdruck gebracht worden. Die Idee war, daß man sich zumindest vorstellen könne, Menschen im Hinblick auf alle relevanten Bedingungen zu klassifizieren. Natürlich war nur allzu sichtbar, daß es schwierig ist, diese Idee praktisch zu verfolgen. Dazu heißt es bei Kaufmann:[14]

„Als ideales Endziel einer solchen Zergliederung der statistisch behandelten Massen erscheint die Zerlegung des statistischen Zahlenmaterials in streng homogene Gruppen. ‘Es liegt nahe’ – so äußert sich über den Gegenstand v. Bortkiewicz – ‘in der Zerlegung des statistischen Materials so weit fortzufahren, bis man zu streng homogenen Gruppen gelangt. Derartige Gruppen würden bei weiterer Zerlegung Koeffizienten ergeben, die numerisch nicht mehr voneinander differieren würden (wobei von Differenzen, die durch ‘zufällige Ursachen’ bedingt werden, abgesehen werden muß) ... einem für solch eine homogene Gruppe ermittelten Koeffizienten käme gleichsam eine definitive Bedeutung zu, und es ist, als ob die betreffenden statistischen Koeffizienten für beliebige Gruppen, die ein und demselben Merkmalskomplex entsprechen, sich gleich hochstellen müßten.’ Wie aber v. Bortkiewicz am konkreten Beispiele einer eingehenden Analyse der statistischen Daten über kinderlose Ehen nachweist, ‘läßt sich eine Gleichheit aller Faktoren mit Ausnahme desjenigen, dessen Wirkungsweise im gegebenen Fall untersucht wird, im Bereich der Statistik durch entsprechende Zerlegung des Materials, streng genommen, niemals herstellen.’ Wieweit – bei der Begrenztheit der technischen Hilfsmittel der Statistik – eine Zerlegung des statistischen Zahlenmaterials auch geführt wird, dadurch kann eine sich nur in großen Zügen bewahrende Homogenität der Teilmassen erreicht werden, – eine Homogenität, welche sich nur auf einige wenige bei der gruppenweisen Tabellierung in Betracht gezogenen Merkmale bezieht.“ (Kaufmann 1913, S. 102f)

Hier wird der Eindruck erweckt, daß das Ideal – „streng homogene Gruppen“ zu bilden – infolge von unvermeidlichen Beschränktheiten des Datenmaterials nicht realisiert werden kann.[15] Man kann sich jedoch überlegen, daß das Problem nicht aus einem Mangel an Daten resultiert. Bei jeder endlichen statistischen Gesamtheit ist es grundsätzlich möglich, hinreichend viele Variablen zu definieren, so daß am Schluß eine Klassenbildung möglich wird, bei der sich in jeder Klasse nur noch ein Mitglied der Gesamtheit befindet. Dann wäre das theoretische Ideal offenbar erreicht. Aber es ist klar, daß dann statistische Begriffsbildungen, die auf eine Charakterisierung konditionaler Häufigkeitsverteilungen zielen, sinnlos werden.

[14] Man vgl. auch Zizek 1930, S. 14f.

[15] In der neueren statistischen Literatur wird das Problem gelegentlich als „curse of high dimensionality“ diskutiert: „Consider, e.g. a typical case in medical diagnosis where you are entering discrete information about the indicants of patients into a contingency table of perhaps 100 dimensions. By the time you take a small sample of about a billion patients you find that most of the cells are empty.“ (Good 1976, S. 337)

6. Hier setzt indessen ein neuer Gedankengang ein, der aus der Basisvorstellung der theoretischen Sozialstatistik resultiert: daß man sich die Prozesse, durch die soziale Sachverhalte und Ereignisse hervorgebracht werden, als Realisationen von Zufallsgeneratoren vorstellen könne. Als Andeutung kommt der Gedanke bereits in dem eben angeführten Zitat von Kaufmann zum Ausdruck: daß eine statistische Gesamtheit im Hinblick auf einen Merkmalsraum $\tilde{Y}$ dann „streng homogen“ genannt werden könne, wenn die Werte der korrespondierenden Variablen Y in der Gesamtheit nur noch „zufällig“ variieren. Man kann sich zwar leicht klarmachen, daß auch diese Vorstellung in eine Sackgasse führt; denn man kann sich ja grundsätzlich für jede statistische Variable einen hypothetischen Zufallsgenerator denken, der ihre Werte hervorgebracht haben könnte, und dann folgt aus der Denkvoraussetzung, daß ihre Werte nur „zufällig“ variieren. Da indessen der Gedankengang in der Geschichte der theoretischen Sozialstatistik von zentraler Bedeutung geworden ist, soll darauf im folgenden noch etwas näher eingegangen werden.

6.2.3 Probabilistische Homogenität

1. Ein probabilistischer Homogenitätsbegriff entsteht mit der Vorstellung, daß Eigenschaften einer Menge von Dingen, Menschen oder Situationen durch Prozesse zustandegekommen sind, die als Realisationen eines Zufallsgenerators gedeutet werden können. Man denke an eine statistische Variable

$$Y : \Omega \longrightarrow \tilde{Y}$$

Jedes Mitglied von Ω hat einen bestimmten Wert im Merkmalsraum $\tilde{Y}$, und im allgemeinen gibt es auch inbezug auf diesen Merkmalsraum keine statistische Homogenität. Jedenfalls ist dies eine Sinnvoraussetzung für statistische Begriffsbildungen, mit denen gezeigt werden soll, wie die Merkmalswerte in der Gesamtheit Ω variieren; diesem Zweck dient die Berechnung und Darstellung einer Häufigkeitsverteilung für die Variable Y. Die in den vorangegangenen Abschnitten besprochenen Überlegungen können erst einsetzen, wenn weitere Variablen zur Verfügung stehen, durch die die Mitglieder von Ω unterschieden werden können. Dann kann man konditionale Verteilungen bilden und statistische Vergleiche anstellen.

2. Dagegen beginnt die Idee probabilistischer Homogenität mit einer gedanklichen Bezugnahme auf Prozesse, durch die die Mitglieder von Ω jeweils bestimmte Werte im Merkmalsraum $\tilde{Y}$ bekommen. Die Idee ist: die Mitglieder von Ω sind homogen bezüglich des Merkmalsraums $\tilde{Y}$, wenn bzw. insoweit die Prozesse, durch die sie jeweils bestimmte Werte bekommen, gleich (gleichartig) sind. Die Idee einer probabilistischen Homogenität knüpft an diese Vorstellung durch die Annahme an, daß es sich bei den Prozessen um Realisationen eines Zufallsgenerators handelt. Dann gelangt man zu folgender Definition: Die Mitglieder von Ω sind bezüglich des Merkmalsraums $\tilde{Y}$ *probabilistisch homogen*, wenn jedes Mitglied von Ω seinen Merkmalswert durch den gleichen Zufallsgenerator bekommen hat; oder in der Chancenrhetorik von Lexis: wenn alle Mitglieder von Ω im Hinblick auf ihre Werte in $\tilde{Y}$ dem gleichen Chancensystem unterliegen.

3. Eine vergleichsweise klare Definition dieser Idee probabilistischer Homogenität findet sich bei J. v. Kries (1886, S. 103f):

> „Unter der Gleichartigkeit soll hier verstanden sein, dass alle Fälle unter einen und denselben Begriff zu subsummiren sind und dieselben Möglichkeiten des Verlaufs darbieten; ausserdem soll vorausgesetzt werden, dass für unbegrenzt fortzusetzende Reihen gewisse dauernd constante, also für sämmtliche Einzelfälle zutreffende Bedingungen bestehen.“

Hier werden also Dinge, Situationen oder Ereignisse als gleichartig bezeichnet, wenn sie *durch den gleichen Zufallsgenerator* hervorgebracht werden. Zwar verwendet v. Kries nicht das Wort ‘Zufallsgenerator’; zur Erläuterung seiner Definition verwendet er jedoch Hinweise auf Zufallsspiele.

4. Ob es sich um eine nützliche Definition handelt, sei dahingestellt. Man kann ihr jedenfalls dann einen Sinn geben, wenn man auf einen empirisch explizierbaren Zufallsgenerator verweisen kann, durch den Sachverhalte entstanden sind. Ein Problem entsteht jedoch in der theoretischen Sozialstatistik, denn die von ihr unterstellte Zufallsgeneratoren können nicht empirisch expliziert werden, sondern sind begriffliche Fiktionen, um durch sie Entstehungsprozesse empirisch ermittelter Häufigkeitsverteilungen zu deuten. Infolgedessen wird aber auch aus der Idee probabilistischer Homogenität eine reine Spekulation. Denn da es diese Zufallsgeneratoren nicht gibt, gibt es auch keine empirischen Anhaltspunkte, um zu entscheiden, ob eine Menge von Daten aus dem gleichen oder aus unterschiedlichen Zufallsgeneratoren resultiert. Es ist deshalb nicht verwunderlich, daß die Idee probabilistischer Homogenität in der Sozialstatistik hauptsächlich nur zu merkwürdigen Spekulationen geführt hat.[16]

5. Zunächst war auch die Unterscheidung zwischen statistischer und probabilistischer Homogenität nicht immer klar. Noch in den 1920er Jahren gab es ausgedehnte Diskussionen über die Begriffe „Gleichartigkeit“ und „Homogenität“.[17] Ein Nachvollzug dieser Diskussionen ist deshalb schwierig, weil die Begriffsbildungen fast immer mit schwer verständlichen Vorstellungen über Ursachen verbunden wurden. Hier noch einmal ein Zitat von F. Zizek:

> „Es sind jene Einheiten (Individuen usw.) auszuscheiden, welche das infrage stehende Erhebungsmerkmal überhaupt nicht aufweisen, vielleicht gar nicht aufweisen können. Erst wenn dies geschehen ist, liegt eine homogene Masse vor, während sich anderenfalls ein Mittelwert ergibt, der keinen einheitlichen Ursachenkomplex widerspiegelt [...].
>
> Die eben besprochene ‘Ausscheidung unbeteiligter Teilmassen’ genügt oft nicht, um hinsichtlich der darzustellenden Erscheinung homogene Massen zu erzielen.

[16]Exemplarisch sei auf L. H. C. Tippett (1937) hingewiesen: „When the binomial and Poisson distributions fit any given experimental data, it is inferred that the variations are due to chance and cannot be reduced or controlled unless the whole character of the data can be altered by introducing some method of selecting individuals of a given kind.“ (S. 53) Wie schon in Abschnitt 2.5.1 besprochen worden ist, war in der Frühzeit der Sozialstatistik der Gedanke verbreitet, daß auch die Normalverteilung in ähnlicher Weise ein Argument für „zufällige Variation“, d.h. probabilistische Homogenität, liefern könne. Dazu heißt es bei Tippett (S. 61): „The normal curve has also been deduced mathematically as the distribution that results from the combination of an infinite number of very small random errors, and this fact together with the fact that the distribution is a special case of the binomial gives it a fundamental status. It is believed by some to represent the deviations due to experimental errors in measurements of a physical constant, and a quantity that is distributed according to this law is sometimes held to be subject only to chance causes that cannot be controlled. Conversely, where a distribution is skew, it is often inferred that superimposed on the chance variations are some larger ones due to a few important causes that may be controlled. We are dubious of this use of the normal distribution. Doubtless normality may be made a definition, and hence a test, of ‘randomness’ just in the way that conformity to the binomial and Poisson distributions has been interpreted, but it is doubtful if for the normal curve such a course has any practical significance. Frequently, when man has done all he can to control the variation in a character and the remaining variations are practically random, the resulting distribution is far from normal.“

[17]Einen exemplarischen Eindruck vermitteln die Diskussionsbeiträge von Flaskämper (1928, 1929) und Zizek (1928, 1930).

Denn auch innerhalb der beteiligten Gesamtheit gibt es oft wieder Teilmassen, welche die betreffende Erscheinung infolge der Verschiedenheit allgemeiner Ursachen in sehr verschiedenem Ausmaße aufweisen. Die beiden Geschlechter, die verschiedenen Altersstufen, Berufe, Wohlstandsschichten, Stadt und Land usw. zeigen meist wesentliche Unterschiede. Da ist es wünschenswert, 'besondere' Mittelwerte für solche homogenere Teilmassen zu gewinnen, die dann als zahlenmäßiger Ausdruck eines bestimmten einheitlichen Ursachenkomplexes gelten können.“ (Zizek 1921, S. 142f)

Ausführungen dieser Art kann man mit einem statistischen Homogenitätsbegriff rekonstruieren, so wie dies im vorangegangenen Abschnitt anhand des Beispiels von Kaufmann schon verdeutlicht worden ist. Die „besonderen“ Mittelwerte Zizeks sind dann nichts anderes als durch eine Reihe weiterer Variablen bedingte Mittelwerte. Die statistische Begriffsbildung läßt sich in diesem Fall von der Ursachenrhetorik trennen. Wenn Zizek davon spricht, daß ein bedingter Mittelwert der zahlenmäßige Ausdruck für einen „Ursachenkomplex“ ist, handelt es sich einfach um eine nachträglich angehängte Interpretation, die durch statistische Variablen erfaßbare Bedingungen für soziale Prozesse als deren Ursachen deutet.

6. Aber durch die theoretische Sozialstatistik ist das Reden von Ursachen noch in einer weiteren Weise modifiziert und unklar gemacht worden. Es hat sich nämlich eine Rhetorik entwickelt, in der sich der Ursachenbegriff nicht auf statistische Variablen, sondern auf die hypothetischen Zufallsgeneratoren bezieht, die zur Deutung der Entstehungsprozesse sozialer Sachverhalte und Ereignisse angenommen werden. Und mit diesem Reden von Ursachen verbindet sich die Idee der probabilistischen Homogenität. Soweit ich es verstehen kann, liegt der Rhetorik folgende Vorstellung zugrunde. Man hat als Ausgangspunkt eine Reihe statistischer Gesamtheiten bzw. Variablen:

$$Y_j : \Omega_j \longrightarrow \tilde{Y}$$

Die Fragestellung bezieht sich auf den Entstehungsprozeß der für diese Variablen definierbaren Häufigkeitsverteilungen (oder daraus ableitbare Mittelwerte usw.). Das Reden von Ursachen knüpft an die Vorstellung an, daß man für jede Variable einen Zufallsgenerator $\mathcal{G}_j$ unterstellen kann, durch den ihre Werte zustandegekommen sind. Diese Zufallsgeneratoren werden als die „wesentlichen“ oder „allgemeinen“ (Zizek) oder „gesamtheitlichen“ (Winkler) Ursachen der empirischen Häufigkeitsverteilungen bezeichnet. Um das Reden von Ursachen in diesem Kontext von statistischen Ursachen (s-Ursachen) zu unterscheiden, spreche ich von *p-Ursachen*. Folgt man also dieser Rhetorik, verdankt sich die Verteilung

einer Variablen, inbezug auf deren Merkmalsraum eine Gesamtheit probabilistisch homogen ist, per Definition der gleichen p-Ursache, und es kann sich die Frage anschließen, ob sich die Häufigkeitsverteilungen der Variablen Y_j der gleichen oder unterschiedlichen p-Ursachen verdanken.

7. Die Unterscheidung zwischen statistischer und probabilistischer Homogenität ist in der sozialstatistischen Literatur oft schwer erkennbar. Als Beispiel beziehe ich mich auf eine Abhandlung über „Gleichartigkeit, Homogenität und Gleichwertigkeit in der Statistik“ von F. Zizek (1928):

„Eine erste Kategorie bildet die formale (begriffliche, methodische) Gleichartigkeit [...]. Statistische Einheiten, die durch statistische Erhebung erfaßt wurden, sind gleichartig, wenn sie demselben Begriffe der Erhebungseinheit entsprechen; sie stimmen dann hinsichtlich aller Begriffsmerkmale der Erhebungseinheit überein. Nur solche gleichartige Einheiten – mag es sich um ortsanwesende Personen oder um Arbeitslose oder Betriebe oder Warensendungen oder Streiks usw. handeln – bilden eine statistische Masse, die einheitlich weiterbearbeitet und zahlenmäßig charakterisiert werden kann.“ (Zizek 1928, S. 394)

Insoweit handelt es sich einfach darum, eine statistische Gesamtheit durch irgendeine Menge von Merkmalen begrifflich abzugrenzen. Dann unterscheidet Zizek zwei Varianten des Begriffs 'homogen'. Bei seiner ersten Variante handelt es sich tatsächlich nur um eine Weiterführung seines Begriffs der Gleichartigkeit:

„Wir sprechen in der Statistik häufig bei der Gruppenbildung auf Grund eines sachlich-qualitativen Merkmals von 'homogeneren Teilmassen', wir sind bei der Gruppenbildung – z.B. wenn wir die Berufe klassifizieren oder Gruppen von Warengattungen oder Todesursachen bilden – bestrebt, 'möglichst homogene' Gruppen zu bilden, d.h. die sachlich zusammengehörigen, verwandten Einzelfälle – und *nur* diese – gruppenweise zu höheren sachlichen Einheiten zusammenzufassen [...]. *Welche Fälle* in der einzelnen konkreten Statistik als einander nahestehend anzusehen und in dieselbe Gruppe zu vereinigen sind, hängt von der die Gruppenbildung beherrschenden materiellen Zielsetzung ab; ein bestimmtes Gruppierungsprinzip entscheidet darüber, welche Berufe, Gewerbe, Warengattungen, Krankheiten usw. als 'einander nahestehende', als 'miteinander verwandte' zu behandeln sind, damit die uns interessierenden Gruppen zur Aussonderung gelangen.“ (Zizek 1928, S. 395)

Die in diesem Zitat erläuterte Variante von „Homogenität“ fällt wie die zuvor erläuterte „Gleichartigkeit“ unter unseren Begriff statistischer Homogenität: Gesamtheiten werden anhand von Werten statistischer Variablen gebildet bzw. gegliedert.

8. Dann führt Zizek jedoch eine zweite Bedeutungsvariante des Ausdrucks „homogen" ein, die sich deutlich unterscheidet:

„Einen ganz anderen Sinn hat die statistische Homogenität, wenn man dabei an die Verursachung der untersuchten Erscheinung, an die sie bestimmenden 'allgemeinen' oder – in Winklers Terminologie – 'gesamtheitlichen', auch 'wesentlichen', 'konstanten' Ursachen denkt. Unter homogenen Massen in diesem Sinne verstehe ich Massen mit einem nach bestimmten Richtungen, die noch angeführt werden sollen, einheitlichen Komplex allgemeiner (gesamtheitlicher) Ursachen." (Zizek 1928, S. 396)

Dazu gibt es folgende Anmerkung:

„Diese Ursachenkategorie habe ich in meiner Abhandlung 'Ursachenbegriffe und Ursachenforschung' im 17. Bd. dieser Zeitschrift (S. 388ff) eingehend besprochen [vgl. Zizek 1928a]; es handelt sich dabei immer um sehr mannigfaltige, niemals erschöpfend aufzählbare Ursachen, um äußerst komplizierte 'Ursachen – oder Bedingungskomplexe' (z.B. der Sterblichkeit oder Arbeitslosigkeit); jede statistische Zahl bringt das Ergebnis des Zusammenwirkens ihrer allgemeinen Ursachen, bringt die Resultierende ihres Ursachenkomplexes zum Ausdruck. Den Gegensatz zu den 'allgemeinen Ursachen' bilden die 'zufälligen (individuellen, variablen) Ursachen', die sich, wenn das Gesetz der großen Zahlen zur Geltung kommt, gegenseitig aufheben, d.h. die statistische Zahl nicht beeinflussen; diese soll ja kein Zufallsergebnis sein, sondern die Wirksamkeit der allgemeinen, wesentlichen Ursachen zum Ausdruck bringen." (ebda.)

Wenn ich diese Ausführungen richtig verstehe, zielt Zizeks zweite Bedeutung von „homogen" auf, in meiner Formulierung, probabilistische Homogenität. Aber um das genauer diskutieren zu können, muß zuerst besprochen werden, wie in der Wahrscheinlichkeitstheorie das Reden von p-Ursachen – und eine korrespondierende Unterscheidung zwischen „wesentlichen" und „zufälligen" Ursachen – vorbereitet worden ist. Im Anschluß daran wird man nicht nur die Ausführungen von Zizek (und ganz ähnliche Formulierungen bei vielen anderen Autoren) besser verstehen können, sondern auch ihren Zusammenhang mit der von Lexis vorbereiteten Vorstellung, daß man auch Chancen als Ursachen betrachten könne.

6.3 Zufallsgeneratoren als Ursachen

Wir besprechen zuerst die aus der Wahrscheinlichkeitstheorie stammende Unterscheidung zwischen „wesentlichen" und „zufälligen" Ursachen und gehen dann darauf ein, wie sie in der theoretischen Sozialstatistik als ein Deutungsschema verwendet worden ist.

6.3.1 Wesentliche und zufällige Ursachen

1. Die Unterscheidung hat ihre Ursprünge in der Wahrscheinlichkeitstheorie. Der Kontext besteht darin, daß man sich auf Ereignisse oder Sachverhalte bezieht, die durch Zufallsgeneratoren zustande kommen. Die Frage, wie dann zu erklären sei, warum bestimmte Sachverhalte oder Ereignisse auftreten, motiviert die Sprechweisen. Die Eigenschaften des Zufallsgenerators, soweit sie in einer bestimmten Wahrscheinlichkeitsverteilung zum Ausdruck kommen, gelten als die „wesentlichen Ursachen", die übrigen Umstände, die dazu führen, das einmal dies, einmal jenes Ereignis eintritt, gelten als die „zufälligen Ursachen".

2. Eine Fülle von Beispielen für diesen Sprachgebrauch findet sich zunächst bei Laplace (1814). Dann verbreiten sich die Sprechweisen in der Literatur zur Wahrscheinlichkeitsrechnung. Zum Beispiel beginnen die „Vorlesungen über Wahrscheinlichkeitsrechnung" von A. Meyer, die 1879 von E. Czuber ins Deutsche übersetzt wurden, mit folgenden Bemerkungen:

„Jedes Ereignis ist die Folge des Zusammenwirkens zweier Arten von Umständen; die einen, bekannt oder unbekannt, sind *nothwendig* zu seiner Hervorbringung, während die anderen stets unbekannten nur zufällig dazu beitragen. Die Umstände der ersten Art nennt man Ursachen oder Chancen des Ereignisses; die anderen in ihrer Gesammtheit bilden das, was man als Zufall bezeichnet. Beispiel. Besteht das Ereignis in dem Ziehen einer weissen Kugel aus einer Urne, welche weisse und schwarze Kugeln enthält, so sind Ursachen oder Chancen für das Eintreffen einer weissen Kugel: die Gesammtzahl aller Kugeln und die Zahl der weissen; die Anordnung der Kugeln in der Urne sowie der Vorgang beim Ergreifen einer derselben fallen in den Bereich des Zufalls." (Meyer 1879, S. 1)

Meyer erläutert weiterhin, wie er sich den Zusammenhang zwischen Ursachen und Wahrscheinlichkeiten vorstellt:

„Unter den Umständen, welche an der Hervorbringung eines Ereignisses mitwirken, gibt es solche, die sich mit jedem Augenblicke ändern und deren Kenntnis uns gänzlich abgeht, wie etwa die Bewegung, welche die Hand den Würfeln ertheilt; die Vereinigung dieser Umstände ist es, was wir als Zufall bezeichnen. Die anderen Umstände sind constant, wie etwa die Geschicklichkeit der Spieler, die Zahl der gleichen Seiten eines Würfels u.s.w. Diese führen die objectiven Wahrscheinlichkeiten der Ereignisse herbei, und die mehr oder minder ausgebreitete Kenntnis, welche wir von ihnen besitzen, bestimmt deren subjective Wahrscheinlichkeiten. Allein die constanten Umstände reichen zur Hervorbringung der Ereignisse nicht hin, dazu ist vielmehr ihre Vereinigung mit den variablen Umständen erforderlich; sie können blos die Wahrscheinlichkeiten der Ereignisse vergrössern, ohne deren Eintreffen mit Nothwendigkeit herbeizuführen.“ (Meyer 1879, S. 7)

3. Soweit sich aus der Literatur ersehen läßt, hat sich dieser merkwürdige Sprachgebrauch in den Schriften zur Wahrscheinlichkeitstheorie und theoretischen Sozialstatistik weitgehend verbreitet. Ein wichtiger Vermittler war wiederum Quetelet. Unter Berufung auf v. Kries stellte v. Bortkiewicz (1895, S. 321) fest:

„Es entspricht dem Standpunkte der Wahrscheinlichkeitsrechnung, als maßgebend für das Eintreten oder Ausbleiben eines Ereignisses, um dessen Wahrscheinlichkeit es sich handelt, gewisse Umstände als ‘Allgemeinbedingungen’ und andere wiederum als ‘zufällige Ursachen’ hinzustellen.“

Solche und ähnliche Formulierungen finden sich auch oft in der statistischen Literatur. Es ist auch bemerkenswert, wie sich bei vielen Autoren die Unterscheidung zwischen „wesentlichen“ und „zufälligen“ Ursachen mit der Vorstellung verbindet, daß jedes Ereignis durch eine undurchdringliche Vielzahl von Ursachen hervorgerufen wird. Z.B. in den folgenden Ausführungen von J. H. Mueller, K. F. Schuessler und H. L. Costner (1970, S. 209f):

“We conceive of any event as a resultant of innumerable factors, which may be classified into two broad categories: *determining* factors and *chance* factors. The determining factors are those which have been isolated and to which we attribute the explanation or causation of the event. Thus, we have learned that divergent cultural backgrounds of married couples may predispose toward divorce; broken homes are associated with specific types of delinquency; occupational mobility is dependent upon educational attainment. Identification of determining factors is never complete and in fact may consist only of plausible hypotheses to be checked. In any case, the focus of science is always on the determining factors, since they supply the basis for understanding and predicting events.
Chance factors are the remaining or unknown factors affecting an event.”

4. In der frühen Literatur zur Sozialstatistik wurde die Unterscheidung zumeist im Zusammenhang mit einem „Gesetz der großen Zahlen“ rezipiert; z.B. von W. Winkler mit seiner Idee einer „statistischen Wesensform“ (vgl. Abschnitt 2.5.1), die er auch als Ausdruck der wesentlichen, von ihm „gesamtheitlich“ genannten Ursachen interpretiert;[18] aber

[18] Vgl. auch Winkler 1925, S. 11ff.

auch von F. Zizek und vielen anderen Autoren.[19] Um die Rezeption zu illustrieren, sei hier auf Ausführungen von P. Flaskämper hingewiesen, die er auch bis in die letzte Auflage seines Lehrbuchs (1959) übernommen hat:

„Das Gesetz der großen Zahlen in der mathematischen Formulierung läßt sich mit völliger Strenge beweisen. Warum gilt es aber mit weitgehender Annäherung in der Erfahrung, und in welchen Fällen gilt es da? Die Bedingungen sind am durchsichtigsten bei den Glücks- und Zufallsspielen. Wir haben es hier zu tun mit zweierlei Ursachenkomplexen, die verschieden benannt werden und die wir uns an dem Urnenbeispiel klarmachen wollen. Daß sich bei einer genügenden Zahl von Ziehungen immer wieder dasselbe oder annähernd dasselbe Verhältnis ergibt, jenes, das dem Mischungsverhältnis der Kugeln in der Urne entspricht, und zwar um so genauer, je größer die Zahl der Ziehungen ist, das hängt zusammen mit den Spielbedingungen (konstantes Mischungsverhältnis, zurücklegen der nichtgezogenen Kugeln und erneutes Mischen). Wir nennen diese Ursachen die allgemeinen Ursachen oder den *allgemeinen Ursachenkomplex* (es handelt sich ja nicht um *eine* Ursache, sondern eine ganze Gruppe solcher). [...]
Daß sich nun aber nicht immer haarscharf dasselbe Ergebnis einstellt, daß bei kleinen Serien von Ziehungen sogar sehr starke Abweichungen vorkommen, hängt zusammen mit wechselnden ‘zufälligen’ Einflüssen, in dem Urnenbeispiel davon, wie der Ziehende ‘zufällig’ in die Urne greift, welche Kugeln ‘zufällig’ nach der Mischung nach oben zu liegen kommen usw. Wir nennen diese meist sehr viel undurchsichtigeren und schwerer beschreibbaren Ursachen den *zufälligen* oder *besonderen Ursachenkomplex*. Wir stoßen hier auf das Wort Zufall, das in der Statistik eine sehr große Rolle spielt. Zufall in diesem Sinne bedeutet nicht ursachloses Geschehen; gemeint sind vielmehr Zusammenhänge, deren Verkettung im einzelnen völlig undurchsichtig ist. Der Zufall überlagert die durch die allgemeinen Ursachen bedingte Gestaltung des Ergebnisses und verzerrt es. Daher die Abweichungen der Ergebnisse von den wahrscheinlich-

[19] Hier zum besseren Verständnis eine Erläuterung von F. Zizek (1928a, S. 389): „Die methodologische Bedeutung der ‘allgemeinen Ursachen’ können wir richtig nur verstehen, wenn wir uns das ‘Gesetz der großen Zahlen’ vor Augen halten. Dieses Gesetz besagt, daß wir durch Zusammenfassung größerer Mengen von Beobachtungen zu Zahlenwerten gelangen, welche ds – uns als Statistiker interessierende – Ergebnis der in der betreffenden Beobachtungsmasse wirksamen ‘allgemeinen’ (‘konstanten’) Ursachen widerspiegeln, also das für die betreffende Masse Allgemeingültige zum Ausdruck bringen und unserer Erkenntnis zugänglich machen, während einzelne Fälle und kleine Beobachtungsmassen nicht nur von den allgemeinen, sondern auch von individuellen (‘zufälligen’, ‘störenden’) Momenten beeinflußt werden und daher mannigfache Abweichungen aufweisen.“ Der Gedankengang setzt ersichtlich die gedankliche Bezugnahme auf einen Zufallsgenerator voraus, der die Elemente der statistischen Gesamtheit bzw. ihre jeweils bestimmten Eigenschaft hervorgebracht hat.

keitstheoretisch zu erwartenden.“ (Flaskämper 1959, S. 25f)

5. Versuchen wir, die Unterscheidung zu verstehen. Gegeben ist ein Zufallsgenerator $\mathcal{G}$, den man verwenden kann, um n Werte einer Zufallsvariablen X^* zu erzeugen. Den realisierten Werten entspricht eine statistische Variable X, deren Häufigkeitsverteilung man berechnen kann. Man kann dann fragen: Wie ist es zu dieser Häufigkeitsverteilung gekommen? Dann kann man zwei Aspekte unterscheiden. Zunächst ist klar, daß die Ursache dafür, daß überhaupt eine Menge von Sachverhalten entstanden ist, für deren Charakterisierung hinterher eine Häufigkeitsverteilung berechnet werden kann, darin liegt, daß der Zufallsgenerator von irgendeinem Akteur aktiviert worden ist. Eine andere Frage bezieht sich darauf, warum die empirische Häufigkeitsverteilung eine bestimmte Form angenommen hat. Man könnte antworten: weil beim erstenmal ein Sachverhalt s_1 mit dem Merkmalswert x_1, beim zweitenmal ein Sachverhalt s_2 mit dem Merkmalswert x_2, usw. entstanden ist; schließlich sind auf diese Weise die Sachverhalte $s_1, \ldots, s_n$ entstanden, und die Häufigkeitsverteilung resultiert aus dem Rechenverfahren, das auf die korrespondierenden Merkmalswerte $x_1, \ldots, x_n$ angewendet wird. Eine andere Art von Erklärung wird dadurch möglich, daß die Sachverhalte mit einem Verfahren erzeugt worden sind. Man kann versuchen, die resultierende Häufigkeitsverteilung durch Eigenschaften des Verfahrens zu erklären. Da es sich um einen Zufallsgenerator handelt, gelingt das nicht unmittelbar. Aber man kann den Zufallsgenerator durch ein Verteilungsgesetz charakterisieren und dann dieses Verteilungsgesetz verwenden, um eine probabilistische Erklärung für die empirische Häufigkeitsverteilung anzubieten (vgl. Abschnitt 2.1).

6. Allerdings wäre es irreführend, dieses Verteilungsgesetz die Ursache für die empirisch realisierte Häufigkeitsverteilung zu nennen. Denn sie ist ersichtlich nicht durch das Verteilungsgesetz erzeugt worden, sondern durch Akteure, die den Zufallsgenerator aktiviert haben. Das Verteilungsgesetz charakterisiert nur das Verfahren, das für die Erzeugung der Sachverhalte verwendet worden ist. Mit dem Wort ‘Verteilungsgesetz’ wird also nicht auf Ursachen verwiesen, sondern *auf Bedingungen* für Prozesse, die jeweils bestimmte Sachverhalte erzeugen.

7. Man kann also das Reden von „wesentlichen“ oder „allgemeinen“ Ursachen in gewisser Weise als einen in der Formulierung verfehlten Hinweis auf Eigenschaften eines Zufallsgenerators verstehen. Dagegen ist das Reden von „zufälligen Ursachen“ vollständig obskur. Es ist bereits schwierig, eine Fragestellung zu finden, auf die diese Redeweise bezogen

werden könnte. In dem oben angeführten Zitat bezieht Flaskämper das Reden von „zufälligen Ursachen“ auf die Frage, warum die empirische Häufigkeitsverteilung von der Wahrscheinlichkeitsverteilung, die den Zufallsgenerator charakterisiert, abweicht. Dies ist indessen gleichermaßen eine Eigenschaft des (jedes) Zufallsgenerators. Die meisten Autoren, so auch in dem eingangs angeführten Zitat von A. Meyer, beziehen sich dagegen auf eine andere Frage: warum bei der Aktivierung eines Zufallsgenerators ein ganz bestimmter Sachverhalt ensteht oder entstanden ist. Als Erklärung wird angeboten, daß der Sachverhalt durch die Anwendung eines Verfahrens zustande kommt, das jedoch so beschaffen ist, daß sich auch eine Vielzahl von „zufälligen Einflüssen“ geltend machen kann.

8. Schon die Fragestellung widerspricht indessen der Idee eines Zufallsgenerators. Ein Zufallsgenerator wird verwendet, und sei es auch nur als eine gedankliche Konstruktion, um einen Kontext zu erzeugen, in dem die Frage, warum es bei einer Aktivierung des Zufallsgenerators zu einer bestimmten Realisierung kommt, nicht sinnvoll gestellt werden kann. Wenn man sich gleichwohl auf diese Frage bezieht und ein jeweils realisiertes Ergebnis einerseits der Verwendung des Zufallsgenerators und andererseits „dem Zufall“ zuschreibt, resultiert nur eine obskure Antwort auf eine sinnlose Frage. Es ist auch bemerkenswert, daß diese Frage in der Wahrscheinlichkeitstheorie gar nicht gestellt wird, jedenfalls nicht in ihrem Theoriegebäude, das vielmehr darauf beruht, daß die Frage nach Ursachen eines jeweils bestimmten Ergebnisses im Rahmen ihres theoretischen Ansatzes nicht sinnvoll gestellt werden kann. Man kann zwar im Rahmen dieser Theorie sinnvoll davon sprechen, daß Sachverhalte und Ereignisse zufällig entstehen. Aber damit ist dann nur gemeint, daß sie tatsächlich oder in Gestalt einer gedanklichen Fiktion durch die Aktivierung eines Zufallsgenerators zustandegekommen sind. Wenn dennoch von „zufälligen Ursachen“ gesprochen wird, ist das nur ein obskurer und schon in der sprachlichen Formulierung verfehlter Hinweis darauf, daß man an ein deterministisches Weltbild glaubt und diesem Glauben rhetorischen Tribut zollen möchte.[20]

[20]Deutlich wird dies bereits in dem oben angeführten Zitat von Flaskämper. Noch etwas ausführlicher hat sich E. Czuber (1908, S. 7f) so geäußert: „Bei der Analyse des Urnenbeispiels haben wir neben einem positiven Wissen Umstände erkannt, die bei der Realisierung der allgemeinen Bedingungen, d.i. bei Ausführung eines Zuges, auftretend sich einzeln und im Zusammenwirken unserer Kenntnis entziehen und uns verhindern, die Notwendigkeit für das Eintreten oder für das Ausbleiben eines bestimmten Erfolges zu erkennen. Wenn wir den Erfolg oder das Ereignis, auf welches unser Interesse sich richtet, mit Rücksicht auf diese Umstände als ein *zufälliges* bezeichnen, so negieren wir damit *die Notwendigkeit* seines Eintreffens; und wenn wir

6.3.2 Chancen anstelle von Ursachen

1. Die Rhetorik der „zufälligen Ursachen“ kann im weiteren auch deshalb außer Betracht bleiben, weil sich statistische Ursachenforschung erklärtermaßen auf die „wesentlichen“ oder „allgemeinen“ Ursachen richtet. Wie schon in den Abschnitten 6.1 und 6.2 anhand von Beispielen und Zitaten deutlich gemacht wurde, verlangt die Suche nach solchen Ursachen nicht unbedingt die Annahme, daß Sachverhalte und Ereignisse durch einen Zufallsgenerator zustande kommen. Man kann von statistischen Variablen ausgehen und versuchen, durch die Bildung bedingter Häufigkeitsverteilungen Hinweise auf Bedingungen zu bekommen, unter denen sich soziale Prozesse entwickeln. — Eine neue Ursachenrhetorik entsteht in der theoretischen Sozialstatistik durch die Annahme, daß man die Prozesse, durch die soziale Sachverhalte und Ereignisse entstehen, als Realisationen von Zufallsgeneratoren auffassen könne. Dadurch kann das sozialstatistische Reden von Ursachen an in der Wahrscheinlichkeitstheorie entwickelte Sprechweisen anschließen, insbesondere an die Unterscheidung zwischen „wesentlichen“ und „zufälligen“ Ursachen.

sagen, sein Eintreffen hänge vom *Zufall* ab, so meinen wir hierunter das Zustandekommen einer solchen Verbindung der voneinander unabhängigen Umstände, welche das Ereignis wirklich herbeiführt. Ist beispielsweise unser Interesse dem Erscheinen der weißen Kugel zugewandt, so bezeichnen wir es als zufällig, daß bei der uns unbekannten Lagerung der Kugeln und der davon unabhängigen Bewegung der Hand diese gerade die weiße Kugel erfassen sollte, und wenn wir ein solches Geschehen als durch den Zufall hervorgebracht erklären, so fassen wir den *Zufall* als ein *gesetzlos Wirkendes* im Gegensatze zu einem *gesetzmäßig Wirkenden*, einer *Ursache*, auf; aus der gesetzlosen Wirksamkeit des Zufalls erklären wir auch die *Unregelmäßigkeit* in einer Reihe von Erfolgen.
Indessen ist die Vorstellung von dem Walten eines ‘blinden Zufalls’ unvereinbar mit unserer Überzeugung von dem kausalen Zusammenhange alles Geschehens, eine bloße Fiktion. Jedes Geschehen ist die Folge eines vorausgegangenen andern und die Veranlassung eines künftigen, und so ist alles Geschehen nach dem Prinzip der Ursächlichkeit untereinander verbunden. [...]
Mit dem Worte Zufall soll also der kausale Zusammenhang nicht geleugnet, durch seinen Gebrauch vielmehr das Zugeständnis gemacht werden, daß uns der Zusammenhang in dem betreffenden Falle überhaupt nicht erkennbar (Auftreten einer Epidemie und Sichtbarkeit eines Kometen) oder so verwickelt sei (wie etwa beim Würfeln), daß wir über seine bloße Konstatierung nicht hinauskommen.
Es ist üblich, die in das Gebiet der Wahrscheinlichkeitstheorie fallenden Ereignisse, Erscheinungen, Tatbestände dadurch allgemein zu charakterisieren, daß ihr Eintreffen oder ihre Existenz einerseits von bekannten und bleibenden und andererseits von unbekannten und wechselnden Umständen abhänge. Die Umstände der ersten Art bezeichnet man als *Ursachen*, die der zweiten Art als *Zufall*.“

2. Um diskutierbar zu machen, wie der Gedankengang vielleicht verstanden werden kann, sei zunächst von einer statistischen Variablen

$$(X,Y) : \Omega \longrightarrow \tilde{X} \times \tilde{Y}$$

ausgegangen. Der Einfachheit halber sei auch angenommen, daß Y eine Indikatorvariable mit dem Merkmalsraum $\tilde{Y} = \{0,1\}$ ist, mit der das Auftreten von Ereignissen erfaßt wird. Ein parametrisches Regressionsmodell kann dann folgendermaßen konzipiert werden:

$$h(Y = 1 \,|\, X = x) \approx g(x;\theta)$$

D.h. es wird nach einer parametrischen Funktion $g(x;\theta)$ gesucht (z.B. in der Form einer Logitfunktion), mit der dargestellt werden kann, wie die relativen Häufigkeiten, mit denen die Variable Y den Wert 1 annimmt, von Werten der Variablen X abhängen. Als Motiv für diese Modellkonstruktion dient die Vermutung, daß die Werte von X als Bedingungen (s-Ursachen) für Prozesse interpretiert werden können, die die durch $\tilde{Y}$ erfaßten Ereignisse hervorgebracht haben.

3. Ein formal paralleler Gedankengang der theoretischen Sozialstatistik beginnt mit der Annahme, daß es sich bei diesen Prozessen um Realisationen eines Zufallsgenerators handelt. Die Modellformulierung nimmt dann folgende Form an:

$$\mathrm{P}(Y^* = 1 \,|\, X^* = x) \approx g(x;\theta)$$

Hier ist jetzt Y^* eine Zufallsvariable, so daß man von Wahrscheinlichkeitsaussagen sprechen kann.[21] Das verlangt allerdings, daß ein Zufallsgenerator eingeführt wird, durch den die Zufallsvariable Y^* erläutert werden kann. Wie in Abschnitt 5.4.2 besprochen worden ist, kann folgendes Bild verwendet werden:

$$x \longrightarrow \boxed{\mathcal{G}(x)} \longrightarrow y$$

Die Idee ist, daß es für jeden Wert $x \in \tilde{X}$ einen Zufallsgenerator $\mathcal{G}(x)$ gibt, der – wenn dieser Wert als Bedingung gegeben ist – Werte der Zufallsvariablen Y^* erzeugen kann.

[21]Der logische Status von X^* hängt davon ab, ob man zunächst von der Vorstellung einer zweidimensionalen Zufallsvariablen ausgeht oder von vornherein eine Konditionalisierung vornimmt, wodurch X^* zu einer logischen Variablen für den Merkmalsraum $\tilde{X}$ wird.

4. An diesem Bild setzt nun das Reden von „probabilistischen Ursachen“ an: Werte im Merkmalsraum $\tilde{X}$ sind jetzt mögliche p-Ursachen für die Wahrscheinlichkeitsverteilung von Y^*. Offenbar verlaufen die Gedankengänge weitgehend analog. Der Unterschied kommt auch zunächst nur in einer unterschiedlichen Rhetorik zum Ausdruck: im ersten Fall sind Werte in $\tilde{X}$ mögliche s-Ursachen, jetzt sind es mögliche p-Ursachen. Der Übergang zur p-Rhetorik impliziert jedoch eine Idee, die das Reden von Ursachen noch einmal verdreht: daß es zu jedem Wert $x \in \tilde{X}$ einen Zufallsgenerator $\mathcal{G}(x)$ gibt, der die Werte der Variablen Y^* erzeugt. Als p-Ursachen erscheinen infolgedessen die fiktiven Zufallsgeneratoren $\mathcal{G}(x)$. Die Rhetorik kann natürlich unterschiedliche Formen annehmen. Wohl am meisten verbreitet hat sich das von Lexis vorgeschlagene Reden von Chancen. Die möglichen Werte im Merkmalsraum $\tilde{X}$ verweisen dann auf Chancensyteme $\mathcal{G}(x)$, die die Wahrscheinlichkeiten bestimmen, mit denen Y^* ihre möglichen Werte annimmt.

5. Hier können sich auch Überlegungen über probabilistische Homogenität anschließen. Im Anschluß an die in Abschnitt 6.2.3 gegebene Definition kann der Gedanke jetzt so formuliert werden: Die Zufallsvariable Y^* ist probabilistisch homogen in bezug auf den Merkmalsraum $\tilde{X}$, wenn gilt:

$$\text{Für alle } x, x' \in \tilde{X}: \; \mathrm{P}(Y^* = 1 \,|\, X^* = x) \approx \mathrm{P}(Y^* = 1 \,|\, X^* = x')$$

Da es sich in unserem Beispiel bei Y^* um eine Indikatorvariable handelt, ist diese Bedingung mit der folgenden äquivalent:[22]

$$\text{Für alle } x, x' \in \tilde{X}: \; \mathrm{E}(Y^* \,|\, X^* = x) \approx \mathrm{E}(Y^* \,|\, X^* = x')$$

Mithilfe dieser Definition kann man z.B. folgende Fragestellung bei v. Bortkiewicz (1895, S. 333) verstehen:

„... ob etwaige Differenzen zwischen je zwei Werten einer bestimmten Relativzahl [relativen Häufigkeit] oder Schwankungen bei einer Reihe solcher Werte aus der Wirkung zufälliger Ursachen oder aber aus der Verschiedenheit der den betreffenden Werten zu Grunde liegenden abstrakten Wahrscheinlichkeitsgrößen resp. Allgemeinbedingungen zu erklären seien.“

6. Auch v. Bortkiewicz ist sich allerdings im Klaren darüber, daß dies mit Ursachenforschung im üblichen Sinne des Wortes wenig zu tun hat; denn er bemerkt in der gleichen Abhandlung (1895, S. 331), daß

[22]Bei allgemeinen Zufallsvariablen kann man unterschiedliche Arten probabilistischer Homogenität dadurch unterscheiden, welche Eigenschaften einer Wahrscheinlichkeitsverteilung für den Vergleich berücksichtigt werden sollen.

„... es im höchsten Grade Sache subjektiven Ermessens ist, wie man sich die Beschaffenheit der Allgemeinbedingungen resp. der zufälligen Ursachen vorzustellen habe.“[23]

Dann fügt er jedoch hinzu:

„In Wirklichkeit kommt es aber darauf an, daß die Allgemeinbedingungen nicht nur oder überhaupt nicht ihrem Inhalte nach, sondern zugleich oder ausschließlich *ihrem wahrscheinlichkeitsrechnerischen Ausdruck nach* fixiert sind.“

Wenn ich diese Bemerkung richtig verstehe, wird mit ihr gesagt, daß auch das Erkenntnisinteresse nicht auf Einsichten in Verursachungszusammenhänge zielt, sondern auf Wahrscheinlichkeitsverteilungen, die für Erwartungsbildungen verwendet werden können.[24]

7. Hier zeigt sich noch einmal, daß man sich aus unterschiedlichen Gründen für „sozialstatistische Ursachenforschung“ interessieren kann. Einerseits kann sich das Erkenntnisinteresse darauf beschränken, daß man sozialstatistische Modelle für Erwartungsbildungen nutzen möchte. Dann kann man einen rein „instrumentalistischen“ Standpunkt einnehmen, bei dem das einzige Erfolgskriterium darin besteht, ob und wie gut sich ein statistisches Modell für Voraussagen verwenden läßt. Andererseits kann man versuchen, statistische Methoden und Modelle als Hilfsmittel zu verwenden, um zu Einsichten in soziale Prozesse zu gelangen. Erst wenn man von diesem letzteren Erkenntnisinteresse ausgeht, wird es wichtig, über ein genaues Verständnis der Möglichkeiten und Grenzen statistischer Begriffsbildungen nachzudenken.

[23]Ähnlich wiederholt F. Böhm (1916, S. 135): „Es ist nun allerdings Sache des Übereinkommens, was man unter allgemeinen Bedingungen verstehen will.“

[24]Dem entspricht folgende Formulierung von E. Czuber (1908, S. 171): „Es ist in der Wahrscheinlichkeitstheorie üblich geworden, die für die Erwartungsbildung maßgebenden Umstände im Hinblick auf ein ungewisses Ereignis seine *Ursache* zu nennen. In diesem Sinne spricht man auch von den möglichen Ursachen eines beobachteten Ereignisses und meint darunter die verschiedenen Bedingungskomplexe, aus welchen das Ereignis hervorgegangen sein kann.“

6

Literatur

Aitchison, J., Brown, J. A. C. 1957. The Lognormal Distribution with Special Reference to its Uses in Economics. Cambridge: Cambridge University Press.

Alterman, H. 1969. Counting People. The Census in History. New York: Harcourt, Brace & World.

Altschul, E. 1913. Studie über die Methode der Stichprobenerhebung. Ein Beitrag zur Charakteristik der modernen Strömungen in der theoretischen Statistik. Archiv für Rassen- und Gesellschaftsbiologie 10, 110–157.

Anderson, O. 1929. Zur Problematik der empirisch-statistischen Konjunkturforschung. Bonn: Schroeder.

Anderson, O. 1935. Einführung in die mathematische Statistik. Wien: Julius Springer.

Anderson, O. 1947. Zum Problem der Wahrscheinlichkeit a posteriori in der Statistik. Schweizerische Zeitschrift für Volkswirtschaft und Statistik 83, 489–518.

Anderson, O. 1961. Das „Als Ob“ in der statistischen Methodenlehre. Studi in Onore di Corrado Gini, I, 43–51. Reprint in: Ausgewählte Schriften, Band 2, 952–960. Tübingen: Mohr 1963.

Anderson, O. 1957. Probleme der statistischen Methodenlehre in den Sozialwissenschaften. Würzbuerg: Physica-Verlag.

Anderson, O. 1963. Ausgewählte Schriften. 2 Bände. Tübingen: Mohr.

Arminger, G. 1990. Testing against Misspecification in Parametric Rate Models. In: K. U. Mayer, N. B. Tuma (eds.), Event History Analysis in Life Course Research, 253–266. Madison: University of Wisconsin Press.

Bartholomew, D. J. 1995. What is Statistics? Journal of the Royal Statistical Society A 158, 1–20.

Berger, P. L. 1963. Einladung zur Soziologie. Eine humanistische Perspektive. München: DTV 1977.

Bernert, C. 1983. The Career of Causal Analysis in American Sociology. British Journal of Sociology 34, 230–254.

Bernoulli, J. 1713. Wahrscheinlichkeitsrechnung (Ars conjectandi). Dt. Übers. von R. Haussner, Leipzig 1899. Reprint Frankfurt: H. Deutsch 1999.

Billeter, E. P. 1970. Grundlagen der repräsentativen Statistik. Wien: Springer-Verlag.

Blaschke, E. 1906. Vorlesungen über mathematische Statistik. Leipzig: Teubner.

Blind, A. 1953. Probleme und Eigentümlichkeiten sozialstatistischer Erkenntnis. Allgemeines Statistisches Archiv 37, 301–313.

Böhm, F. 1916. Über das Wesen, die Aufgaben und die Ziele der mathematischen Statistik. Allgemeines Statistisches Archiv 10, 119–140.

Bortkiewicz, L. v. 1895. Kritische Betrachtungen zur theoretischen Statistik, 2. Artikel. Jahrbücher für Nationalökonomie und Statistik 10, 321–360.

Bortkiewicz, L. v. 1898. Das Gesetz der kleinen Zahlen. Leipzig: Teubner.

Bortkiewicz, L. v. 1915. Realismus und Formalismus in der mathematischen Statistik. Allgemeines Statistisches Archiv 9, 225–256.

Bortkiewicz, L. v. 1917. Die Iterationen. Ein Beitrag zur Wahrscheinlichkeitstheorie. Berlin: Springer-Verlag.

Bowley, A. L. 1926. Elements of Statistics. 5th ed. London: King & Son.

Box, G. E. P. 1976. Science and Statistics. Journal of the American Statistical Association 71, 791–799.

Bruns, H. 1906. Wahrscheinlichkeitsrechnung und Kollektivmasslehre. Leipzig: Teubner.

Burgdörfer, F. (Hg.) 1940. Die Statistik in Deutschland nach ihrem heutigen Stand. 2 Bände. Berlin: Verlag für Sozialpolitik, Wirtschaft und Statistik.

Carlsson, G. 1951. Sampling, Probability and Causal Inference. Theoria 17, 139–154.

Carver, H. C. 1921. The Mathematical Representation of Frequency Distributions. Quarterly Publication of the American Statistical Association 17, 720–731, 885–892.

Charlier, C. V. L 1920. Vorlesungen über die Grundzüge der mathematischen Statistik. Lund: Verlag Scientia.

Cochran, W. G. 1977. Sampling Techniques. 3rd. ed. New York: Wiley.

Cochran, W. G., Mosteller, F., Tukey, J. W. 1954. Principles of Sampling. Journal of the American Statistical Association 49, 13–35.

Cohen, J. B. 1938. The Misuse of Statistics. Journal of the American Statistical Association 33, 657–674.

Coleman, J. S. 1981. Longitudinal Data Analysis. New York: Basic Books.

Coolidge, J. L. 1927. Einführung in die Wahrscheinlichkeitsrechnung. Leipzig: Teubner.

Cox, D. R. 1990. Role of Models in Statistical Analysis. Statistical Science 5, 169–174.

Cox, D. R., Hinkley, D. V. 1974. Theoretical Statistics. London: Chapman & Hall.

Cox, D. R., Snell, E. J. 1981. Applied Statistics. Principles and Examples. London: Chapman & Hall.

Cox, D. R., Wermuth, N. 1996. Multivariate Dependencies. Models, Analysis and Interpretation. London: Chapman & Hall.

Czuber, E. 1908. Wahrscheinlichkeitsrechnung. Band 1: Wahrscheinlichkeitstheorie, Fehlerausgleichung, Kollektivmasslehre. Leipzig: Teubner.

Czuber, E. 1921. Die statistischen Forschungsmethoden. Wien: Seidel.

Czuber, E. 1923. Die philosophischen Grundlagen der Wahrscheinlichkeitsrechnung. Leipzig: Teubner.

d'Alembert, J. 1756. Artikel „Experimentell“. In: Artikel aus der von Diderot und d'Alembert herausgegebenen Enzyklopädie, 433–437. Hrsg. von M. Naumann. Frankfurt: Röderberg 1984.

D'Agostino, R. B., Stephens, M. A. 1986. Goodness-of-Fit Techniques. New York: Dekker.

Daston, L. 1988. Classical Probability in the Enlightenment. Princeton: Princeton University Press.

Deming, W. E. 1950. Some Theory of Sampling. New York: Dover.

Dempster, A. P. 1987. Probability and the Future of Statistics. In: I. B. MacNeill, G. J. Umphrey, eds., Foundations of Statistical Inference, 1–7. Dordrecht: Reidel.

Dobson, A. J. 1990. An Introduction to Generalized Linear Models. London: Chapman & Hall.

Durkheim, E. 1895. Die Regeln der soziologischen Methode. Hrsg. und eingel. von R. König. Frankfurt: Suhrkamp 1984.

Durkheim, E. 1950. Physik der Sitten und des Rechts. Vorlesungen zur Soziologie der Moral. Dt. Übers., hrsg. von H.-P. Müller. Frankfurt: Suhrkamp 1991.

Edgeworth, F. Y. 1908. On the Probable Errors of Frequency-Constants. Journal of the Royal Statistical Society 71, 381–397, 499–512, 651–678.

Elderton, W. P. 1938. Frequency Curves and Correltion. 3rd ed. Cambridge: Cambridge University Press.

Estes, W. K. 1991. Statistical Models in Behavioral Research. Hillsdale: Lawrence Erlbaum.

Fechner, G. T. 1897. Kollektivmasslehre. Hrsg. von G. F. Lipps. Leipzig: Engelmann.

Fienberg, S. E. 1985. Comments on and Reactions to Freedman, Statistics and the Scientific Method. In: W. M. Mason, S. E. Fienberg, eds., Cohort Analysis in Social Research, 371–383. New York: Springer-Verlag.

Fisher, A. 1928. The Mathematical Theory of Probabilities and Its Application to Frequency Curves and Statistical Methods, 2nd ed. New York: Macmillan.

Fisher, R. A. 1922. On the Mathematical Foundations of Theoretical Statistics. Philosophical Transactions of the Royal Society of London. Series A, Vol. 222, 309–368.

Fisher, R. A. 1951. Statistics. In: A. E. Heath, ed., From Scientific Thought in the Twentieth Century (London: Watts 1951). Reprint in: Collected Papers of R. A. Fisher, V, 185–207. Ed. by J.H. Bennett. University of Adelaide 1973.

Fisher, R. A. 1955. Statistical Methods and Scientific Induction. Journal of the Royal Statistical Society B 17, 69–78.

Fisher, R. A. 1966. The Design of Experiments (8th ed). New York: Hafner Publ.

Flaskämper, P. 1927. Die Statistik und das Gesetz der grossen Zahlen. Allgemeines statistisches Archiv 16, 501–514.

Flaskämper, P. 1928. Beitrag zu einer Theorie der statistischen Massen. Allgemeines Statistisches Archiv 17, 538–556.

Flaskämper, P. 1929. Das Problem der „Gleichartigkeit'" in der Statistik. Allgemeines Statistisches Archiv 19, 205–234.

Flaskämper, P. 1959. Grundriß der Sozialwissenschaftlichen Statistik. Teil I: Allgemeine Statistik. Hamburg: Meiner.

Forcher, H. 1913. Die statistische Methode als selbständige Wissenschaft: Leipzig: Veit & Comp.

Freedman, D. 1985. Statistics and the Scientific Method. In: W. M. Mason, S. E. Fienberg, eds., Cohort Analysis in Social Research, 343–366. New York: Springer-Verlag.

Freedman, D. 1997. Some Issues in the Foundation of Statistics. In: B. C. van Fraassen, ed., Topics in the Foundations of Statistics, 63–67. Dordrecht: Kluwer.

Freudenthal, H., Steiner, H.-G. 1966. Aus der Geschichte der Wahrscheinlichkeitstheorie und der mathematischen Statistik. In: H. Behnke, G. Bertram, R. Sauer, Grundzüge der Mathematik, Band IV: Praktische Methoden und Anwendungen der Mathematik, 149–195. Göttingen: Vandenhoeck & Ruprecht.

Fry, T. C. 1928. Probability and Its Engineering Uses. New York: van Nostrand.

Galton, F. 1889. Natural Inheritance. London: Macmillan.

Galton, F. 1908. Inquiries into Human Faculty and Its Development, 2nd ed. London and New York. Reprint by AMS Press, New York 1973.

Galtung, J. 1967. Theory and Methods of Social Research. London: Allen & Unwin.

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., Krüger, L. 1989. The Empire of Chance. How Probability Changed Science and Everyday Life. Cambridge: Cambridge University Press.

Good, I. J. 1976. Diskussionsbemerkung. In: W. L. Harper, C. A. Hooker, eds., Foundations and Philosophy of Statistical Inference, 337. Dordrecht: Reidel.

Gumbel, E. J. 1978. Bortkiewicz, Ladislaus von. In: W. H. Kruskal, J. M. Tanur, eds., International Encyclopedia of Statistics, Vol. 1, 24–27. New York: Free Press.

Hagood, M. J. 1941. The Notion of a Hypothetical Universe. Reprint in: D. E. Morrison, R.Ė. Henkel, eds., The Significance Test Controversy, 65–78. Chicago: Aldine Publ. 1970.

Hagood, M. J., Price, D. O. 1952. Statistics for Sociologists. Revised Ed. New York: Holt, Rinehart and Winston 1952.

Hansen, M. H., Dalenius, T., Tepping, B. J. 1985. The Development of Sample Surveys of Finite Populations. In: A. C. Atkinson, S. E. Fienberg, eds., A Celebration of Statistics. The ISI Centenary Volume, 327–354. New York: Springer-Verlag.

Hansen, M. H., Hurwitz, W. N., Madow, W. G. 1953. Sample Survey Methods and Theory. 2 Volumes. New York: Wiley.

Hartwig, H. 1956. Naturwissenschaftliche und sozialwissenschaftliche Statistik. Zeitschrift für die gesamte Staatswissenschaft 112, 252–266.

Heidelberger, M. 1987. Fechner's Indeterminism: From Freedom to Laws of Chance. In: L. Krüger, L. J. Daston, M. Heidelberger, eds., The Probabilistic Revolution, I, 117–156. Cambridge: MIT Press.

Heiss, K.-P. 1978. Lexis, Wilhelm. In: W. H. Kruskal, J. M. Tanur, eds., International Encyclopedia of Statistics, Vol. 1, 507–512. New York: Free Press.

Helm, G. 1902. Die Wahrscheinlichkeitslehre als Theorie der Collectivbegriffe. Annalen der Naturphilosophie 1, 364–381.

Henry, N. W. 1997. Thoughts on the Concept and Application of Statistical Models. In: B. C. van Fraassen, ed., Topics in the Foundations of Statistics, 63–67. Dordrecht: Kluwer.

Hinkley, D. V., Reid, N. 1991. Statistical Theory. In: D. V. Hinkley, N. Reid, E. J. Snell, eds., Statistical Theory and Modelling, 1–29. London: Chapman & Hall.

Hochstädter, D. 1991. Statistische Methodenlehre. Ein Lehrbuch für Wirtschafts- und Sozialwissenschaftler. Frankfurt: Verlag H. Deutsch.

Jeffreys, H. 1948. Theory of Probability. 2nd ed. Oxford: Clarendon Press.

Jensen, A. 1926. Report on the Representative Method in Statistics. Proceedings of the International Statistical Institute 22, 359–451.

Jevons, W. S. 1877. The Principles of Science. Reprint of 2nd ed., introduced by E. Nagel. New York: Dover 1958.

John, V. 1884. Geschichte der Statistik. Erster Teil: Von dem Ursprung der Statistik bis auf Quetelet (1835). Stuttgart: Enke. Reprint Wiesbaden: Sändig 1968.

John, V. 1895. Statistik und Probabilität. Allgemeines Statistisches Archiv 4, 1–46.

Johnson, N. L., Kotz, S. 1997. Leading Personalities in Statistical Science. From the Seventeenth Century to the Present. New York: Wiley.

Kaufmann, A. 1913. Theorie und Methoden der Statistik. Tübingen: Mohr.

Kellerer, H. 1953. Theorie und Technik des Stichprobenverfahrens. Einzelschriften der Deutschen Statistischen Gesellschaft, Nr. 5. München.

Kellerer, H. 1960. Statistik im modernen Wirtschafts- und Sozialleben. Reinbek: Rowohlt.

Kendall, M., Stuart, A. 1977. The Advanced Theory of Statistics, Vol. 1, 4th ed. London: Charles Griffin & Comp.

Kiär, A. N. 1898. Die repräsentative Untersuchungsmethode. Allgemeines Statistisches Archiv 5, 1–22.

Kish, L. 1965. Survey Sampling. New York: Wiley.

Knies, C. G. A. 1850. Die Statistik als selbständige Wissenschaft. Zur Lösung des Wirrsals in der Theorie und Praxis dieser Wissenschaft. Reprint Frankfurt: Sauer & Auvermann 1969.

Kries, J. von 1886. Die Principien der Wahrscheinlichkeitsrechnung. Freiburg: Mohr.

Kriz, J. 1973. Statistik in den Sozialwissenschaften. Reinbek: Rowohlt.

Kreyszig, E. 1968. Statistische Methoden und ihre Anwendungen (3. Auflage). Göttingen: Vandenhoeck & Ruprecht.

Krüger, L., Daston, L. J., Heidelberger, M. (eds.) 1987. The Probabilistic Revolution, Vol. 1. Cambridge: MIT Press.

Krüger, L., Gigerenzer, G., Morgan, M. S. (eds.) 1987. The Probabilistic Revolution, Vol. 2. Cambridge: MIT Press.

Krümmel, K. 1921. Anthropometrie und Statistik. Allgemeines Statistisches Archiv 13, 97–108.

Kruskal, W., Mosteller, F. 1980. Representative Sampling IV: the History of the Concept in Statistics, 1895–1939. International Statistical Review 48, 169–195.

Laplace, P. S. de 1814. Philosophischer Versuch über die Wahrscheinlichkeit. Hier zitiert nach der von R. von Mises hrsg. dt. Übers. Leipzig: Akademische Verlagsgesellschaft 1932.

Lexis, W. 1875. Einleitung in die Theorie der Bevölkerungsstatistik. Strassburg.

Lexis, W. 1877. Theorie der Massenerscheinungen in der menschlichen Gesellschaft. Freiburg.

Lexis, W. 1903. Abhandlungen zur Theorie der Bevölkerungs- und Moralstatistik. Jena: Fischer.

Lexis, W. 1909. Anthropologie und Anthropometrie. In: J. Conrad, L. Elster, W. Lexis, E. Loening, Handwörterbuch der Staatswissenschaften, Bd. 1, 523–544. Jena: Fischer.

Linder, A. 1964. Statistische Methoden für Naturwissenschaftler, Mediziner und Ingenieure. Basel: Birkhäuser.

Lindner, R., Wohak, B., Zeltwanger, H. 1984. Planen, Entscheiden, Herrschen. Vom Rechnen zur elektronischen Datenverarbeitung. Reinbek: Rowohlt.

Lindsey, J. K. 1973. Inferences from Sociological Survey Data. A Unified Approach. Amsterdam: Elsevier.

Lippert, P. 1910. Quetelet, Lambert Adolf Jakob. In: Handwörterbuch der Staatswissenschaften, 3. Aufl., Bd 6, 1277–1279. Jena: Fischer.

Loether, H. J., McTavish, D. G. 1974. Inferential Statistics for Sociologists. Boston: Allyn and Bacon.

Longford, N. T. 1993. Random Coefficient Models. Oxford: Clarendon Press.

Mallows, C. 1998. The Zeroth Problem. The American Statistician 52, 1–9.

Marschak, J. 1950. Statistical Inference in Economics. In: T. C. Koopmans, ed., Statistical Inference in Dynamic Economic Models, 1–50. New York: Wiley.

Mayer, K. U., Huinink, J. 1990. Alters-, Perioden- und Kohorteneffekte in der Analyse von Lebensverläufen oder: Lexis ade? In: K. U. Mayer, Lebensverläufe und sozialer Wandel, 442–459. Opladen: Westdeutscher Verlag.

Mayr, G. v. 1895. Statistik und Gesellschaftslehre. Band 1: Theoretische Statistik. Freiburg und Leipzig: Mohr.

Mayr, G. v. 1907. Die Berechtigung der „Moralstatistik“. Allgemeines Statistisches Archiv 7, 1–20.

Menges, G. 1959. Stichproben aus endlichen Gesamtheiten. Frankfurt: Klostermann.

Meyer, A. 1879. Vorlesungen über Wahrscheinlichkeitsrechnung. Dt. Übers. von E. Czuber. Leipzig: Teubner.

Mill, J. S. 1877. System der deductiven und inductiven Logik. Dt. Übers. von J. Schiel. Braunschweig: Vieweg.

Mises, R. 1957. Probability, Statistics and Truth. London: Allen & Unwin. Reprint: Dover (New York, 1981).

Moivre, A. de 1756. The Doctrine of Chances: or, A Method of Calculating the Probabilities of Events in Play (3rd ed). London: A. Millar.

Mood, A. M., Graybill, F. A., Boes, D. C. 1974. Introduction to the Theory of Statistics. 3rd ed. Singapore: McGraw-Hill.

Morgan, M. S. 1990. The History of Econometric Ideas. Cambridge: University Press.

Morrison, D. E., Henkel, R. E. (eds.) 1970. The Significance Test Controversy. Chicago: Aldine Publ. 1970.

Most, O. 1962. Allgemeine Statistik. Baden-Baden: Verlag August Lutzeyer.

Mueller, J. H., Schuessler, K. F., Costner, H. L. 1970. Statistical Reasoning in Sociology. 2nd ed. New York: Houghton Mifflin.

Nadler, S. (ed.) 1993. Causation in Early Modern Philosophy. University Park: Pennsylvania State University Press.

Nagel, E. 1939. Principles of the Theory of Probability. International Encyclopedia of Unified Science, Vol. I, No. 6. Chicago: University of Chicago Press.

Nelson, W. 1982. Applied Life Data Analysis. New York: Wiley.

Newman, J. R. 1956. The World of Mathematics. 4 vol. New York: Simon and Schuster.

Neyman, J. 1934. On the Two Different Aspects of the Representative Method: the Method of Stratified Sampling and the Method of Purposive Selection. Journal of the Royal Statistical Society 97, 558–625.

Neyman, J. 1937. Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability. Philosophical Transactions of the Royal Society (A), No. 236, 333–380.

Neyman, J. 1950. First Course in Probability and Statistics. London: Constable and Company.

Neyman, J. 1971. Foundations of Behavioristic Statistics. In: V. P. Godambe, D. A. Sprott, eds., Foundations of Statistical Inference, 1–19. Toronto: Holt, Rinehart and Winston.

Neyman, J., Pearson, E. S. 1928. On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference, Part I. Biometrika 20, 175–240.

O'Muircheartaigh, C. A. 1977. Statistical Analysis in the Context of Survey Research. In: C. A. O'Muircheartaigh, C. Payne, The Analysis of Survey Data, Vol. 1, 1–40. Ney York: Wiley

Ondar, K. O. (ed.) 1981. The Correspondence Between A. A. Markov and A. A. Chuprov on the Theory of Probability and Mathematical Statistics. New York: Springer-Verlag.

Pearson, K. 1896. Mathematical Contributions to the Theory of Evolution, III: Regression, Heredity, and Panmixia. Philosophical Transactions of the Royal Society of London. Ser. A 187, 254–318.

Pearson, K. 1900. The Grammar of Science. London: Adam and Charles Black.

Petty, W. 1986. Schriften zur politischen Ökonomie und Statistik. Übers. und hrsg. von W. Görlich. Berlin: Akademie-Verlag.

Pfanzagl, J. 1972. Allgemeine Methodenlehre der Statistik I (5. Auflage). Berlin: de Gruyter.

Polya, G. 1954. Patterns of Plausible Inference. Mathematics and Plausible Reasoning, vol. II. Princeton: Princeton University Press.

Porter, T. M. 1985. The Mathematics of Society: Variation and Error in Quetelet's Statistics. British Journal for the History of Science 18, 51–69.

Porter, T. M. 1986. The Rise of Statistical Thinking 1820–1900. Princeton: Princeton University Press.

Quenouille, M. H. 1950. Introductory Statistics. Oxford: Pergamon Press.

Quetelet, A. 1835. Über den Menschen und die Entwicklung seiner Fähigkeiten oder Versuch einer Physik der Gesellschaft. Übers. von V. A. Riecke. Stuttgart 1838.

Ramsey, F. P. 1931. The Foundations of Mathematics and other Logical Essays. London: Routledge & Kegan Paul.

Rao, C. R. 1995. Was ist Zufall? Statistik und Wahrheit. München: Prentice Hall.

Rassem, M., Stagl, J. (eds.) 1994. Geschichte der Staatsbeschreibung. Berlin: Akademie-Verlag.

Reichenbach, H. 1951. Probability Methods in Social Science. In: D. Lerner, H. D. Lasswell, eds., The Policy Sciences, 121–128. Stanford: Stanford University Press.

Ross, F. A. 1934. The Use of Statistical Data and Techniques in Sociology. In: L. L. Bernard, ed., The Fields and Methods of Sociology, 458–475. New York: Ray Long & Richard Smith.

Sakamoto, Y., Ishiguro, M., Kitagawa, G. 1986 Akaike Information Criterion Statistics. Dordrecht: Reidel.

Särndal, C.-E., Swensson, B., Wretman, J. 1992. Model Assisted Survey Sampling. New York: Springer-Verlag.

Schaich, E. 1990. Schätz- und Testmethoden für Sozialwissenschaftler. München: Vahlen.

Scheuch, E. K., Rüschemeyer, D. 1956. Soziologie und Statistik. Kölner Zeitschrift fuer Soziologie und Sozialpsychologie 8, 272–291. Hier zitiert nach dem Wiederabdruck in: E. Topitsch (Hg.), Logik der Sozialwissenschaften, 345–363. Köln: Kiepenheuer & Witsch 1972.

Schlittgen, R. 1996. Statistische Inferenz. München: Oldenbourg.

Schmetterer, L. 1966. Einführung in die mathematische Statistik. 2. Aufl. Wien: Springer-Verlag.

Schneider, I. (ed.) 1988: Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfängen bis 1933. Einführungen und Texte. Darmstadt: Wissenschaftliche Buchgesellschaft.

Schnell, R., Hill, P. B., Esser, E. 1999. Methoden der empirischen Sozialforschung. 6. Aufl. München: Oldenbourg.

Shaffer, J. P. (ed.) 1992. The Role of Models in Nonexperimental Social Science: Two Debates. Washington: American Educational Research Association.

Siegel, S. 1956. Nonparametric Statistics for the Behavioral Sciences. London: McGraw-Hill.

Sigwart, C. 1911. Logik. Band 2: Methodenlehre. Tübingen: Mohr.

Sirkin, R. M. 1995. Statistics for the Social Sciences. Thousand Oaks: Sage.

Siskin, B., Staller, J., Rorvik, D. 1989. What are the Chances? Risks, Odds, and Likelihood in Everyday Life. New York: Crown Publ.

Smith, T. M. F. 1976. The Foundations of Survey Sampling: A Review (with Discussion). Journal of the Royal Statistical Society A 139, 183–204.

Stouffer, S. A. 1934. Sociology and Sampling. In: L. L. Bernard, ed., The Fields and Methods of Sociology, 476–488. New York: Ray Long & Richard Smith.

Suppes, P. 1970. A Probabilistic Theory of Causality. Amsterdam: North-Holland.

Tippett, L. H. C. 1937. The Methods of Statistics. An Introduction Mainly for Experimentalists. 2nd ed. London: Williams and Norgate.

Tschuprow, A. A. 1905. Die Aufgaben der Theorie der Statistik. Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft im Deutschen Reich 29, 11–70.

Tschuprow, A. A. 1906. Statistik als Wissenschaft. Archiv für Sozialwissenschaft und Sozialpolitik 23, 647–711.

Uspensky, J. V. 1937. Introduction to Mathematical Probability. New York: McGraw-Hill.

Wagenführ, R. 1970. Wirtschafts- und Sozialstatistik, Band 1. Freiburg: Rudolf Haufe.

Wagner, A. 1864. Die Gesetzmässigkeit in den scheinbar willkurlichen menschlichen Handlungen vom Standpunkte der Statistik. Hamburg: Boyes & Geisler.

Weber, M. 1921. Wirtschaft und Gesellschaft. 1. Halbband, 5. Aufl., hrsg. von J. Winckelmann. Tübingen: Mohr 1976.

Westergaard, H., Nybølle, H. C. 1928. Grundzüge der Theorie der Statistik. Jena: Fischer.

Whitaker, L. 1914. On the Poisson Law of Small Numbers. Biometrika 10, 36–71.

Whittle, P. 1992. Probability via Expectation. New York: Springer-Verlag.

Winkler, W. 1925. Statistik. Leipzig: Quelle & Meyer.

Winkler, W. 1929. Die statistische Wesensform. Allgemeines Statistisches Archiv 18, 420–425.

Winkler, W. 1931. Grundriss der Statistik I: Theoretische Statistik. Enzyklopädie der Rechts- und Staatswissenschaft, Bd. 46. Berlin: Springer-Verlag.

Wolff, H. 1926. Theoretische Statistik. Grundrisse zum Studium der Nationalökonomie, Bd. 20. Jena: Fischer.

Wolff, H. 1932. Die Fiktion in der Statistik. In: A. Seidel (Hg.), Die Philosophie des Als Ob und das Leben, 123–139. Berlin: Reuther & Reichard. Neudruck: Aalen: Scientia Verlag 1986.

Wundt, W. 1908. Logik. Bd. 3: Logik der Geisteswissenschaften. Stuttgart: Enke.

Yule, G. U. 1912. An Introduction to the Theory of Statistics. 2nd ed. London: Griffin.

Yule, G. U., Kendall, M. G. 1950. An Introduction to the Theory of Statistics. 14th ed. London: Griffin.

Zizek, F. 1908. Die statistischen Mittelwerte. Leipzig: Duncker & Humblot.

Zizek, F. 1921. Grundriß der Statistik. München: Duncker & Humblot.

Zizek, F. 1928. Gleichartigkeit, Homogenität und Gleichwertigkeit in der Statistik. Allgemeines Statistisches Archiv 18, 393–420.

Zizek, F. 1928a. Ursachenbegriffe und Ursachenforschung in der Statistik. Allgemeines statistisches Archiv 17, 380–432.

Zizek, F. 1930. Der Begriff der „Gleichartigkeit“ in der Statistik. Allgemeines statistisches Archiv 20, 8–23.

Namenverzeichnis