MÖGLICHKEITEN UND GRENZEN DER TABELLENANALYSE

Inhalt:

- I Einleitung
- II Wahrnehmungs- und kognitionspsychologische Aspekte der Tabellierung
- III Ökonomische Aspekte der Tabellierung
 - IV Analyseverfahren
 - 1) Grundstrukturen von Beziehungen zwischen Variablen
 - 2) Flow-charts zur Unterscheidung verschiedener Beschiedungen
 - 3) Logische und empirische Probleme der Kausalität
 - 4) Die logische Struktur der Korrelation
 - 5) Die logische Struktur konditionaler Beziehungen
 - 6) Tabellenanalyse und Induktionsproblematik
 - V Literatur

I EINLEITUNG

Um dem höchst anspruchsvollen Titel des Aufsatzes auch nur einigermaßen zu genügen, soll hier auf recht abstraktem Niveau die logische Struktur der Tabellenanalyse dargestellt werden, da sich so m.E. die Möglichkeiten und Grenzen dieser Methode gehereller aber auch präziser aufzeigen lassen als etwa durch eine Sammlung von Anwendungsbeispielen.

Unter Tabellenanalyse versteht man bestimmte Analyseverfahren, die auf Daten angewendet werden, die in Tabellenform geordnet sind. Selbstverständlich ließen sich die Analyseverfahren auch auf Daten anwenden, die nicht in tabellierter, sondern in einer anderen Repräsentation vorliegen, die Tabellenform und die Analyseverfahren sind also grundsätzlich voneinander unabhängig und sollen daher auch getrennt besprochen werden; dabei sei aber die Tabellierung von Daten nur kurz behandelt, da sie ja primär von kognitionspsychologischem und ökonomischen, nicht aber statistischem Interesseist, im eigentlichen soll es um die Untersuchung der Analyseverfahren gehen.

Schon seit frühester Zeit ist in der Psychologie das Phänomen der 'Enge des Bewußtseins' bekannt,d.h. der Tatsache,daß in jeder beliebigen Zeiteinheit nur eine begrenzte Menge von Information und in einer hinreichend kleinen Zeiteinheit nur eine einzige Information bewußt sein kann. Von der Informationspsychologie sind diese Angaben inzwischen präzisiert. (1,245-248) Von den ca. 10⁷ bit Infor-

mation, die den optischen Kanal pro sec 'reizen', werden etwa 16 bit/sec bewußt. Die Zeitdauer, welche zur Vergegenwärtigung von 1 bit Information benötigt wird, das subjektive Zeitquant (SZQ), beträgt 1/16 sec. 1 SZQ ist auch der Mindestabstand, den zwei Ereignisse haben müssen, damit sie noch sicher als aufeinanderfolgend wahrgenommen werden.

Inwiefern sind diese Ergebnisse nun relevant für die kognitionspsychologische Beurteilung der Tabellierung von Daten? Auf Grund
der hier aufgezeigten beschränkten Kapazität der Informationsaufnahme und -verarbeitung des kognitiven Systems des Menschen,ist
folglich die Struktur des Zeichensystems,in dem die Information codiert ist,von entscheidender Bedeutung dafür,wieviel erwünschter
Information innerhalb einer Zeiteinheit aufgenommen werden kann,
denn dieses Maß hängt davon ab,inwieweit das entsprechende Zeichensystem den wahrnehmungspsychologischen Gesetzen unseres kognitiven
Systems adäquat ist. Und die Tabelle, die ja nichts anderes als ein
spezifisches Zeichensystem ist,erweist sich hier als vorteilhaft.
Zur Verdeutlichung seien die gleichen Beziehungen zwischen drei Variablen einmal durch eine Tabelle und einmal in rein linearer Schreibweise dargestellt.

A)	Tabe	7 7	1 6
A	Tabe	1_	Lt

		Х Х		X		
		+	404	+	-	
Y	+	a	ъ	е	f	R ₁
	-	С	d	g	h	R ₂
		S ₁	S	S3	SA	n

B) Lineare Darstellung

t:
$$S_1 + S_2 + S_3 + S_4 = R_1 + R_2 = n$$

 $S_1 = a + c$
 $S_2 = b + d$
 $S_3 = e + g$
 $S_4 = f + h$
 $R_1 = a + b + e + f$
 $R_2 = c + d + g + h$
t+: $S_1 + S_2$
t-: $S_3 + S_4$
X: n
X+: $S_1 + S_3$
X-: $S_2 + S_4$
Y: n

Y+: R1 Y-: R2 t + / X + / Y + : at + /X - /Y + : bt + / X + / Y - : ct + /X - /Y - : dt-/X+/Y+: et-/X-/Y+: ft = /X + /Y - : gt - / X - / Y - : ht+/X+: S1 t+/X- : S2 t-/X+ : S₃ t-/X- 8 S4 X + / Y + : a + eX+/Y- : c + gX - / Y + : b + fX-/Y-:d+h

(Man könnte diese Liste noch erweitern, aber die wesentlichen Informationen sind wohl aufgeführt.)

Obwohl diese beiden Zeichensysteme ja in Symbolreservoir und syntaktischen Regeln weitgehend übereinstimmen und sich primär nur in der graphischen Repräsentation unterscheiden, fällt sofort auf, daß die Daten in der 'dreidimensionalen' Tabelle im Gegensatz zur 'eindimensionalen' Aneinanderreihung in einer Weise geordnet sind, die man als übersichtlich, anschaulich, überschaubar etc. bezeichnen würde die Tabelle ermöglicht es, wesentlich mehr relevante Daten bzw. Beziehungen zwischen Daten äufzunehmen als bei der linearen Darstellung. Dies hängt wahrscheinlich mit der Invariantenbildung bei der Wahrnehmung von Superzeichen zusammen, auf diese Fragen soll hier aber nicht weiter eingegangen werden. Jedenfalls ist dies nicht allein ein zeitökonomischer Vorteil der Tabellenanalyse, sondern durch die dichtere, konzentriertere Verarbeitung von Information wird eine tiefere, umfassendere Erkenntnis der strukturellen Abhängigkeiten der Daten erreicht, ohne das dies hier präzisiert werden soll.

III ÖKONOMISCHE ASPEKTE

Selbstverständlich hat die Tabellierung von Daten ökonomische Vorteile; man spart Zeit, Arbeitskraft, Schreibmaterialien und somit natürlich auch Geld. Dies ist zwar nicht von wissenschaftlicher oder wissenschaftstheoretischer, aber erheblicher pragmatischer Relevanz.

IV ANALYSEVERFAHREN

1) Grundstrukturen von Beziehungen zwischen Variablen

Bei den Analyseverfahren geht es darum, die Beziehungen, die zwischen Variablen bestehen, aufzudecken. (In diesem Aufsatz sollen nur Beziehungen zwischen maximal drei Variablen berücksichtigt werden, die Ergebnisse ließen sich aber mit entsprechenden Modifikationen auch auf Beziehungen zwischen mehr Variablen übertragen.)

Grundsätzlich sollen die Beziehungen hinsichtlich zweier Faktoren unterschieden werden:

- A) das Meßniveau der Variablen
- a) nominal
- b) ordinal
- c) metrisch
- B) die Allgemeinheit der Abhängigkeit
- a) deterministisch
- b) statistisch (= wahrscheinlich)

Ist eine Beziehung deterministisch, dann haben alle Objekte, die bezüglich einer Variablen (bzw. auf einer Merkmalsdimension) den gleichen Wert (bzw. die gleiche Merkmalsausprägung) haben, dies auch bezüglich der anderen Variablen, bei statistischen Beziehungen gilt dies nur für einen gewissen Prozentsatz, d.h. für ein einzelnes Objekt nur mit einer gewissen Wahrscheinlichkeit. Deterministische Beziehungen werden durch sog. Allaussagen, statistische durch statistische Aussagen beschrieben. Streng genommen besteht allerdings keine wirkliche Grenze zwischen deterministischen und statistischen Aussagen (bzw. Beziehungen); deterministische Aussagen sind Grenzfälle statistischer Aussagen, sie sind entweder wahr, d.h. sie haben eine Wahrscheinlichkeit von p = 1, oder sie sind falsch, dann haben sie eine Wahrscheinlichkeit von p = 0, dazwischen liegen die statistischen Aussagen.
Berücksichtigt man diese beiden Faktoren, so ergeben sich sechs grund-

Allgemeinheit der Beziehung

M e			deterministisch	5	statistisch
B	nominal	٨	$z (X_{(z)} \rightarrow Y_{(z)})$	р($X_{(z)} \rightarrow Y_{(z)} = p$
i	ordinal	٨	$z \wedge w (X_{(z,w)} \rightarrow Y_{(z,w)})$	p ($(z,w) \rightarrow Y_{(z,w)} = \mathbf{r}$
e a	metrisch	٨	$z (mX_{(z)} \rightarrow nY_{(z)})$		$mX(z) \rightarrow nY(z) = p$

Zur Erläuterung sei für jeden der deterministischen Fälle ein (rein willkürliches) Beispiel gebracht.

- a) Für alle Menschen gilt: Wenn sie Schmerzen haben, sind sie unglücklich.
- b) Für alle Amerikaner gilt: Wenn sie ein höheres Einkommen haben als andere Amerikaner, dann haben sie auch einen höheren Status als diese anderen.
- c) Für alle Deutschen gilt: Wenn sie 50 Jahre alt sind, haben sie eine Lebenserwartung von 25 Jahren.

Folgende Bemerkungen sind hier notwendig:

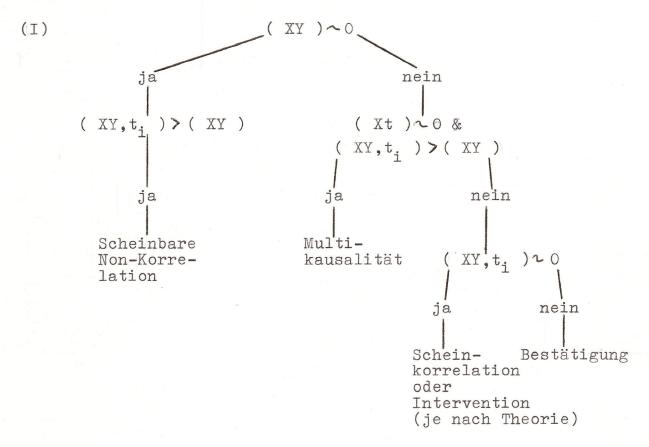
- A) Es soll in diesem Aufsatz immer wieder auf die formale Logik (Aussagen-, Prädikaten- und Quantorenlogik) zurückgegriffen werden, um ein möglichst hohes Maß an Klarheit und Präzision zu erreichen.
- B) Die hier verwendeten Symbole sind z.T. in der Logik unüblich; für z $(X_z \to Y_z)$ schreibt man gewöhnlich: x $(F_x \to G_x)$.Um aber eine Anpassung vorzunehmen an die bei der Tabellenanalyse übliche Bezeichnung der beiden Grundvariablen mit X und Y, wurden die Symbole ausgewechselt.
- C) Die statistische Aussage $p(X_z \rightarrow Y_z)^{\frac{1}{2}}$ schreibt man gewöhnlich p(X,Y) = r; die erstgenannte Schreibweise ist aber m.E. adäquater.
- D) Für den positiven Wert einer Variablen: X + schreibt man in der Logik nur das Symbol: X,für den negativen Wert: X die Verneinung: ¬ X.Will man allgemein die Variable X kennzeichnen, so schreibt man: X v ¬ X.
- E) Im Rahmen dieses Aufsatzes soll primär vom einfachsten und grundlegensten Fall, nämlich einer deterministischen Beziehung zwischen 2 oder 3 nominal gemessenen, dichotomischen Variablen ausgegangen werden, die Resultate sind aber prinzipiell auf die anderen Fälle übertragbar.
- 2) Flow-charts zur Unterscheidung verschiedener Beziehungen
 Bei MAYNTZ et al. (6/192) werden drei grundsätzliche 'kausale' Beziehungen zwischen 2 oder 3 Variablen angegeben:
- A) der monokausale Zusammenhang
- B) der multikausale Zusammenhang
- C) die Kausalkette

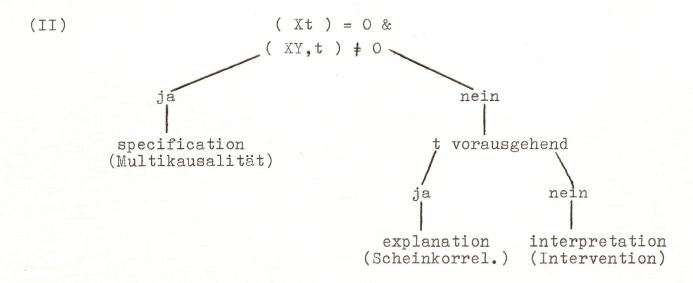
Außerdem unterscheidet man (6/199-209) fünf generelle Möglichkeiten, die sich ergeben können, wenn man in eine 2-dimensionale Tabelle eine 3. Variable, die sog. Testvariable t, einführt:

1 Anmerk.: In diesem Aufsatz wird bewußt sehr oft Bezug genommen auf das Buch von MAYNTZ et al. - Einführung in die Methoden der empirischen Soziologie,Opladen 1972,3.A.,da es wohl das verbreiteste und modernste Lehrbuch der empirischen Sozialforschung (in der BRD) ist.

- a) Bestätigung (des monokausalen Zusammenhangs)
- b) Multikausalität
- c) Scheinkorrelation
- d) Intervention (Kausalkette)
- e) Scheinbare Non-Korrelation

Rein statistisch gesehen unterscheiden sich diese fünf Möglichkeiten durch die unterschiedlichen Korrelationen zwischen den 3 Variablen.Man kann diese Unterschiede z.B. in einem Flow-chart zur Bestimmung der vorliegenden Struktur darstellen. Anschließend seien zwei solcher Flow-charts gebracht: das erstist zusammengestellt nach Angaben von MAYNTZ et al. (6/199-209),das zweite nach Angaben von LAZARSFELD (5/122).





3) Logische und empirische Probleme der Kausalität

Die Unbefangenheit.mit der etwa bei MAYNTZ et al. (6) der Kausalitätsbegriff bei der Tabellenanalyse verwendet wird, mutet wissensschaftstheoretisch gesehen recht problematisch an. In Wirklichkeit zeigt sich nämlich gerade hier sehr deutlich eine wesentliche Grenze der Tabellenanalyse; was diese tatsächlich leisten kann, ist primär die Aufdeckung von Korrelationen zwischen Variablen. Unter welchen Bedingungen Korrelationen aber kausale Beziehungen zugrunde gelegt werden können, ist eine äußerst schwierige, und noch keineswegs hinreichend geklärte Problematik, allein schon deswegen, weil der Kausalitätsbegriff selbst unpräzise und inkonsistent ist. Deshalb beschränkt man sich in den Naturwissenschaften ja auch nur von fünktionalen Abhängigkeiten, nicht aber Kausalbeziehungen zu reden. Dahinter steht die grundsätzliche Frage, inwieweit es in der Wirklichkeit überhaupt so etwas wie Kausalität gibt, ob Kausalität nicht nur eine Struktureigenschaft unseres Denkens (bzw. unseres neurophysiologischen Systems) ist, die wir auf die Wirklichkeit projektieren. Dafür spricht z.B., daß der Kausalitätsbegriff trotz seiner Schwierigkeiten offensichtlich wissenschaftlich unverzichtbar bleibt; es ist aber notwendig, dann jeweils so genau wie möglich zu definieren, was man darunter versteht. Selbstverständlich kann hier nicht ausführlich auf die höchst umfangreiche Kausalitätsproblematik eingegangen werden, es sollen nur einige, für die Tabellenanalyse relevante Probleme, aufgezeigt werden. Hierbei sollen unter Kausalbeziehungen -wie dies auch wissenschaftstheoretisch üblich ist- nur deterministische Beziehungen verstanden werden, von stochastischem Kausalzusammenhang zu reden (wie z.B. in (6/193)) ist im Grunde ein Widerspruch in sich. Außerdem soll der Einfachheit halber immer von dichotomen Variablen ausgegangen werden. Wie schon gesagt, unterscheidet man in der statistischen Literatur streng zwischen Korrelation und Kausalität. Unter positiver Korrelation wird verstanden (bei dichotomen Variablen), daß X+ und Y+ bzw. X- und Y- zusammen auftreten, unter negativer Korrelation entsprechend das gemeinsame Auftreten von X+ und Y- bzw. X- und Y+. Wie sich das logisch präzisieren läßt, was man normalsprachlich so leichthin 'gemeinsam auftreten' nennt, darauf soll später eingegangen werden; hier soll es primär um Probleme der Kausalität gehen. Bei MANTZ et al. (6/193) wird folgende Kausalitätsdefinition gegeben: X ist Ursache von Y, wenn gilt: $(X \rightarrow Y) & (\neg X \rightarrow \neg Y)$; dies ist logisch äquivalent $(X \rightarrow Y)$ & $(Y \rightarrow X)$, dies ist aber genau die Definition der logischen Äquivalenz: X \ Y.Hierbei ist folgendes zu be-

rücksichtigen: X -> Y bedeutet ja: X ist hinreichende Bedingung für Y

und Y ist notwendige Bedingung für X. X + Y bedeutet dann entsprechend: X ist hinreichende und notwendige Bedingung für Y. und Y ist hinreichende und notwendige Bedingung für X.Dies kann aber wohl kaum eine befriedigende Definition von Kausalität sein; Wenn man z.B. sagt: 'Person a starb, weil (=Kausalkonjunktion) sie ertrank' oder allgemeiner: 'Wenn ein Mensch ertrinkt, dann sirbt er', wird man gemeinhin akzeptieren, daß das Ertrinken das Sterben kausal bewirkt. Nach obiger Definition von MAYNTZ et al. läge hier aber keine Kausalität vor, denn es gilt ja nicht, daß wenn ein Mensch nicht ertrinkt, daß er dann auch nicht stirbt, d.h. das Ertrinken ist nur eine hinreichende, aber keine notwendige Bedingung des Sterbens, entsprechend ist das Sterben nur eine notwendige, aber keine hinreichende Bedingung des Ertrinkens. Das Problem liegt hier also darin, daß eine Wirkung verschiedene Ursachen haben kann; hat dagegen eine 'Ursache' verschiedene, alternativ auftretende Wirkungen, so liegt keine Kausalität yda es sich ja dann nicht mehr um eine deterministische Beziehung handelt. Aber selbst wenn X 😝 Y wahr ist, d.h. Y dann und nur dann auftritt, wenn X auftritt, und umgekehrt, bedeutet dies dennoch nicht, daß Y auftritt, w e i l X auftritt. Es ist auch nicht zufällig, daß es in der Logik keinen Junktor für Kausalität gibt;die meisten Logiker sind der Auffassung, daß sich Kausalität nicht formalisieren läßt. Wichtig ist hier allerdings die Unterscheidung zwischen der Ursache eines Sachverhalts und seiner kausalen Erklärung, die sich beide auf die Frage 'warum' als Antwort geben lassen. Nimmt man z.B. den banalen Sachverhalt: 'zn ist sterblich.' Als kausale Erklärung für diesen Sachverhalt würde man in der Wissenschaftstheorie, gemäß dem HO-Schema, angeben:

	enschen sind ein Mensch	sterblich	formal:	A 2	$z (X_z \rightarrow X_{z_n})$	Y _z)
z _n ist	sterblich				_	Yzn

Man reduziert hier also im Grunde Kausalität auf logische Ableitbarkeit. Man würde aber dennoch nicht sagen, daß die Konjunktion von Gesetzeshypothese Λ z (X $_z\to Y_z$) und empirischer Randbedingung X $_z$ die Ursache des singulären Sachwerhaltes Y $_z$ ist. Sondern man würde sich hier auf eine biologische Alternstheorie berufen, sondern man könnte etwa anführen, daß die Ribosomen von z $_n$ nach einer bestimmten Zeit anfangen werden, Nonsense-Protein zu produzieren, dadurch fehlerhafte Enzyme gebildet werden, wodurch die Homöostasie des Stoffwechsels progressiv gestört wird, bis schließlich die Gehirnwellenaktivität erliegt. Hier stellt sich das Problem der Kausalkette, man kann ja immer

weiter 'warum' fragen, nach der Ursache jeder Ursache fragen. Reicht es z.B. aus zu sagen, die Ursache dafür, daß zn aggressiv ist, ist daß zn frustriert wurde, oder muß man weiter angeben, warum zn frustriert wurde usw., wobei die Menge der aufzuzählenden Sachverhalte schnell ins Ungeheure anwachsen würde und vorallem auch bald Fragen auftreten würden, die sich wissenschaftlich gar nicht beantworten lassen, da sie metaphysischer Natur sind (z.B. Warum gibt es in der Natur das Prinzip der Arterhaltung?).

Immerhin muß man sich die Frage stellen, ob nicht jeder Sachverhalt letztlich in einem Beziehungsgeflecht mit unendlich vielen anderen Sachverhalten steht; dabei braucht es sich nicht nur um kausal lineare Beziehungen zu handeln, wahrscheinlicher ist vielmehr, daß rückkoppelnde Regelkreisabhängigkeiten vorliegen.

Natürlich ist auch die 'Natur' des Sachverhalts von Relevanz: handelt es sich um einen logisch-analytischen oder um einen empirischsynthetischen Sachverhalt, und besteht der empirische Sachverhalt in einem zeitlichen Prozeß oder in einem 'zeitlosen' Zustand. Man wird z.B. den Sachverhalt ' z_n ist Deutscher', bei dem es sich ganz einfach um die mengentheoretische (oder umfangslogische) Beziehung handelt, daß z_n Element der Menge aller Deutschen ist, anders kausal begründen müssen als etwa den Sachverhalt '1804 wurde Napoleon zum Kaiser gekrönt', der sich nur inhaltslogisch adäquat erfassen läßt.

Ein weiteres Problem betreffend die Kausalität kommt von einer ganz anderen Seite. Bekanntlich lassen sich in der Mikrophysik nur statistische Gesetzmäßigkeiten angeben, und die überwiegende Auffassung ist heute, daß hier die Natur tatsächlich nur statistische Ordnung aufweist und sich nicht hinter den statistischen noch deterministische Gesetzmäßigkeiten verbergen. Die 'Unschärfe' elementarer Prozesse wird allerdings durch die große Zahl, mit der sie makrophysikalisch in Erscheinung treten, soweit ausgeglichen, daß für makrophysikalische Vorgänge daraus annähernd deterministische Gesetzmäßigkeiten resultieren. Aber auch in makrophysikalischen Systemen können noch indeterminierte, quantenhafte Prozesse auftreten. Dies ist von großer Relevanz gerade für die Kausalanalyse sozialer Phänomene; denn als naturwissenschaftlich denkender Soziologe kann man nicht anstreiten, daß sich soziales Verhalten höchstwahrscheinlich letztlich nur durch die ihm jeweils entsprechenden bzw. mit ikm identischen (psychophysisches Problem) Abläufe im neuro-endokrinen System erklären lassen; zwar muß man das NES als ein makrophysikalisches System ansehen, da die Erregungsleitung wesentlich nicht in der Zustandsänderung eines oder weniger Atome.sondern in den Strömungsverhältnissen ganzer Kolektive von Tausenden von

Na- und K-Ionen und anderen ähnlichen Diffusionsvorgängen besteht. Aber von verschiedenen Forschern (z.B. P.JORDAN) wird dennoch der Einfluß quantenhafter Prozesse postuliert. Außerdem zeigen Ergebnisse der Neurokybernetik, daß die Prozesse im ZNS stests mit einer gewissen 'Streuung' behaftet sind und insofern nur als 'teildeterminiert' angesehen werden können. ((2/48-51) Man muß also jedenfalls mit der Möglichkeit rechnen, daß soziale Verhaltensweisen prinzipiell nicht streng kausal-deterministisch erklärt werden können. Abschließend hierzu sei festgehalten: Was Kausalität genau ist,inwieweit sie überhaupt real existent ist usw., dies sind höchstkomplexe und keineswegs gelöste Probleme. Die berechtigte Forderung nach Vollständigkeit einer kausalen Erklärung könnte dazu führen, daß zur Erklärung jedes singulären Sachverhalts eine völlig unvorstellbare. alle Sachverhalte und alle Beziehungen zwischen Sachverhalten umfassende kosmische Theorie notwendig wäre. Zumindestens ist es jedenfalls unhaltbar, ein so komplexes Geschehen wie menschliches Verhalten durch monale Beziehungen (z.B. Monokausalität) k a u s a l erklären zu wollen. Wenn menschliches Verhalten überhaupt kausal erklärbar ist, dann letzlich wohl nur durch eine physikalisch-kybernetische Theorie,in die alle relevanten soziologischen, psychologischen, biologischen, neurophysiologischen ... Informationen eingegangen sind. Hier liegt m.E. die wesentliche Grenze der Tabellenanalyse: sie kann eben nicht kausale oder regelkreisartige Beziehungen aufdecken, sondern nur Korrelationen, inwieweit es berechtigt ist, aus bestimmten Korrelationskombinationen darüberhinausgehende Beziehungen abzuleiten. soll im folgenden erörtert werden.

4) Die logische Struktur der Korrelation

Auch wenn die Auffassung von der Tabellenanalyse als einer Methode der Kausalanalyse nicht haltbar ist,könnte man es doch für möglich halten,mit ihr konditionale Beziehungen aufzudecken,in verschiedenen Varianten wie 'Monokonditionalität', 'Multikonditionalität, 'interventionale Konditionalität' etc. Dazu ist es aber erst einmal erforderlich,genau anzugeben,was unter Korrelation und unter Konditionalität zu verstehen ist und wie sich die beiden Begriffe gegeneinander abgrenzen lassen. Dazu soll versucht werden, die logische Struktur von Korrelation und Konditionalität aufzuzeigen. Dies erweist sich als keineswegs trivial, sondern im Gegenteil als so kompliziert, daß hier nur Vorschläge gemacht werden können, die keineswegs Anspruch erheben, hinreichend adäquat zu sein.

Dabei soll wieder vom Modell einer deterministischen Beziehung zwischen zwei dichotomen Variablen ausgegangen werden. Man könnte kri-

tisieren, daß dies nicht typisch für die wissenschaftliche Praxis ist, schon gar nicht für die sozialwissenschaftliche. Die Logik der Tabellenanalyse ist aber grundsätzlich unabhängig vom Meßniveau der Variablen und der Allgemeinheit der Beziehung; mit entsprechenden Modifikationen, auf die hier allerdings hur kurz eingegangen werden kann, sind die Ergebnisse generalisierbar. Unter einer deterministischen Beziehung in einer Tabelle soll nur verstanden werden, daß alle untersuchten Objekte (Stichprobe) die Beziehung realisieren, nicht aber notwendig alle Objekte der betreffenden Klasse (Grundgesamtheit) überhaupt; das Problem, ob und wie sich die an der Stichprobe gewonnenen Tabellenwerte verallgemeinern lassen, soll am Schluß des Aufsatzes besprochen werden.

Unter Korrelation zweier Variablen X und Y versteht man, daß X und Y voneinander abhängig sind, präziser: unter positiver Korrelation, daß hohe Werte von Werte von X mit hohen Werten von Y und niedrige Werte von X mit niedrigen Werten von Y einhergehen, unter negativer Korrelation, daß hohe X-Werte mit niedrigen Y-Werten und niedrige X-Werte mit hohen Y-Werten einhergehen. Bei deterministischen Beziehungen wischen dichotomen Variablen hat die positive Korrelation natürlich immer den Wert $\mathbf{k}_p=1$ und die negative $\mathbf{k}_n=-1$, denn es gilt dann eben: entweder geht die Wahrheit von X mit der Wahrheit von Y einher oder die Wahrheit von X mit der Falschheit von Y.

Diese normalsprachlichen Formulierungen könnte man formal logisch folgendermaßen ausdrücken:

$$(1) (X \rightarrow Y) & (\neg X \rightarrow \neg Y) \dot{v} (X \rightarrow \neg Y) & (\neg X \rightarrow Y)$$

Dieser Ausdruck ist logisch äquivalent mit

(2)
$$(\neg Y \rightarrow \neg X) & (Y \rightarrow X) \dot{v} (\neg Y \rightarrow X) & (Y \rightarrow \neg X)$$

Es wird also ausgedrückt, daß Korrelation eine wechselseitige Beziehung ist, daß nicht Y mit X korreliert, sondern X und Y zusammen korrelieren. Am besten schreibt man kurz:

(3)
$$(X \leftrightarrow Y) \dot{v} (X \leftrightarrow \neg Y)$$

Die beiden Teilausdrücke werden durch das ausschließende oder : \dot{v} verbunden, da positive Korrelation ja negative Korrelation ausschließt und umgekehrt.

Das Problem, das sich nun ergibt, ist folgendes: der Ausdruck (3) ((1) und (2) natürlich auch) ist eine Tautologie, er ist also immer wahr, unabhängig davon, welche Wahrheitswerte X und Y haben, d.h. definiert man 'X und Y korrelieren zusammen' durch diesen Ausdruck, gilt also: 'X und Y korrelieren zusammen' Tautologie. Diese Äquivalenzdefinition ist aber nur dann wahr, wenn beide Seiten wahr oder beide Seiten falsch sind, da die Tautologie immer wahr ist, muß also auch

'X und Y korrelieren zusammen' immer wahr sein, es kann also nie der Fall eintreten, daß keine Korrelation zwischen X und Y besteht. Folgende Interpretationen dieses Ergebnisses sind möglich:

A) Der Ausdruck (3) ist adäquat, denn bei deterministischen Beziehungen zwischen dichotomen Variablen kann die Korrelation nicht den Wert k = o annehmen.

- B) Der Ausdruck ist falsch, denn die Korrelation muß in jedem Fall auch o werden können; es muß ein anderer logischer Ausdruck verwendet werden.
- C) Der Ausdruck ist falsch,aber er kann nicht durch einen anderen aussagenlogischen Ausdruck adäquat ersetzt werden,denn die Dreiheit von positiver,keiner und negativer Korrelation ist innerhalb einer 2-wertigen Logik nicht ausdrückbar,sondern man benötigte dazu eine 3-wertige Logik.

M.E. ist C) höchst unwahrscheinlich; die Wahrheitstafeln der 2-wertigen Aussagenlogik beruhen zwar betreffend ihren Begründungszusammenhang auf Definition, aber die Aussagenlogik hat sich rein pragmatisch gesehen als sehr leistungsfähig zur Beschreibung empirischer und analytischer Sachverhalte erwiesen, durch sie läßt sich die gesamte Mathematik aufbauen, sie liegt der Schaltelektronik der Computer zugrunde, aber sie entspricht auch unserem intuitiven logischen Denken, was dadurch gestützt wird, daß auch die Schaltungen der Neuronen sich aussagenlogisch darstellen lassen; außerdem kann ja die oben genannte Dreiheit reduziert werden auf die Zweiheit: Korrelation – keine Korrelation, und Korrelation auf die Zweiheit: positive Korrelation – negative Korrelation.

Die Entscheidung zwischen A) und B) ist dagegen problematisch; man könnte etwa auch folgenden Ausdruck als logischen Korrelationsbegriff rechtfertigen:

$$(4) (X \rightarrow Y) \rightarrow (\neg X \rightarrow \neg Y) \vee (X \rightarrow \neg Y) \rightarrow (\neg X \rightarrow Y)$$

Hier würde man also, betr. die positive Korrelation, aussagen, daß die Tatsache, daß X Y impliziert, eine hinreichende Bedingung dafür ist, daß $\neg X$ $\neg Y$ impliziert; die Formel (4) ist aber nicht tautologisch, die gerade aufgezeigte Problematik würde also nicht auftreten. Allerdings müßte man eigentlich fordern, daß wenn $(X \rightarrow Y) \rightarrow (\neg X \rightarrow \neg Y)$ gilt, daß dann auch $(\neg X \rightarrow \neg Y) \rightarrow (X \rightarrow Y)$ gilt, als Gesamtausdruck ergäbe sich dann:

(5)
$$(X \rightarrow Y) \leftrightarrow (\neg X \rightarrow \neg Y) v (X \rightarrow \neg Y) \leftrightarrow (\neg X \rightarrow Y)$$

Dies ist aber wieder eine Tautologie.

Umgekehrt könnte man fragen, ob sich ein brauchbarer Begriff für Nicht-Korrelation logisch konstruieren läßt.ist (3) richtig, dann muß eine Formel für Nicht-Korrelation logisch widersprüchlich sein.d.h.

immer falsch, unabhängig vom Wahrheitswert von X und Y, wie z.B. der adäquat erscheinende Ausdruck:

(6) $(X \rightarrow Y \& \neg Y) \& (\neg X \rightarrow Y \& \neg Y)$

Das gleiche gilt für die Negation von (3):

 $(7) \neg ((X \Leftrightarrow Y) \dot{V} (X \Leftrightarrow \neg Y))$

Damit spricht sehr vieles dafür, daß A) richtig ist, d.h. daß die Korrelation zwischen dichotomen Variablen bei deterministischen Beziehungen nicht o werden kann.

Dafür gibt es auch noch andere Argumente: Eine deterministische Beziehung ist ja gerade dadurch definiert, daß für alle Elemente der betr. Klasse dasselbe gilt, d.h. daß entweder alle z, die einen positiven X-Wert haben, auch einen positiven Y-Wert haben, der alle z mit positivem X-Wert einen negativen X-Wert haben (k=-1), nur wenn für eine Teilmenge der Menge aller z mit X+ Y+ gilt, für die Ergänzungsmenge aber Y-, ergeben sich Korrelationen zwischen -1 und +1, und zwar ist die Korrelation genau dann o, wenn die beiden Teilmengen gleich groß sind; dies ist aber der Fall, wenn gilt: $p(X_Z \rightarrow Y_Z) = 0.5$, also $X \rightarrow Y$ eine Wahrscheinlichkeit von p=0.5 hat, und damit per definitionem (s. IV,1) keine deterministische, sondern eine statistische Aussage ist.

Wie beurteilt man dann aber eine Aussage wie: Λ z ($X_Z \to Y_Z$ v $-Y_Z$)? Formal ist dies eine deterministische Aussage, und zwar eine tautologische, sie ist also wahr, unabhängig vom Wahrheitswert von Y (und X), also sowohl wenn Y_Z wahr ist als auch wenn $\neg Y_Z$ wahr ist; es kann z.B. gelten: für 60% aller z, für die X_Z wahr ist, ist Y_Z wahr, für 40% $\neg Y_Z$, oder für 20% Y_Z und für 80% $\neg Y_Z$. Jede denkbare Verteilung ist möglich, dabei treten natürlich Korrelationen zwischen k=1 und k=-1 auf, und bei einem Verhaltnis von 50%: 50% eine Korrelation von o.

Dies widerspricht natürlich den gerade oben gezogenen Folgerungen. Welche Bedeutung ist diesem Ergebnis nunzuzumessen?

Dazu sei noch einmal kurz die Funktion aller hier (in Punkt 4) bisher angestellten Überlegungen für das Thema des Aufsatzes rekapituliert: Um aufzuzeigen, welche Möglichkeiten die Tabellenanalyse als Korrelationsanalyse bietet und inwiefern sie darüberhinaus gestattet, andere Beziehungen aufzudecken, ist es bei korrekter Argumentationsweise erforderlich, erst einmal den unpräzisen, normalsprachlichen Korrelationsbegriff logisch zu formalisieren und damit zu präzisieren. Es soll also nicht versucht werden, einen neuen Korrelationsbegriff einzuführen, sondern die impliziten semantischen Regeln des vorliegenden Korrelationsbegriffes explizit zu machen. Die aufgetretenen Schwierigkeiten machen deutlich, daß dies keineswegs eine triviale, sondern eine recht komplizierte Aufgabe ist. So ist es bis hierher nicht einmal gelungen, die fundamentale Frage eindeutig zu entscheiden, ob bei dem Grundmodell

einer Beziehung, nämlich einer deterministischen Beziehung zwischen zwei dichotomen Variablen ein tautologischer oder ein nicht tautologischer Ausdruck dem normalsprachlichen Korrelationsbegriff adäquat ist.

Es gibt aber noch eine andere Möglichkeit zu überprüfen,welche logische Eigenschaften der Korrelationsbegriff besitzt: denn in den mathematischer Korrelationskoeffizienten liegen ja präzise Korrelationsausdrücke vor, deren Brauchbarkeit erwiesen ist,wenn es auch fragwürdig ist,inwieweit sie exakt dem normalsprachlichen Korrelationsbegriff entsprechen. Es soll also versucht werden,den logischen Korrelationsausdruck (3) $(X \leftrightarrow Y)$ v $(X \leftrightarrow \neg Y)$ mit Hilfe eines Korrelationskoeffizienten für nomial skalierte Variablen,und zwar $Q = \frac{\text{ad} - \text{bc}}{\text{ad} + \text{bc}}$, zu überprüfen,indem man die Wahrheitswerte aus der Wahrheitstafel von (3) in Tabellenwerte transformiert und dann mit Q die Korrelation berechnet.

Die Wahrheitswertetafel von (3) lautet:

$$(XX \leftrightarrow Y) \dot{v} (\neg X \leftrightarrow Y)$$

Es ist einleuchtend, daß man zwei Tabellen benötigt, eine für positive Korrelation, Logisch X \(\lefta\) Y, und eine für negative Korrelation, logisch \(\times\) Y. Die Tabelle für positive Korrelation sieht folgendermaßen aus:

	X			
		+		
Y -	+	a=o v > o (a=o → d>o)	b=o	a
	-	C=0	d=o v > o (d=o → a>o)	đ
		a	đ	n

Dies sei kurz erklärt: Vorweg sei daran erinnert, daß folgende Ausdrücke gleichbedeutend sind:

dichotomisches Merkmal X : X v ¬X

positive Merkmalsausprägung X+ : X / negative Merkmalsausprägung X- : ¬X Die Wahrheitstafel von X \longleftrightarrow Y, oder prädikatenlogisch X $_{\rm Z}$ \longleftrightarrow Y $_{\rm Z}$, unterscheidet vier logisch mögliche Fälle: X $_{\rm Z}$ \longleftrightarrow Y $_{\rm Z}$

Zuerst seien der 2. und 3. Fall erklärt, da diese unproblematischer sind: 2./3. Fall: der 2. Fall besagt: wenn z die Eigenschaft X hat, X_z also wahr ist, so ist die Folgerung falsch, daß z nicht die Eigenschaft Y hat, X_z also falsch ist, und umgekehrt; d.h. es kann keine z geben, die auf der

Merkmalsdimension X die Ausprägung + haben,aber auf der Merkmalsdi-Y die Ausprägung -; daher muß die Zelle c der Tabelle den Wert o erhalten. Entsprechend kann es beim 3. Fall keine z geben, für die X_z falsch ist, Y_z aber wahr; folglich muß auch die Zelle b den Wert o haben.

Nun zum 1. und 4. Fall: Der 1. Fall sagt aus, daß wenn z die Eigenschaft X hat, dies hinreichend und notwendig dafür ist, daß z auch die Eigenschaft Y hat, und umgekehrt. Der 4. Fall behauptet entsprechend: \neg X \leftrightarrow \neg Y.

- 4. Fall, so ergeben sich folgende Möglichkeiten:
- A) Alle untersuchten Menschen sind frustriert und aggressiv.
- B) Alle untersuchten Menschen sind nicht frustriert und nicht aggressiv
- C) Ein Teil der untersuchten Menschen ist frustriert und aggressiv, der andere Teil ist nicht frustriert und nicht aggressiv.

Ob eine der ersten beiden oder die Möglichkeit C) zutrifft,ist von entscheidender Bedeutung:

Bei der Möglichkeit A) gilt nämlich: a>o & d=o,

dann folgt: $Q = \frac{a.0 - 0.0}{a.0 + 0.0} = 0$

Bei der Möglichkeit B) gilt: a=o & d>o,

dann folgt: $Q = \frac{0.d - 0.0}{0.d - 0.0} = 0$

d.h. bei Möglichkeit A) und B) besteht keine Korrelation.

Bei Möglichkeit C) gilt dagegen: a> o & d>o,

also $Q = \frac{a \cdot d - 0 \cdot o}{a \cdot d + 0 \cdot o} = 1$, es ergibt sich also eine positive Korrelation von k=1.

Dieses Ergebnis erscheint aber inadäquat. Denn gemäß dem normalsprachlichen Korrelationsbegriff würde man doch zumindestens auch bei der Möglichkeit A) eine positive Korrelation von k=1 postplieren. Folglich ist entweder der Korrelationskoeffizient Q, oder der logische Ausdruck (3). oder beide, nicht dem normalsprachlichen Korrelationsbegriff semantisch äquivalent, oder es liegt ein Störfaktor oder Fehler vor. Wahrscheinlich läßt sich die Diskrepanz folgendermaßen erklären: Der normalsprachliche Begriff für positive Korrelation erfordert zwar, daß $(X \to Y)$ & $(\neg X \to \neg Y)$ wahr ist, er wird aber bereits angewendet, wenn sich $(X \to Y)$ v $(\neg X \to \neg Y)$ nachweisen läßt, da man (unzulässigerweise) $(X \to Y)$ \leftrightarrow $(X \to \neg Y)$ voraussetzt. Damit wäre der normalsprachliche Begriff für positive Korrelation aber äquivok und natürlich nicht durch einen einzigen logischen Ausdruck wiederzugeben. Und die logische Struktur $(X \to Y)$ & $(\neg X \to \neg Y)$ wird eben im Gegen-

satz zu (X \Longrightarrow Y) $\mathring{\mathbf{v}}$ (-X \Longrightarrow -Y) erst dann empirisch voll realisiert, wenn Fälle von X_z u n d -X_z auftreten, nicht aber nur X_z o d e r nur -X_z.

Für die negative Korrelation X +> -Y gilt entsprechendes:

Die Wahrheitstafel lautet: Die dazu gehörige Tabelle:

		X		
		+	-	
¬Y	+	a=0	b=0 v >00 (b=0 -> c > 0)	ъ
	-	$ \begin{array}{c} c=0 & \mathring{v} > 0 \\ (c=0 & \rightarrow b > 0) \end{array} $	d=o	С
		С	Ъ	n

Bei der Möglichkeit C) (s. oben) ergibt sich also: b> 0 & c> 0, dann folgt: $Q = \frac{0.0 - b.c}{0.0 + b.c} = -1$

Wie bereits bemerkt wurde, lassen sich die für deterministische Beziehungen zwischen nominal skalierten Variablen gültigen Ergebnisse modifiziert verallgemeinern: Berücksichtigt man die in IV,1 gegebene Unterscheidung von Beziehungen nach ihrer Allgemeinheit und dem Meßniveau der Variablen, so ergeben sich folgende sechs Korrelationsausdrücke:

I Allgemeinheit der Beziehung

```
M
                                        deterministisch
e
                                        \Lambda_z ((X_z \leftrightarrow Y_z) \dot{V} (X_z \leftrightarrow -Y_z))
            nominal
B
n
                                        \land z \land w ((X_{z,w} \leftrightarrow Y_{z,w}) \dot{v} (X_{z,w} \leftrightarrow \neg Y_{z,w}))
             ordinal
i
                                         \land z \ ((\texttt{mx}_z \iff \texttt{ny}_z) \ \dot{\texttt{v}} \ (\texttt{mx}_z \iff \texttt{nny}_z)) 
            metrisch
e
u
                                        statistisch
M
                                       p((X_z \leftrightarrow Y_z) \dot{v} (X_z \leftrightarrow \neg Y_z)) = r
            nominal
B
n
                                        p((X_{z,w} \leftrightarrow Y_{z,w}) \dot{v} (X_{z,w} \leftrightarrow \neg Y_{z,w})) = r
             ordinal
i
                                        p(((X_Z \leftrightarrow Y_Z) \mathring{v} (X_Z \leftrightarrow \gamma Y_Z)) = r
            metrisch
```

Das Problem, ob bei deterministischen Beziehungen Korrelationen zwischen 1 und -1 und von o auftreten können, ist bei metrisch skalierten Variablen natürlich eindeitig entscheidbar; denn je nachdem, welche mathematischen Formeln man für m und n einsetzt, ergeben sich beliebige Korrelationen. Im übrigen soll nicht weiter auf zusätzliche oder veränderte Aspekte und Probleme der anderen Korrelationsausdrücke eingegangen werden.

Als Schlußergebnis von IV,4 sei festgehalten: Der Korrelationsbegriff $(x \leftrightarrow Y)$ \dot{v} $(X \leftrightarrow \neg Y)$ ist offensichtlich dem entsprechenden normal-

sprachlichen Korrelationsbegriff nicht semantisch völlig äquivalent, was aber daran liegen kann, daß dieser äquivok ist. Jedenfalls soll im Folgenden aber von dem logischen Ausdruck (3) als präzisiertem Korrelationsbegriff ausgegangen werden. da mir z.Z. kein adäquatere zur Verfügung steht.

Das Problem, ob deterministische Beziehungen zwischen dichotomen Variablen eine Korrelation von K=0 haben können, kann auch durch die Übersetzung von (X- \limes Y) v (X \limes -Y) in Q nicht eindeutig geklärt werden, denn gerade die Fälle, wo Q gleich o ist, erscheinen fragwürdig. Zur Lösung dieses Problems wäre es wohl erforderlich, zusätzlich zu der hier durchgeführten Präzisterung des Korrelationsbegriffes auch eine Bedeutungsanalyse oder Begriffsexplikation des Determinismusbegriffs vorzunehmen, was aber in diesem Rahmen nicht geleistet werden kann. Wichtig ist, daß der hier konstruierte Logische Korrelationsbegriff keinerlei Aussagen über Konditionalität oder gar Kausalität impliziert, sondern als rein beschreibend aufgefaßt werden kann. Inwieweit bestimmte Korrelationskombinationen es gestatten, über die Beschreibung von funktionalen Abhängigkeiten hinauszugehen und konditionale oder kausale Interpretationen und Erklärungen vorzunehmennund was exakt unter Konditionalität zu verstehen ist, das soll im Folgenden erörtert werden.

5) Die logische Struktur konditionaler Beziehungen

Nachdem nun eindeutig geklärt ist, was sich unter Korrelation verstehen läßt und damit, was die Tabellenanalyse als Korrelationsanalyse über die Beziehungen von Variablen auszusagen ermöglicht, muß gefragt werden, inwiefern die Tabellenanalyse noch mehr leisten kann als die Aufdeckung und Berechnung von Korrelationen. So wird etwa -wie schon erwähnt- bei MAYNTZ et al. (6/200-209) von bestimmten Korrelationskombinationen auf kausale Beziehungen geschlossen. Es ist dabei interessant, daß MAYNTZ et al. Kausalität genau so definieren, wie hier positive Korrelation definiert wurde, nämich : (X → Y) & (¬X → ¬Y); das dies aber keine befriedigende Definition von Kausalität sein kann, wurde in IV, 3 schon gezeigt. Hier soll nun aber nicht versucht werden, einen adäquateren logischen Ausdruck für Kausalität zu konstruieren, da dies auf Grund der in IV.3 aufgewiesenen Schwierigkeiten zu problematisch erscheint. Aber man könnte dennoch postulieren, daß sich außer über die Korrelation von Variablen mit hinreichender Sicherheit auch noch etwas über andere Beziehungen aussagen läßt, die man im weitesten Sinne als Folgebeziehungen bezeichnen könnte.

Nachfolgend sei eine Übersicht über verschiedene Folgebeziehungen gebracht; diese Übersicht ist selbstverständlich nur provisorisch zu verstehen, sie ist nicht nur unvollständig und unzureichend unterteilt,

sondern berührt eine Fülle grundlegenster erkenntnis- und wissenschaftstheoretischer Probleme, auf die hier aber nur punktuell eingegangen werden kann.

FOLGEBEZIEHUNGEN (Y 'folgt' X):

1) logisch

a) hinreichende Bedingung : X -> Y

b) notwendige Bedingung : $\neg X \rightarrow \neg Y$

c) Äquivalenz : X 😝 Y

d) Korrelation : $(X \leftrightarrow Y) \dot{v} (\neg X \leftrightarrow Y)$

e) logische Folge : $X \Rightarrow Y$

A) modus ponens : $(X \rightarrow YY) \& X \Rightarrow Y$

B) modus tollens : $(X \rightarrow Y) \& \neg Y \Rightarrow \neg X$

2) temporal : Y folgt zeitlich auf X

(X geht Y zeitlich voraus)

3) konditional : X bedingt Y

(X ist eine Teilursache von Y, Y folgt also nur dann aus X, wenn außer X auch noch die anderen

Teilursachen zutreffen)

4) kausal : X ist (eine) Ursache von Y

(Ursache = Summe aller Teilursachen)

5) rückwirkend (feedback) : X und Y bedingen einander wechselseitig

6) teleologisch (final) : Y ist Zweck ('Zielursache') von X

Diese verschiedenen Folgebeziehungen lassen sich sprachlich fast alle mit 'wenn - dann' ausdrücken, dies ist wohl zugleich Symptom und (eine) Ursache dafür, daß diese Beziehungen höchst selten genau auseinandergehalten werden, selbst in der Analytischen Wissenschaftstheorie. Im Weiteren sollen hier außer den logischen nur die konditionale Folgebeziehung berücksichtigt werden, da sie von besonderer Bedeutung für die Wissenschaft ist und am ehesten einer tabellenanalytischen Interpretation zugänglich zu sein scheint. Selbstverständlich ist eine konditionale Beziehung immer zugleich auch eine temporale Beziehung, so daß auch diese Berücksichtigung ffinden muß.

Es egibt sich nun folg endes Problem: Wenn man untersuchen will, ob sich mittels der Tabellenanalyse auch konditionale Beziehungen eruieren lassen und wenn man mittels logischer Analyse vornehmen will, muß man fragen, inwieweit sich konditionale Beziehungen überhaupt logisch ausdrücken lassen; es bestehen nämlich zwischen logischen und konditionalen Beziehungen zumindestens zwei entscheidende Unterschiede:

- (I) Logische Beziehungen sind zeitlos, dagegen besteht bei Konditionalität zwischen Teilursache und Wirkung immer ein -wenn auch noch so geringer- Zeitabstand.
- (II) Logische Beziehungen sind 'symetrisch', das soll hier bedeuten:

Wenn Y in irgendeiner Weise aus X folgt, dann gilt auch, daß X (oder -X) in irgendeiner Weise aus Y (oder -Y) folgt. Konditionale Beziehungen sind dagegen asymetrisch!: Wenn X Teilursache von Y ist, dann kann Y nicht Teilursache von X sein, da eben die Ursache der Wirkung zeitlich vorausgehen muß. (Ausnahmen hiervon gibt es in der Mikrophysik.) Hintergrund dieser Fragestellungen ist das fundamentale Problem: Sprache, Logik/ Wirkklichkeit oder: Satz, Aussage/ Sachverhalt. Können logische Beziehungen nur zwischen Sätzen oder auch zwischen Sachverhalten bestehen und können umgekehrt konditionale Beziehungen nur zwischen Sachverhalten oder auch zwischen Sätzen bestehen? Und ist es möglich, konditionale Beziehungen zwischen Sachverhalten adäquat durch logische Beziehungen zwischen Sätzen wiederzugeben.oder ist es z.B. nicht berechtigt, von der Zeitlichkeit konditionaler Beziehungen zu abstrahieren?Es ist sicher richtig, daß in diesem Aufsatz bisher mehr Fragen aufgeworfen, als beantwortet wurden, aber dies ist wohl angemessener als fragwürdige Pseudoantworten.

Festzuhalten bleibt: TATSÄCHLICH FESTSTELLEN KANN MAN von den verschiedenen aufgezeigten Folgebeziehungen NUR ZEITLICHE UND LOGISCHE BEZIEHUNGEN, (und dies auch nur für vergangene oder gegenwärtige Sachverhalte, also in Form von finiten Sätzen (s. IV,6)), ALLE DARÜBER HINAUSGEHENDEN AUSSAGEN SIND letztlich INTERPRETATIONEN. Dadurch, daß man Kausalität einfach als logische Ableitbarkeit (aus allgemeinen Gesetzen und Randbedingungen) neu definiert, ist natürlich nichts gewonnen, denn dies entspricht offensichtlich nicht dem Gehalt des normalsprachlichen, ursprünglichen Kausalitätsbegriff.

Trotz der aufgezeigten Problematik soll versucht werden, bestimmte Be ziehungen, die sich durch Tabellenanalyse aufdecken lassen und die bei MAYNTZ et al. als Monokausalität, Multikausalität, Scheinkorrelation und Intervention bezeichnet werden, hier aber nur als konditionale Interpretattionen aufgefaßt werden sollen, logisch zu formulieren.

A) MMonokonditionalität'

Ich will mich hier der wissenschaftstheoretisch üblichen Formulierung $X \to Y$ anschließen,allerdings mit Einwänden: Denn es erschiene eiw www gentlich adäquater,den Sachverhalt: X ist Teilursache von Ywff durch Y \to X auszudrücken; denn wenn X Teilursache von Y ist, fwf bedeutete dies ja (allerdings nur, wenn es für Y nur eine Gesamtursache gibt), daß X logisch notwendig ist für Y,Y aber logisch hinreichend für X, und genau das wird eben durch Y \to X und nicht durch X \to Y ausgedrückt.

Mögliche Varianten der Monokonditionalität sind: $X \to \neg Y$, $\neg X \to Y$, $\neg X \to \neg Y$.

Man müßte nun überprüfen,inwieweit diese Ausdrücke die z.B. bei MAYNTZ et al. (6/200-209) angegebenen Korrelationsbedingungen für 'Mono-kausalität' (bzw. Bestätigung) (vgl.IV,2) erfüllen. Dies soll hier aber nur am Fall der Multikenditionalität exemplarisch vorgenommen werden (s.dort).

Essei jedoch betont: Daß $X \rightarrow Y$ gilt (in der oben erläuterten Auffassung),ist nur eine notwendige,aber keine hinreichende Bedingung dafür, daß eine monokonditionale Beziehung vorliegt; die betreffende Beziehung könnte auch rein zufällig $X \rightarrow Y$ realisieren. In diesem Fall muß aber keineswegs eine andere konditionale Beziehung, z.B. Multikonditionalität, vorliegen, sondern es kann eben auch überhaupt keine Konditionalität vorliegen.

B) 'Mültikonditionalität'

Als Grundform könnte man etwa

```
X & t \rightarrow Y
             oder (X \rightarrow Y) & (t \rightarrow Y) annehmen, was nicht logisch äqui-
                        ww
w w w
       TAT TAT
                              W
                                 W
                                    w w
                            ff
       ff
                        ff
w w
       W W
                        ff
wff
                                     wf
       wf
                     W
ffw
                     f
       W W
                        WW
                                     W W
ffw
                     f
                        wf
                                     ff
       wf
                                 f
fff
       W W
                        wf
                                     wf
       Wf
                              W
```

valent ist (vgl. Wahrheitstafeln). Auf die vielen möglichen Varianten soll hier nicht eingegangen werden.

Nach LAZARSFELD (5/122) gelten im Fall der Multikonditionalität die Bedingungen: (XY,t) \neq o und (Xt) = o.MAYNTZ et al. nennen (6/206): (XY,t_i) (XY) und ebenfalls (Xt) \sim o.Es bestehen zwei Möglichkeiten, zu überprüfen, inwieweit die logischen Ausdrücke für Multikausalität diese Bedingungen erfüllen. (Im Folgenden soll nur der Ausdruck (X \rightarrow Y) & (t \rightarrow Y) berücksichtigt werden.)

- A) Man transformiert die Wahrheitswerte des logischen Ausdrucks für Multikausalität in Tabellenwerte und berechnet dann mit einem Korrelationskoeffizienten die verschiedenen Korrelationen zwischen den drei Variablen (vgl. IV,4), die dann eben mit den Korrelationsbedingungen übereinstimmen müßten.
- b) Man übersetzt die Korrelationsbedingungen selbst in logische Aus-

drücke und überprüft, in welchen logischen Beziehungen diese Ausdrücke zu dem logischen Ausdruck für Multikausalität stehen.

Hier soll nur der zweite Weg ansatzweise beschritten werden.

(I) $(XY,t) \neq 0$

Diese Bedingung besagt ja, daß wenn man t einführt und konstant hält, die bedingte Korrelation zwischen X und Y üngleich o ist; das könnte man folgendermaßen formalisieren: $t \to (X \longleftrightarrow Y)$ v $(X \longleftrightarrow \neg Y)$. Je nachdem, ob man (XY, t) = 0 als notwendige oder hinreichende Bedingung für Multikondition: ansieht, ergeben sich natürlich unterschiedliche Konsequenzen; m.E. ist es nur akzeptabel, (I) als notwendige Bedingung zu betrachten. Dann müßte also gelten:

 $(X \to Y) & (t \to Y) \Rightarrow \exists t \to (X \leftrightarrow Y) \dot{v} (X \leftrightarrow Y).$ Dies trifft auch zu, da der Korrelationsbegriff eben tautologisch ist.

(II) (XY,t) > (XY)

Bei MAYNTZ et al. wird (II) anstatt (I) gefordert.Offensichtlich ist dabei gemeint, daß (XY,t) entweder eine höhere positive oder eine höhere negative Korrelation aufweisen muß, jedenfalls soll das hier so verstanden werden.Man könnte (II) mit dem 'ordinalen Korrelationsbegriff' (vgl. S.16) ausdrücken, man kann aber auch den statistischen 'nominalen Korrelationsbegriff' verwenden.Man muß allerdings dann erst einmal festlegen, welches Verhältnis zwischen der Korrelation und der Wahrscheinlichkeit einer Beziehung besteht.Nach meinen Überlegungen müßte folgende Formel adäquat sein: k = 2p - 1 (k= Korrelation, p= Wahrscheinlichkeit).Esgilt also z.B. dann: wenn p=0,dann k=-1,wenn p=0,5,dann k=0,wenn p=1,dann k=1.Istp>0,5,besteht positive Korrelation (kp),ist p<0,5,negative Korrelation (kn);und zwar gilt für kp: je größer p, desto höher kp,und für kn: je kleiner p,desto höher kn.Also müßte man (XY,t)> (XY) folgendermaßen logisch formulieren können:

 $p(t \rightarrow (X \leftrightarrow Y)) > p(X \leftrightarrow Y) \dot{v} p(t \rightarrow (X \leftrightarrow -_{1}Y)) < p(X \leftrightarrow -_{1}Y)$

In welcher logischen Beziehung dieser Ausdruck zum Ausdruck für Multikonditionalität steht, kann hier nicht weiter untersucht werden. (III) (Xt) = 0

Diese Forderung berührt natürlich wieder die Problematik, ob die Korrelation bei deterministischen Beziehungen zwischen dichotomen Wariablen überhaupt den Wert o annehmen kann. Wenn man nämlich keine Korrelation einfach durch Verneinung des Korrelationsausdruckes (3) formalisiert, also $\neg((X \longleftrightarrow Y) \ \dot{v} \ (X \longleftrightarrow \neg Y))$, so ergibt sich ein logischer Widerspruch, was wissenschaftlich natürlich nicht akzeptabel ist. (vgl. IV,4) Auch hier entsteht dagegen kein Problem, wenn man mit dem statistischen Korrelationsbegriff arbeitet. Gemäß der Formel k = 2p - 1 ergibt sich für k = 0 p = 0,5, es müßte also gelten: $p((X \longleftrightarrow Y) \ \dot{v} \ (X \longleftrightarrow \neg Y)) = 0$,5. Ist die Wahrscheinlichkeit also in beiden Tabellen (für positive und negative Korrelation) gleich 0,5, ergibt sich p = 0,5 v 0,5, woraus logisch

eben p=0,5 folgt. Allerdings erscheint auch hier eine Überprüfung, ob dieser Ausdruck aus dem Ausdruck für Multikausalität logisch ableitbar ist, zu aufwendig.

C) Scheinkorrelation und Intervention

Als Grundformen sind denkbar: für Scheinkorrelation: ($t \to X$) & ($t \to Y$), bzw. der äquivalente Ausdruck: $t \to X$ & Y; für Intervention:

 $(X \rightarrow t)$ & $(t \rightarrow Y)$. Jedenfalls haben diese beiden Beziehungen nicht die gleiche logische Struktur, obwohl man die gleichen Korrelationsbedingungen an sie stellt.

Zusammenfassend sei festgehalten: Um zu untersuchen, ob die Tabellenanalyse auch über die Korrelation hinaus etwas über (Folge-) Beziehungen zwischen Variablen auszusagen ermöglicht, wurden die wichtigsten Folgebeziehungen aufgeführt und ihre logischen Eigenschaften diskutiert. Dabei wurden die konditionalen Beziehungen als diejenigen konstatiert, über die Tabellenanalyse am ehesten Aussagen aufzustellen berechtigt, während dies für kausale Beziehungen nicht zutrifft. Es wurde versucht, entsprechend wie in IV,4 für den Korrelationsbegriff, den Begriff der konditionalen Beziehung durch logische Formalisierung präzise und eindeutig zu bestimmen und damit erst wissenschaftlich brauchbar zu machen. Ebenso wurden die in der Literatur angegebene Korrelationsbedingungen für kausale Beziehungen logisch umgeformt und es wurde überprüft, inwieweit die logischen Konditionalitätsausdrücke diesen Bedingungen genügen!Insgesamt traten bei diesen Bemühungen beträchtliche Probleme auf, die nur zum Teil oder recht unbefriedigend gelöst werden konnten, was auch daran liegen mag, daß mir keinerlei Literatur speziell zu diesem Themenbereich zur Verfügung stand. Dennoch konnten folgende Hauptergebnisse gewonnen werden: Tatsächlich feststellen kann man durch Tabellenanalyse nur logisch-mathematische Beziehungen, etwa Korrelation. Die Interpretation bestimmter solcher Beziehungen als konditional, d.h. teilkausal, ist letztlich immer eine Spekulation, die vielleicht nur auf der spezifischen Struktur unseres neuromentalen Systems und nicht der der Wirklichkeit beruht. Auch die etwa von MAYNTZ et al. angegebenen Korrelationsbedingungen sind -einmal ganz abgesehen von ihrer generellen Problematik- nur notwendige, aber keinesfalls hinreichende Bedingungen für Konditionalität, wenn überhaupt. Selbst perfekte Korrelationen zwischen zwei Variablen, die in endlich vielen Realisationen einer Beziehung auftraten, können rein zufällig sein und durch spätere Fälle nahezu wieder völlig aufgehoben werden. (vgl. IV, 6) Dadurch, daß man einfach abweichend vom normalsprachlichen Konditionalitätsbegriff neu definiert, eine Beziehung sei dann konditional, wenn sie die und die Korrelationsbedingungen erfüllt, ist natürlich nichts gewonnen, sondern nur eine naive Scheinlösung volzogen.

Wenn auch in diesem Aufsatz nicht viele konkrete Ergebnisse erbracht wer-

den können, so sollte er doch wenigstens durch die aufgezeigten Probleme bei der logischen Präzisierung demonstrieren, daß die Begriffe und Sachverhalte, um die es hier geht, keinesfalls einfach und unproblematisch sind, dann wäre es nämlich auch einfach, sie formal darzustellen. Gerade deshalb ist die logische Formalisierung so wichtig, weil sie die scheinbare Leichtverständlichkeit normalsprachlicher Formulierungen als unpräzise und mehrdeutig entlarvt.

6) Tabellenanalyse und Induktionsproblematik

Angesichts des beachtlichen Umfangs, den der Aufsatz schon erreicht hat, kann die enorm komplexe und schwierige Problematik dieses Themenbereiches nur ganz oberflächlich gestreift werden. Alle Probleme, um die es hier geht, hängen letztlich mit dem sogenannten Induktionsproblem zusammen, d.h. dem Problem, ob und inwiefern man von Prämissen auf einen Schlußsatz schließen kann, dessen Informationsgehalt größer ist als der der Konjunktion der Prämissen. Im einzelnen gehts es um:

- A) Bestätigung eines (mono)konditionalen Zusammenhangs
- B) Signifikanz einer Korrelation, Generalisierung von an einer Stichprobe gewonnen Tabellenwerten auf die Grundgesamtheit
- C) Induktiv-statistische Erklärung

Als ein mögliches Ergebnis bei der Einführung eines Brittfaktors gilt die Bestätigung des angenommenen (mono)kausalen Zusammenhangs bzw. (mono)konditionalen Zusammenhangs (gemäß der hier vertretenen Auffassung) Angenommen, man hat festgestellt, daß für alle bisher beobachtete z die (als monokonditional interpretierte) Beziehung $X_z \to Y_z$ zutrifft, man stellt nun die Hypothese auf: Λ z ($X_z \to Y_z$). Hier stellt sich nun die Frage: a) Inwiefern ist es berechtigt, von einer Teilmenge aller z auf alle z zu schließen? Eine eng damit zusammenhängende Frage ist: b) Wie läßt sich die Aussage Λ z ($X_z \to Y_z$) begründen, überprüfen und bestätigen?

Die Beantwortung dieser Fragen hängt entscheidend davon ab,ob die Menge aller z endlich oder unendlich ist,bzw. ob man Λ z $(X_Z \to Y_Z)$ als raum-zeitlich beschränkte (finite) Allaussage oder als unbeschränkte (infinite) Allaussage auffaßt. In der Wissenschaftstheorie ist man sich heute weitgehend einig,daß sich Naturgesetze nur adäquat als infinite Sätze (entsprechend Beziehungen) ansehen lassen und daß man auch in den Sozialwissenschaften solche Sätze anstreben sollte,obwohl deren Aufstellung beim heutigen Forschungsstand noch kaum möglich ist. Für infinite Allsätze sprechen folgende Argumente:

a) Man stellt Allsätze im wesentlichen dazu auf,um aus ihnen mittels Randbedingungen Erklärungen oder Prognosen logisch abzuleiten. Man vermute z.B., daß der Satz Λ Menschen (frustriert_{Mensch} \longrightarrow aggressiv_{M.})

wahr ist, d.h. entsprechend, daß die Beziehung, die er beschreibt, in der

Wirklichkeit realisiert ist. Faßt man diesen Satz nun als zeitlich beschränkte Allaussage auf, die nur alle heute lebenden und in der Vergangenheit gelebt habenden Menschen miteinbezieht, dann ist es logisch nicht zu rechtfertigen, daraus Prognosen abzuleiten über Menschen, die schon morgen (oder später) geboren werden; dies ist aber höchst inadäquat.

Ebenso sind Erklärungen mit räumlich begrenzten Allaussagen problematisch: In dem Beispiel: Alle Personen in diesem Raum sind männlich

Person a befindet sich in diesem Raum

Person a ist männlich

wird die Männlichkeit von a dadurch erklärt, daß sie sich in einem besimmten Raum befindet, was offensichtlich unhaltbar ist.

Man benötigt also in der Wissenschaft unbeschränkte Allaussagen, um konditionale Beziehungen auszudrücken.

Aus endlich vielen Beobachtungen, die der Mensch als endliches Wesen aber immer nur vollziehen kann, lassen sich aber keine Aussagen über . eine (aktual oder potential) unendliche Menge ableiten, aus endlich vielen Beobachtungssätzen kann man nicht logisch auf einen infiniten Allsatz schließen, d.h. ein solcher Satz ist nicht verifizierbar. Es läßt sich aber auch nichts über die Wahrscheinlichkeit aussagen. daß ein unbeschränkter Allsatz wahr ist (außer daß sie größer o ist), denn jede beliebige endliche Menge ist angesichts einer unendlichen Menge verschwindend klein. Aus der Tatsache, daß alle Menschen, die bisher gelebt haben, sterblich waren, folgt -vorrausgesetzt die Menge aller Menschen in Vergangenheit, Gegenwart und Zukunft ist unendlich- für die Wahrscheinlichkeit, daß alle Menschen sterblich sind, nur daß sie größer o ist. Es ist also nicht möglich, konditionale Beziehungen, die durch einen unbeschränkten Allsatz beschrieben werden, vollständig zu bestätigen. Dagegen genügt bereits ein singulärer Satz.um den Allsatz bzw. die von ihm beschriebene Beziehung streng logisch nach dem Modus tollens $(X \rightarrow Y) \& \neg Y \rightarrow \neg X \text{ (vgl. S.18)}$ zu falsifizieren. Bei infiniten statistischen Sätzen ist die Sachlage noch komplizierter,

denn sie sind nicht nur nicht verifizierbar, sondern auch nicht falsifizierbar, da sie nichts 'verbieten'. Der Satz: '99,9 % aller Menschen
sind unsterblich' wird dadurch, daß bisher alle Menschen sterblich
waren, nicht falsifiziert, (angenommen, die Menge aller Menschen ist
unendlich), denn o,1 % einer unendlichen Menge ist eben immer noch eine
unendliche Menge. Man kann es auch so formulieren: Aus der relativen
Häufigkeit einer Beziehung in einem endlichen Abschnitt einer unendlichen Folge kann man nicht auf den wahrscheinlichen Grenzwert der relativen Häufigkeit in der gesamten Folge schließen.

Die Bestätigung konditionaler Beziehungen ist eine höchst unsichere An-

gelegenheit: Angenommen, man stellt für eine endliche Menge von z eine

positive Korrelation bezüglich der Variablen X und Y fest, es gilt also: $V z (X_2 \rightarrow Y_2)$; man stellt die Hypothese auf, daß hier eine konditionale Beziehung vorliegt, die für alle z wahr ist, es soll also gelten : $\bigwedge z (X_z \rightarrow Y_z)$. Dieser Satz ist aber nicht verifizierbar. Aber selbst wenn man ihn verifizieren könnte, wäre damit die monokonditionale Beziehung noch nicht als selbstständige Beziehung bestätigt, da sie ja nur die Folge einer monokonditionalen Beziehung sein könnte, denn es gilt ja: $(X \rightarrow Y)$ & $(t \rightarrow Y) \Rightarrow (X \rightarrow Y)$. Um dies aber zu überprüfen, müßte man eine unendliche Menge von Testfaktoren einführen, was nicht möglich ist. Und dennoch könnte man selbst dann nicht sicher sein, daß die bestätigte Beziehung tatsächlich konditionaler und nicht zufälliger Natur ist. Nungeht man trotz der aufgezeigten Schwierigkeiten in den Sozialwissenschaften meistens von finiten Sätzen aus. Und hier ist die Situation etwas verändert: Finite Sätze sind nämlich prinzipiell verifizierbar. aus praktischen Gründen ist dies aber kaum je durchführbar, sondern man überprüft sie mittels einer Zufallstichprobe. Eine Zufallsstichprobe ist nur dann möglich, wenn die Grundgesamtheit physisch oder symbolisch geg e nwärtig und manipulierbar ist, weshalb es auch nicht möglich ist, aus einer unendlichen Population eine Zufallsstichprobe zu ziehen. Auf die umfangreiche Problematik der Zufallsstichprobe kann hier nicht eingegangen werden; es sei nur kurz bemerkt: Nach der mathematischen Wahrscheinlichkeitstheorie ist es möglich, aus der Häufigkeitsverteilung einer Beziehung in einer Stichprobe innerhalb berechenbarer Fehlergrenzen auf die wahrscheinliche Häufigkeitsverteilung in der Population zu schließen. Je mehr Sicherheit man aber für diesen Schluß verlangt ('Vertrauensintervall'), desto geringer wird der Informationsgehalt des Schlußsatzes. Nimmt man nicht einen Informationsgehalt von o in Kauf, (was natürlich absurd wäre), dann bleibt immer die Gefahr des sog. Typ - I - Fehlers bzw. Typ - II - Fehlers bestehen, d.h. die Gefahr, eine wahre Hypothese irrtümlich zu verwerfen oder eine falsche Hypothese irrtümlich für wahr zu halten. Inwieweit die Anwendungsbedingungen der mathematischen Wahrscheinlichkeitstheorie in der wissenschaftlichen Praxis realisierbar sind, ist ein weiteres Problem. Von ähnlichen Überlegungen geht man auch bei der Überprüfung der Sicherheit (Signifikanz) einer Beziehung in einer Tabelle aus, indem man die Differenz zwischen beobachteter und bei Unabhängigkeit der Variablen zu erwartenden Häufigkeitsverteilung berechnet und daraus Wahrscheinlichkeitsaussagen ableittet, Selbstverständlich erlaubt die Signifikanz einer Stichprobenbeziehung in einer Tabelle nur dann einen Rückschluß auf die Population, wenn eine Zufallsstichprobe vorliegt.

V LITERATUR

(Literaturangaben im Text: Die erste Zahl verweist auf die Nummer im Literaturverzeichnis, die zweite auf die Seitenzahl in der betreffenden Publikation.)

- 1 Frank, H. Informationspsychologie in: H. Frank (Hg.) - Kybernetik Frankfurt 1970, 7. A., S. 243-258
- 2 Keidel, W.D./ Elementar funktionen und Informationsverarbeitung innerhalb des Zentralnervensystems in: H. Frank (Hg.) Kybernetik Frankfurt 1970, 7.A., S. 45-62
- 3 Klaus, G. Moderne Logik, Berlin 1967, 4.A.
- 4 Kutschera, F. Wissenschaftstheorie, Bd. I.II, München 1972
- 5 Lazarsfeld, P.F- Interpretation of Statistical Relations as a Research Operation in: P.F.Lazarsfeld, M.Rosenberg (Hg.) The Language of Social Research Glencoe 1955, S.115 174
- 6 Mayntz,R. et al. Einführung in die Methoden der empirischen Soziologi Opladen 1972,3.A.
- 7 Roghmann, K. Methoden der empirischen Soziologie in: E.Fels et al. - Methoden der Sozialwissensch. München und Wien 1967, S. 197-198, 201-203
- 8 Sahner, H. Schließende Statistik Stuttgart 1971, S. 124-140
- 9 Stegmüller, W. Das Problem der Induktion: Humes Herausforderung und moderne Antworten in: H. Lenk (Hg.) Neue Aspekte der Wissenschaftstheorie, Braunschweig 1971, S. 13-74
- 10 Zeisel, H. Die Sprache der Zahlen Berlin und Köln, 1970