BEDINGTE WAHRSCHEINLICHKEIT

Robert Müller, Wien

§ 0 VORBEMERKUNGEN

Blättert man in der von W. PESCHEK erstellten Untersuchung (vgl. L 16), so muß man feststellen, daß Wahrscheinlichkeitsrechnung, insbesonders Bedingte Wahrscheinlichkeit, zu jenen Themen gehört, welche im Unterrichtsgeschehen allzu oft nicht behandelt werden. Im Zuge der Neuorientierung des Mathematikunterrichtes scheint gerade diesem Thema in naher Zukunft besondere Bedeutung zuzukommen. Wir wollen uns hier nicht dem Grundlagenproblem widmen, welches spätestens seit KOLMOGOROV für die Mathematik als "gelöst" gilt, wenn man die Form der Axiomatisierung als Lösung akzeptieren will. Vor zwei Jahren wurde hieorts von R. LAUSSERMAYER (vgl. L 14) dazu referiert. Wir stellen uns hier auf einen naiv-axiomatischen Standpunkt insoferne, als Wahrscheinlichkeit als undefinierter Grundbegriff Verwendung findet. Aber anders als in einer strengen HILBERTschen Axiomatisierung, in der nicht das Wesen der Dinge,sondern bloß die Beziehung zwischen den Dingen von Bedeutung ist, kommt es uns sehr wohl darauf an, dem Schüler eine Vorstellung vom Begriff "Wahrscheinlichkeit" zu vermitteln.

Wir leben in einer Welt, deren Ablauf durch das Eintreten bestimmter Ereignisse bestimmt wird. Für das Überleben in dieser Welt ist es unerläßlich, daß wir über das Eintreten oder Nichteintreten von Ereignissen im notwendigen Umfang Voraussicht besitzen. Wahrscheinlichkeit "mißt" in gewissem Sinn diese unsere Erwartung. Unsere Erwartungshaltung resultiert aus unserer Erfahrung, und zwar aus unserer stammesgeschichtlichen wie persönlichen. Insoferne ist Wahrscheinlichkeit ein Begriff, den jeder von uns mehr oder weniger reflektiert benützt, und benützen muß. Offensichtlich leben wir in einer hoffnungslos komplizierten Welt, deren Problemen wir nur durch weitestgehende Vereinfachung begegnen können. Das Operieren mit Wahrscheinlichkeiten ist das Evolutionsprodukt dieser Bestrebungen.

jedes Eintreten eines Ereignisses (zumindest im Makrokosmos) wird durch das vorherige Eintreten anderer Ereignisse bedingt. Eine erste Vereinfachung wird darin bestehen, für das Eintreten eines ins Auge gefaßten Ereignisses von vornherein nur bestimmte, explizit aufgezählte Bedingungen als relevant zuzulassen. Die einem Ereignis so "zugemessene" Erwartung für sein Eintreten unter (vor-) bestimmten Bedingungen bezeichnet man als Bedingte Wahrscheinlichkeit. Die Unbedingte Wahrscheinlichkeit, oder kurz Wahrscheinlichkeit steht als letzte Konsequenz am Ende unserer Vereinfachungsbestrebungen insoferne, als keine Voraussetzungen mehr explizit angegeben werden, obwohl es solche gibt. In diesem Sinne scheint der Begriff der Bedingten Wahrscheinlichkeit der natürliche(re) zu sein. Trotzdem gehen alle dem Verfasser bekannten Bücher zur Wahrscheinlichkeitsrechnung vom (einfacheren?) Begriff der (unbedingten) Wahrscheinlichkeit aus.

Die folgenden Ausführungen sollen skizzieren, daß aus unterrichtspraktischer Sicht keine Notwendigkeit besteht, der Lehrbuchpraxis zu folgen. Vielmehr sprechen viele Gründe dafür eine gleichzeitige Genese beider Begriffe herbeizuführen. Um dies im unterrichtspraktischen wie fachwissenschaftlichen Sinne zu rechtfertigen, wurde das Manuskript zweigeteilt; und zwar in der Absicht, auf der jeweils rechten Seite den Unterrichtsablauf darzustellen, und auf der zugehörigen linken Seite durch didaktische, fachliche und persönliche Bemerkungen die Intentionen und den fachwissenschaftlichen Hintergrund zu beleuchten. Selbstverständlich kann es sich hier nur um eine diskussionswürdige Alternative handeln; gerade das Aufzeigen von Alternativen ist Sinn von Fortbildungsveranstaltungen.

Did. Anm.: Die Komplexität des Einführungsbeispieles mag viele von Ihnen erstaunen. Tatsächlich scheint jedoch erst ein weiter Rahmen einen homogenen Unterricht zu gestatten. An dieser Stelle kann auf dieses Problem nicht weiter eingegangen werden (vgl. WITTENBERG).

6.1 KONSTRUKTION VON BEDINGTEN UND UNBEDINGTEN WAHRSCHEINLICHKEITEN ANHAND VON VIERFELDERTAFELN

Beispiel 1: Es gilt den Zusammenhang zwischen "Wohnort" und "Einkommen" zu untersuchen. Eine erste grobe Analyse sieht für die beiden Merkmale folgende Ausprägungen vor:

"Wohnort" ...Stadt, Land (= nicht städtisch)

"Einkommen"..schlecht, gut (bezogen auf das Durchschnittseinkommen)

Der Stichprobenumfang beträgt 1500 Personen.

Aufgaben:1) Worauf hat man bei der Auswahl der Stichprobe zu achten ?
2) Wie könnte der Fragebogen aussehen? Welche Hinweise müßte er enthalten. ?
3) In welcher Weise könnte man das Umfrageergebnis (Zahlenmaterial) übersichtlich darstellen ?

Wir wollen die folgende Tabelle verwenden:

	Stadt A	Land $\overline{A}$	
Schlecht B	450	600	
Gut $\overline{B}$	150	300	

Abb. 1 Tabelle der Absoluten Häufigkeiten

Aufgaben:1) Errechne, wieviel Personen Städter sind,(am Land wohnen)!
2) Errechne, wieviel Personen gut, wieviel schlecht verdienen !
3) Wo erscheint der Stichprobenumfang ? (Kontrolle!)
4) Faßt man A,B als Mengen auf, so sind $\overline{A}$, $\overline{B}$ ...... ?
5) Welche Mengenoperation führt zu obiger Tabelle ?
6) Wie nennt man eine solche Tabelle im mathematischen Fachjargon ?

Die Lösung der Aufgaben führt schließlich zu folgender Matrix:

Did.Anm.: Zu den vordringlichsten Aufgaben unserer Unterrichtstätigkeit zählt es, den Schüler zum Mitdenken in aktiver und passiver Hinsicht anzuleiten. Zur Zeit scheint das passive Mitarbeiten, das Mitschreiben und Nachvollziehen der dargebotenen Gedankengänge im Unterricht das Übergewicht zu haben. Der vorgelegte Weg versucht, beide Verhaltensweisen gleichermaßen anzusprechen. Aufgabenphasen, in denen die Selbsttätigkeit und die Eigeninitiative der Schüler gefordert, und damit gefördert, wird, wechseln mit Vortragsphasen, in denen der rote Faden weitergesponnen wird, und Ergebnisse in die konventionelle Form gebracht werden. Der erste Block von Aufgaben auf S.3 läßt bewußt einen größeren Antwortenspielraum offen. Der Schüler soll erkennen, daß der ihm vom Lehrer gewiesene Weg nur einen, (wenn auch bewährten) darstellt, um an das Problem heranzugehen.

Die in Abb.1 angegebenen Zahlen sind fiktive, um "schöne" Ergebnisse zu erhalten und solcherart vorerst das Problem sinnvoller, adäquater Genauigkeit nicht thematisieren zu müssen.

Vierfeldertafeln, allgemein rechteckige Matrizen, entstehen "von selbst", wenn man "rechteckige" VENN-Diagramme zum Durchschnitt bringt.

Manche Begriffe, zu denen ich auch die Bedingte Wahrscheinlichkeit rechne, sind, obwohl sie eine einfache Definition besitzen (etwa über die selbst undefinierte Unbedingte Wahrscheinlichkeit - vgl.S.13 (*)), aus dieser heraus dem Schüler eher unverständlich. Das darf einen nicht wundern, läßt man in dieser axiomatischen Vorgangsweise altbekannte Forderungen außer acht. Gemäß FREUDENTHAL's Forderung nach einer "Didaktischen Phänomenologie" und BRUNER's Forderung, daß die Konstitution mentaler Objekte den Begriffen vorauszugehen habe, gilt es vorerst einmal, der Begriffsgenese ausreichend Raum und Zeit zu widmen.

	A	$\overline{A}$	Σ
B	450	600	1050
$\overline{B}$	150	300	450
Σ	600	900	1500

Abb. 2: Tabelle der Absoluten Häufigkeiten samt Randverteilungen

Um eine Vergleichbarkeit mit ähnlichen Untersuchungen von anderem Stichprobenumfang zu ermöglichen, wird man zu relativen Häufigkeiten übergehen.

A) Im ersten Fall sei die vorgegebene Stichprobe die Bezugsmenge.
Aufgabe: Durch welche Rechnung erhält man die relativen Häufigkeiten?
Die Lösung der Aufgabe führt zu folgender Matrix:

	A	$\overline{A}$	Σ
B	$P(A \wedge B) = \frac{450}{1500} = \frac{3}{10}$	$P(\overline{A} \wedge B) = \frac{600}{1500} = \frac{2}{5}$	$P(B) = \frac{7}{10}$
$\overline{B}$	$P(A \wedge \overline{B}) = \frac{150}{1500} = \frac{1}{10}$	$P(\overline{A} \wedge \overline{B}) = \frac{300}{1500} = \frac{1}{5}$	$P(\overline{B}) = \frac{3}{10}$
Σ	$P(A) = \frac{2}{5}$	$P(\overline{A}) = \frac{3}{5}$	1

Abb. 3 Tabelle der Unbedingten Wahrscheinlichkeiten

Fachl.Anm.: Die Verwendung der Relativen Häufigkeit als "vernünftigster Schätzwert" für die Wahrscheinlichkeit ist naheliegend und in gewissem Sinn mathematisch präzisierbar (Maximum Likelihood-Schätzung).Nichtsdestoweniger sind die Schwierigkeiten des Überganges von den Relativen Häufigkeiten zu Wahrscheinlichkeiten prinzipiell nicht behebbar.

Schon die Verwendung Relativer Häufigkeiten und ihre Darstellung als "Verhältniszahlen" ist problematisch, vor allem das Kürzen: Definieren $\frac{2}{5}$ und $\frac{6}{15}$ dieselbe Relative Häufigkeit, dieselbe Wahrscheinlichkeit? Angenommen, zwei Spieler spielen gleich gut, so daß der Zufall die Spiele entscheidet mit $p = \frac{1}{2}$. Gemäß der Binomialverteilung ist die Wahrscheinlichkeit,2 von 5 Spielen zu gewinnen $\binom{5}{2} \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^3 = 0{,}3125$, und 6 von 15 Spielen zu gewinnen analog $\binom{15}{6} \cdot \left(\frac{1}{2}\right)^6 \cdot \left(\frac{1}{2}\right)^9 = 0{,}152$, d.h., gleiche Relative Häufigkeiten definieren verschiedene Wahrscheinlichkeiten. Wo liegt der Fehlschluß?

Did.Anm.: Aufgrund der prinzipiellen Schwierigkeiten der logischen Grundlegung des Wahrscheinlichkeitsbegriffes im Verhältnis zum Begriff der Relativen Häufigkeiten gehen wir von zwei Thesen aus. Erstens begnügen wir uns mit einer psycho-logischen Grundlegung der Begriffe und vermeiden eine voreilige Axiomatisierung. Zweitens versuchen wir, den Zusammenhang zwischen Relativen Häufigkeiten und Wahrscheinlichkeiten durch den Trick einer zweifachen Interpretation ein- und derselben Rechengröße herzustellen. Um dabei den Maßcharakter von Wahrscheinlichkeit herauszustreichen, wurde von vornherein der Begriff der Relativen Häufigkeit seiner kombinatorischen Bedeutung weitgehend entkleidet, indem er durch den Begriff "Anteil" ersetzt wurde. Abgesehen davon, daß der Begriff "Anteil" nicht mehr nur Werte aus der Menge der Rationalen Zahlen zuläßt und so dem Maßbegriff nähersteht, erlaubt er auch die Verwendung der gleichen Notation "P" für Anteil und Wahrscheinlichkeit.

Aufgabe: Drücke mit eigenen Worten aus, welche Bedeutung den eben errechneten Größen zukommt !

Sei E das Ereignis, der Stichprobe anzugehören. Dann wollen wir die verbale Lösung obiger Aufgaben folgendermaßen formal fassen:
Für z.B. "Der Anteil der Städter in der Stichprobe beträgt 600/1500, also 2/5",
wird notiert: $P(A/E)=\frac{2}{5}$ P ... Anteil(engl. part)
und gelesen:" Der Anteil von A in E ist zwei Fünftel" .

Aufgabe: Lies in der Tabelle die Werte ab für: $P(\overline{A}/E)=$
$P(B/E)=$
$P(\overline{B}/E)=$
$P(A\cap B/E)=$
...
$P(E/E)=$

Wenn sich wie in der letzten Aufgabe alle Ereignisse auf dasselbe Ereignis beziehen, dh., für jede Person in jedem Fall die Vor-Bedingung, der Stichprobe anzugehören, erfüllt ist, so verzichtet man in der Notation auf diese "Bedingung", und schreibt statt $P(A/E)$ nunmehr $P(A)$,-(und merkt sich stillschweigend, daß man $A\subseteq E$ vorausgesetzt hat).

Aufgabe: Ergänze in Abb. 3 die Kurz-Notation!
Aufgabe: Diskutiere die Aussage:" Unter der Voraussetzung, der Stichprobe anzugehören, lebt eine (zufällig gewählte) Person mit der Wahrscheinlichkeit $\frac{2}{5}$ in der Stadt".

Im Sinne der obigen Uminterpretation der Zahlenwerte aus Abb. 3 führen wir als neuen Grundbegriff den der "Wahrscheinlichkeit"(engl. probability) ein, und verwenden dieselbe Notation wie bisher, weil es sich um <u>dieselbe Zahl mit zwei Interpretationsmöglichkeiten</u> handelt:
Unter $P(A/E)$ verstehen wir (nun auch) die Wahrscheinlichkeit für das Eintreten des Ereignisses A unter der Bedingung, daß das Ereignis E

eingetreten ist, oder kurz: P(A/E) ist die Bedingte Wahrscheinlichkeit von A unter der Bedingung E. Da das Eintreten von E auch hypothetischer Natur sein kann, spricht man von der Wahrscheinlichkeit von A unter der Hypothese E.

Stützen sich wie hier alle errechneten Wahrscheinlichkeiten auf dieselbe Bedingung, die somit stets erfüllt ist, so sieht man wiederum von dieser "Bedingung" ab und spricht von Unbedingter Wahrscheinlichkeit. Im dem vorliegenden Fall schreibt man z.B. : P(A) für P(A/E).

Did.Anm.: Die Verwendung des gleichen Namens "P" erlaubt den Übergang von Anteilen zu Wahrscheinlichkeiten ohne Änderung der Notation. Man muß sich über die Vor- und Nachteile dieser Vorgangsweise im klaren sein. Einerseits bringt die notationelle Vereinheitlichung den Vorteil, den Übergang, das Uminterpretieren ungemein zu erleichtern, andererseits könnte sie Anlaß sein, auch die Begriffe selbst zu identifizieren. Aus meiner Erfahrung hat die konsequente Übung des Uminterpretierens diesen möglichen Nachteil aber nie zum Tragen kommen lassen. Die Art unserer Vorgangsweise spiegelt deutlich die Art unseres Verständnisses von Wahrscheinlichkeit wieder, nämlich die, aus unserer Erfahrung (Bedingungen) Erwartungen zu bilden. Sie steht im Einklang mit unseren erzieherischen Absichten: Der Schüler wird angehalten, Zahlenmaterial zu interpretieren, seine Interpretation verbal und formal zu formulieren und so seine Urteilsfähigkeit zu entwickeln.

Die Einführung des Begriffes "Anteil", der nur unter Angabe einer Bezugsmenge sinnvoll ist, führt solcherart zwangsläufig zuerst zum Begriff der Bedingten Wahrscheinlichkeit, zumindest zwingt sie zur expliziten Darstellung der Bezugsmenge, welche in der üblichen Notation Unbedingter Wahrscheinlichkeit fehlt.

Fachl.Anm.: Die systematische Erstellung der Tabellen in Abb.4 und Abb.5 hat einen erkenntnistheoretischen Grund. Die beiden Tafeln sind ein naheliegender Ausgangspunkt für Überlegungen zur "Ursache-Wirkungs-Verkettung". Besonders die in Abb.2 dargestellte simultane Koinzidenz von Merkmalsausprägungen findet eine Entsprechung auf der Ebene sukzedaner Koinzidenz. Die systematische Berechnung Bedingter Wahrscheinlichkeiten ist ein Weg, das statische Verständnis Unbedingter Wahrscheinlichkeiten durch ein dynamisches Verständnis über Ereignisabläufe zu ersetzen.

Wir wollen das Errechnen bedingter Wahrscheinlichkeiten einüben, indem wir zu anderen Bezugsmengen übergehen.

B) Im zweiten Fall sei jeweils eines der Merkmale vorausgesetzt.

+ Einerseits kann man das Merkmal Wohnort voraussetzen, dh. seine beiden Ausprägungen als Bezugsmengen voraussetzen. Bei Division jeder Spalte aus Abb. 2 durch die jeweilige Spaltensumme erhält man:

$\swarrow\!\!\swarrow$	A	$\bar{A}$	Σ
B	$P(B/A) = \frac{450}{600} = \frac{3}{4}$	$P(B/\bar{A}) = \frac{600}{900} = \frac{2}{3}$	$\neq 1$
$\bar{B}$	$P(\bar{B}/A) = \frac{150}{600} = \frac{1}{4}$	$P(\bar{B}/\bar{A}) = \frac{300}{900} = \frac{1}{3}$	$\neq 1$
Σ	$= 1$	$= 1$	

Abb. 4 Tabelle der Bedingten Wahrscheinlichkeiten
unter der Bedingung Wohnort

+ Andererseits kann man das Merkmal Einkommen voraussetzen.
Aufgabe: Erstelle die zugehörige Vierfeldertafel !
Lösung: Man erhält durch Zeilennormierung:

$\nearrow\!\!\nearrow$	A	$\bar{A}$	Σ
B	$P(A/B) = \frac{450}{1050} = \frac{3}{7}$	$P(\bar{A}/B) = \frac{600}{1050} = \frac{4}{7}$	$= 1$
$\bar{B}$	$P(A/\bar{B}) = \frac{150}{450} = \frac{1}{3}$	$P(\bar{A}/\bar{B}) = \frac{300}{450} = \frac{2}{3}$	$= 1$
Σ	$\neq 1$	$\neq 1$	

Abb. 5 Tabelle der Bedingten Wahrscheinlichkeiten
unter der Bedingung Einkommen

Fachl. Anm.: Das dynamische Verständnis unserer Welt zwingt uns dazu, nicht mehr nur wie bisher verschiedene Merkmale zum gleichen Zeitpunkt zu betrachten, sondern ebensosehr das gleiche Merkmal zu verschiedenen Zeiten zu untersuchen. Damit ergibt sich aus dem Vorschlag der systematischen Berechnung bedingter Wahrscheinlichkeiten anhand von Vierfeldertafeln ein naheliegender Zugang zu MARKOFF-Ketten. Es sei dies im folgenden skizziert, obwohl die Verfolgung dieses Vorschlages erst zu einem wesentlich späteren Zeitpunkt im konkreten Unterrichtsgeschehen erfolgen sollte.

Did. Anm.: Der in Abb. 4 dargestellte Einfluß des Wohnortes auf das Einkommen, und der in Abb. 5 dargestellte Einfluß des Einkommens auf den Wohnort wird im Sinne der Optimierungsbestrebungen der Lebensbedingungen voraussichtlich zu einem Wohnortwechsel führen, der wegen $P(B/A) > P(B/\overline{A})$ in welche Richtung verlaufen wird?

Die Vierfeldertafel nimmt nun die Gestalt einer Übergangsmatrix an, etwa von der Form:

⇙	$A(t)$	$\overline{A}(t)$	Σ
$A(t+1)$	0,90	0,05	
$\overline{A}(t+1)$	0,10	0,95	
Σ	1	1	

Anhand dieser Übergangsmatrix können in gewohnter Weise die Fragen, wie sie für Markoff-Prozesse gestellt werden, einer Lösung zugeführt werden, und zwar durchaus in einer für die Unterrichtspraxis tauglichen Form. Dies ist jedoch Inhalt eines anderen Vortrages dieser Veranstaltung, und wird hier somit nicht weiter verfolgt.

Aufgaben: 1) Interpretiere die errechneten Werte !
2) Stelle nochmals die Vorgangsweise schematisch dar !

Lösung:

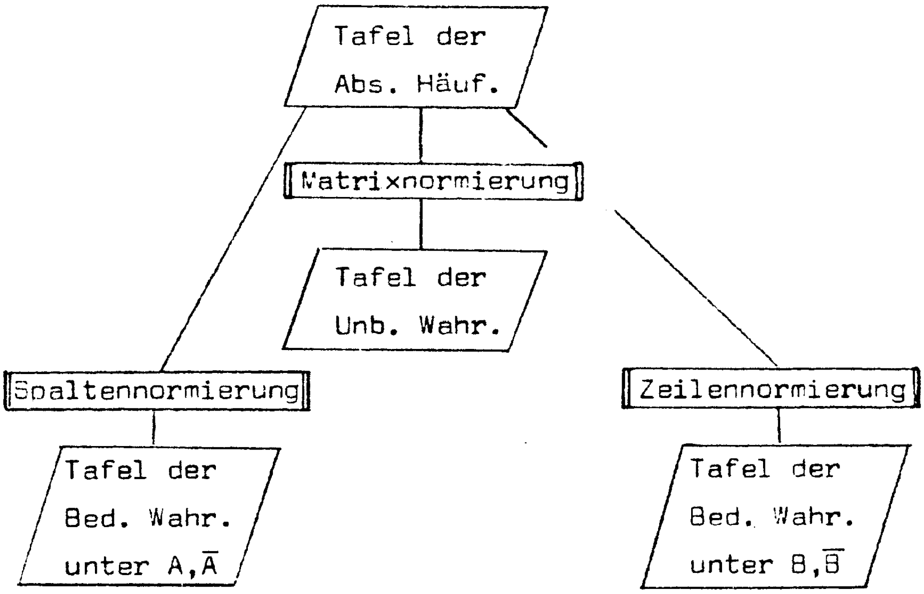

Abb. 6

<u>Beispiel 2:</u> 100 Personen wurden zum Problem "Treu-Untreu" befragt, und zwar mußten sie über sich und über ihre Einschätzung durch den Partner Auskunft geben . Man erhielt folgendes (fiktives Ergebnis):

	Treu	Untreu	Σ
Als treu eingeschätzt	20	65	
Als untreu eingeschätzt	5	10	
Σ			

Aufgabe: 1) Vervollständige die Tafel der Absoluten Häufigkeiten und erstelle die restlichen 3 Tafeln !

2) Beantworte aus dem vorliegenden Zahlenmaterial die Fragen:
   a) Wie groß ist die Wahrscheinlichkeit, treu zu sein, obgleich man als untreu gilt ?
   b) Wie groß ist die Wahrscheinlichkeit wirklich treu zu sein, wenn man als treu gilt.
   c) Vergleiche die Ergebnisse aus a) und b)

Did. Anm.: Beispiel 2) hat nicht nur Einübungscharakter. Neben dem Ziel, aus dem doch eher paradoxen Ergebnis Zündstoff für eine Diskussion über die Gültigkeit des vorgelegten Zahlenmaterials zu gewinnen, verfolgt das Beispiel auch die Absicht, den Begriff "Bedingung" näher betrachten zu müssen. Im Alltag verbindet man mit Bedingung meist die Vorstellung einer Einschränkung. Im vorliegenden Fall sieht man wegen $P(A/B) < P(A) < P(A/\bar{B})$ deutlich, daß eine Bedingung die Wahrscheinlichkeit für das Eintreten eines Ereignisses sowohl verkleinern als auch vergrößern kann.

Wie man sieht, kann man in fast naiver Weise vorgelegte Daten be-"urteilen". Erst nach dieser Phase sollte eine formale Vertiefung erfolgen. Dies geschieht hier im Einklang mit dem genetischen Prinzip. Die Phase des selbst Findens von Beziehungen ist mindestens genauso wichtig wie die des Beweisens. Vierfeldertafeln haben den richtigen Grad an Komplexität um eigenständiges Suchen erfolgreich verlaufen zu lassen, ohne die Ergebnisse als trivial abtun zu müssen. Denn zu Teil herrscht auch bei "Fachleuten" Unkenntnis über manche Zusammenhänge. So berichtet G. SCHRAGE (vgl. L 17) von einem Indizienprozeß, in welchem man auf Grund der Gutachten von Sachverständigen zu einer völlig falschen Einschätzung der Wahrscheinlichkeit für die Täterschaft des Angeklagten kam. So verwendeten die "Experten" z.B. die Beziehung $P(A/B)+P(A/\bar{B})=1$, deren Ungültigkeit aus Abb. 5 unmittelbar abgelesen werden kann. Auf solche Gegebenheiten sollte man im Unterricht hinweisen; Unterrichtsziel ist Erziehung zur Mündigkeit, nicht zur "Expertenhörigkeit". Damit ergibt sich eine natürliche Fortsetzungen unserer Intentionen aus Beispiel 2; Wurde dort das Datenmaterial auf seine Glaubwürdigkeit hin diskutiert, so in diesem zweiten Unterrichtsabschnitt die formalen Beziehungen.

G. SCHRÖPFER (vgl. L 18) hat im Vorjahr an dieser Stelle eine andere Möglichkeit des Zuganges zur Formel von BAYES etc. vorgestellt.

Lösung von 2) a) $P(A/\bar{B}) = \frac{1}{3}$

b) $P(A/B) = \frac{20}{85}$

c) Folgerung??: Die als untreu eingeschätzten Personen sind mit größerer Wahrscheinlichkeit treu als die als treu geltenden.

§ 2 FORMALE FOLGERUNGEN

Die gemäß Abb. 6 ablaufende Berechnung von bedingten und unbedingten Wahrscheinlichkeiten wirft zwei Fragen auf:

1) Welche Kontrollen gibt es, die Richtigkeit der Werte in numerischer Hinsicht prüfen zu können.
2) Nicht immer wird die Tafel der absoluten Häufigkeiten vorliegen. Wie soll man in diesem Falle vorgehen?

Genaugenommen fordern beide Fragen dasselbe, nämlich Auskunft über den inneren Zusammenhang zwischen den Tafeln untereinander:

A) Summenkriterien:

Beh.: $P(A \cap B) + P(A \cap \bar{B}) = P(A)$ laut Abb. 3

Aufgaben: 1) Finde und formuliere analoge Behauptungen (aus Abb. 3)
2) Begründe deine Vermutung verbal !

Aus der Diskussion von Aufgabe 2) sollten sich die Begriffe "unvereinbare Ereignisse" und "Gegenereignis", sowie die Probleme mit der Berechnung von Wahrscheinlichkeiten von Vereinigungsmengen entwickeln. Durch Übertragung der geläufigen Regeln aus der Mengenlehre kann man obige Begriffe und Sachverhalte fixieren. Unter diesen Voraussetzungen wird die folgende Aufgabe lösbar.

Aufgaben: 1) Versuche die obige Behauptung formal zu beweisen !

Aufgaben: 1) Versuche analoge Beziehungen in den Tafeln 4 und 5 zu finden !
2) Versuche die Regel in Worte zu fassen !

Lösung von 2) Nur die Vordergliedkomplettierung führt zum sicheren Ereignis.

B) Zusammenhang zwischen bedingten und unbedingten Wahrscheinlichkeiten

Aus den Tafeln 3 und 5 liest man unmittelbar ab: $P(A/B) = \frac{P(A \cap B)}{P(B)}$ (*)

speziell: 1) $B=E$: $P(A/E) = \frac{P(A \cap E)}{P(E)} = \frac{P(A)}{1} = P(A)$

2) $B=\emptyset$: $P(A/\emptyset) = \frac{0}{0}$. Um solche Ausdrücke zu vermeiden, formt man (*) um zum Multiplikationssatz:

$$P(A \cap B) = P(A/B) \cdot P(B) \qquad (**)$$

Aufgaben: 1) Formuliere und verifiziere (**) für $P(B/\bar{A})$!
2) Beweise mittels (*) $0 \leq P(A/B) \leq 1$ für beliebiges A,B !

Fachl. Anm.: Alle in § 2 hergeleiteten Formeln lassen sich verallgemeinern, und zwar nicht nur, was ihre Aussage betrifft, sondern ebensosehr, was ihre Herleitbarkeit anhand von Matrizen betrifft. Eine erste naheliegende Verallgemeinerung besteht darin, zu Partitionen A_i von E überzugehen.

Die in A) angegebenen Summenkriterien sind als Zeilensummenkriterium bzw. als Spaltensummenkriterium bei stochastischen Matrizen geläufig:

$$\sum_{i=1}^{n} P(A_i) = 1$$

$$P(B) = \sum_{i=1}^{n} P(A_i \cap B)$$

in Analogie zu Abb. 3, und analog zu Abb. 4 und Abb. 5 z.B.

$$\sum_{i=1}^{n} P(A_i/B) = 1 \qquad (+)$$

Der in B) angegebene Multiplikationsatz findet seine Fortsetzung in:

$$P(A_1 \cap A_2 \cap .. \cap A_n) = P(A_1).P(A_2/A_1).P(A_3/A_1 \cap A_2). \ldots . \\ .P(A_n/A_1 \cap A_2 \cap \ldots \cap A_{n-1})$$

Die in C) angegebne Formel von der Totalen Wahrscheinlichkeit lautet nun:

$$P(B) = \sum_{i=1}^{n} P(B/A_i).P(A_i)$$

Sie stellt in gewissem Sinn das Gegenstück zu (+) dar, welche Aussagen über die Vordergliedkomplettierung macht, während jetzt die Hintergliedkomplettierung auf ihre Auswirkung hin untersucht wird.

Die in D) angegebene Formel von BAYES lautet in ihrer allgemeinen Form:

$$P(A_k/B) = \frac{P(B/A_k).P(A_k)}{\sum_{i=1}^{n} P(B/A_i).P(A_i)} \qquad k=1,\ldots,n$$

C) Die Formel von der totalen Wahrscheinlichkeit

Satz: $P(B)=P(B/A).P(A) + P(B/\overline{A}).P(\overline{A})$ (***)

Bew: Anl.: Der Multiplikationssatz (**) liefert auf der rechten Seite die Wahrscheinlichkeitssumme zweier Durchschnitte. Dieser Ausdruck wurde schon in §2,A abgeleitet.

D) Die Formel von BAYES

Wir haben bereits gesehen, daß der Multiplikationssatz den Zusammenhang zwischen den Tafeln 3 und 4 bzw. zwischen den Tafeln 3 und 5 herstellt. Hier gilt es nun den Zusammenhang zwischen den Tafeln 4 und 5 aufzuzeigen. In gewissem Sinn handelt es sich um eine "Umkehraufgabe", wollen wir doch etwa aus dem bekannten $P(A/B)$ den Wert von $P(B/A)$ errechnen, und umgekehrt.

Aus dem Multiplikationssatz (**)

$$\left.\begin{array}{l} P(A\cap B)=P(A/B).P(B) \text{ und analog} \\ P(A\cap B)=P(B/A).P(A) \end{array}\right\}$$

folgt $P(A/B).P(B) = P(B/A).P(A)$

also für $P(B) \neq 0$ bereits die (spezialisierte) Formel von BAYES zu

$$P(A/B) = \frac{P(B/A).P(A)}{P(B)}$$

Meist setzt man für den Nenner gemäß (***) ein und erhält:

$$P(A/B) = \frac{P(B/A).P(A)}{P(B/A).P(A) + P(B/\overline{A}).P(\overline{A})} \qquad (****)$$

Beispiel 3: Ein Krebstest ist sehr zuverlässig. Habe ich Krebs, so ist der Test positiv mit 96% Sicherheit. Habe ich nicht Krebs, so ist der Test mit 94% Sicherheit negativ. Mein Testergebnis ist positiv. Wie groß ist die Wahrscheinlichkeit, tatsächlich Krebs zu haben, wenn 8 Promille der Bevölkerung Krebs haben ?

Lösung: Sei A ....ich habe Krebs , und $\overline{A}$ ... ich habe nicht Krebs
B ....Test positiv , und $\overline{B}$ ... Test negativ

dann lautet die formalisierte Aufgabenstellung:

Geg.: $P(A)= 0,008$, $P(B/A)=0,96$, $P(\overline{B}/\overline{A}) = 0,94$
Ges.: $P(A/B)$

Aufgrund der oben abgeleiteten Beziehungen ist es (bei jeder eindeutigen und widerspruchsfreien Angabe) möglich, die Tafeln 3,4 und 5 zu rekonstruieren. Da dies in Hinblick auf die Fragestellung unnötigen Aufwand verursacht, wollen wir die Formel von BAYES (****) heranziehen, und die Rekonstruktion der Tafeln als Aufgabe - zur Einübung der formalen Zusammenhänge - stellen.

Did. Anm.: Im Beispiel 3 erfolgt scheinbar eine Abkehr vom propagierten Weg der Verwendung von Matrizen. Dem ist jedoch nicht so. Vielmehr handelt es sich darum, den Problemlöseapparat, der auf Vierfeldertafeln beruht, dem allgemeinen Wunsch, möglichst ökonomisch vorzugehen, anzupassen. Offensichtlich ist es so, daß nicht alle Elemente aller Tafeln für die vorliegende Frage von Bedeutung sind. Folglich gilt es als nächsthöhere Stufe intelligenten Verhaltens jene bereits bekannten Beziehungen zwischen den Tafeln zu verwenden, die auf möglichst kurzem Weg genau zu den gewünschten Elementen führen. Wie die Erfahrung zeigte, hat die Einschaltung einer Stufe der systematischen Berechnung aller möglichen unbedingten und bedingten Wahrscheinlichkeiten zwischen die Stufe der Definition der bedingten Wahrscheinlichkeiten und die Stufe ihrer Anwendung in der Formel von BAYES dem Verständnis und der Lösungsfähigkeit von Aufgaben analog zu Beispiel 3 sehr gut getan. Unsere Vorgangsweise entspricht auch dem Ziel, die Lösung eines Problems auf mehreren Stufen abzuwickeln. Damit entsprechen wir bekannten psychologischen Forderungen als auch der Absicht, formales Problemlöseverhalten als Ergebnis von Ökonomiebestrebungen herauszustellen.

In L 1 und L 8 findet man eine Fülle ähnlicher Beispiele.

Im Paragraphen über formale Folgerungen wäre an und für sich Gelegenheit, rückblickend die Wahrscheinlichkeit als Maß-Funktion vorzustellen, und durch die Axiome von KOLMOGOROFF zu definieren. Wir stellen diese Präzisierungen nochmals zurück, um vorerst den Begriff der Unabhängigkeit von Ereignissen zu konstituieren. Damit bleiben wir weiterhin unserem Konzept einer Verwendung von Vierfeldertafeln treu. Die Stringenz unserer Vorgangsweise ist aber nur einer der Gründe. Bei der Herleitung des Unabhängigkeitskriteriums wird man zwangsläufig auf Begriffe wie Verteilung und Zufallsvariable stoßen, die man [illegible] "Nutz"-Begriffe hinreichend vorbereiten kann, um [illegible] einer Präzisierung zuzuführen.

Aufgabe: Vervollständige die folgenden Tafeln:

	A	$\bar{A}$	
B		↗	
$\bar{B}$			
	$\frac{8}{1000}$		

↙	A	$\bar{A}$	
B	0,96		
$\bar{B}$		0,94	

↗	A	$\bar{A}$	
B			
$\bar{B}$			

Gemäß (****) erhält man durch Einsetzen:

$$P(A/B) = \frac{0{,}96 \cdot 0{,}008}{0{,}96 \cdot 0{,}008 + P(B/\bar{A}) \cdot P(\bar{A})}$$

Über das Spaltensummenkriterium $P(B/\bar{A}) + P(\bar{B}/\bar{A}) = 1$ errechnet man für $P(B/\bar{A}) = 0{,}06$,

über $P(A) + P(\bar{A}) = 1$ den Wert von $P(\bar{A}) = 0{,}992$.

Insgesamt erhält man so für $P(A/B) \approx 0{,}11$, dh., es besteht kein Grund zur Panik.

§ 3 UNABHÄNGIGE EREIGNISSE

Beispiel 4: Ein Banküberfall. Ein Zeuge berichtet: "Der Täter hat rote Haare und außerdem Sommersprossen". Die Polizei hat bald darauf eine Person verhaftet, auf die die Beschreibung zutrifft, und beendet darauf hin ihre Suche.
Für den Dorfkiebitz ist die Sache klar:" Rote Haare kommen hier mit 5%, Sommersprossen mit 6% Wahrscheinlichkeit vor, dh., rote Haare haben nur 5% aller Sommersprossigen, mithin 3 Promille aller Personen. Daher findet man mit 3 Promille Wahrscheinlichkeit eine der Personsbeschreibung entsprechende Person, was wiederum bedeutet, daß die Wahrscheinlichkeit, keine weitere Person zu finden, mit 99,7% vorliegt. Also ist der Festgenommene mit 99,7% Wahrscheinlichkeit der Täter. Dies entspricht der statistischen Sicherheit. Er ist es, er muß es sein !"

Diskutiere die Argumentationskette des mathematisch vorgebildeten Dorfkiebitzes!

Lösung: Es wird hier von einer - auch im Alltag oft fälschlich angenommenen - Voraussetzung ausgegangen, nämlich der, zu glauben, daß die Eigenschaft, rote Haare zu haben, keinen Einfluß darauf hat, ob man Sommersprossen besitzt oder nicht. Genau genommen hat man nur folgende Daten zur Verfügung:

Did. Anm.: Die Unabhängigkeit von Ereignissen ist einer der zentralen Begriffe der Wahrscheinlichkeitstheorie überhaupt. Dies scheint unser Vorgehen zu rechtfertigen, nicht von der dürren mathematischen Definition auszugehen, sondern diese gemäß allgemein anerkannter didaktischer Grundsätze und Unterrichtsprinzipien an den Schluß der Betrachtungen zu stellen. Insbesondere gilt es folgende Punkte hervorzuheben:

1) Das stillschweigende Voraussetzen von Unabhängigkeit kann ein Urteil entscheidend beeinflussen. In Beispiel 4 wird versucht dies drastisch vor Augen zu führen. Gleichzeitig wird dem Schüler der Mangel an mathematischem Rüstzeug zur Untersuchung solcher Probleme offenbar.
2) Man sollte erkennen, daß sehr oft - wie in Beispiel 4 - die Angabe unvollständig ist, da bloß die Wahrscheinlichkeiten für das isolierte Auftreten von Merkmalsausprägungen vorliegen. Dies führt geradewegs zur Frage, durch welche Angabestücke eine Vierfeldertafel eindeutig und widerspruchsfrei gegeben ist. Für den Lehrer als Aufgabenerfinder ist diese Frage von besonderer Wichtigkeit.
3) Da in Beispiel 4 nur der statische, simultan-koinzedale Standpunkt vertreten werden kann, und so nicht alleine den didaktischen Zielvorstellungen genügt, wird in Beispiel 5 der Begriff der Unabhängigkeit in seiner dynamischen, sukzedan-koinzedalen Erscheinungsform vorgestellt. Diese Vorgangsweise entspricht unserem bisherigen Weg, der Begriffsgenese breiten Raum zu widmen.
4) Der Gepflogenheit vieler Lehrbücher, den Unabhängigkeitsbegriff anhand des Multiplikationssatzes zu definieren, wollen wir nicht folgen, sondern gerade die spezielle Form des Multiplikationssatzes als Folge unserer Auffassung von Unabhängigkeit vorstellen. Wir wollen ganz analog vorgehen wie bei der naiv-systematischen Herleitung des Wahrscheinlichkeitsbegriffes. Wiederum gehen wir von empirisch vorliegenden Daten aus,und ver-

	rote H. A	kein r.H. $\overline{A}$	
Sommersprossen B	x		60
keine Sommersp. $\overline{B}$			
	50		1000

Aufgaben:
1) Gib die zulässigen Werte für x an !
2) Vergleiche die Werte x=3 und x=45 hinsichtlich ihrer Ausagekraft für die Wahrscheinlichkeit der Täterschaft des Festgenommenen !
3) Offensichtlich hat x zentrale Bedeutung für das vorgelegte Problem. Was "bedeutet" jedoch x ?

Beispiel 5: Ein Vertreter legt einen Rechenschaftsbericht vor, um seine Leistungen in Hinblick auf eine angestrebte Gehaltserhöhung ins rechte Licht zu rücken. Dazu legt er eine Tabelle samt Interpretation vor:

	1980 A	1981 $\overline{A}$	$\sum$
Kauf B	5	15	20
Kein Kauf $\overline{B}$	35	45	80
Verkaufsgesp.	40	60	100

Erläuterungen: A ... Werbemethode 1980
$\overline{A}$ ... Werbemethode 1981

1) Ich habe um 50% mehr Verkaufsgespräche geführt.
2) Ich habe 3mal soviele Maschinen verkauft. (+ 200%)
3) Meine neue Verkaufstrategie ist 4mal so gut, weil 50% Mehraufwand 200% Mehrverkauf bewirkten !

suchen sie - nur aus einem anderen Blickwinkel - zu interpretieren. Wiederum scheint der Begriff der Abhängigkeit der natürliche(re) zu sein, aus dem heraus erst durch unser Streben nach möglichst weitgehender Vereinfachung der Unabhängigkeitsbegriff erwächst. Der systematischen Erstellung aller Werte in den Vierfeldertafeln zur Berechnung von Wahrscheinlichkeiten entspricht nun die systematische Berechnung von verschiedenen, zulässigen Vierfeldertafeln zur Darstellung möglicher Abhängigkeiten.

5) Die systematische Erstellung aller möglichen, speziell der extremen, Vierfeldertafeln bezüglich der fest vorgegebenen Randverteilungen führt schließlich weg von der polaren Sichtweise des Unabhängigkeitsbegriffes hin zu einem Maßbegriff für den Grad der Abhängigkeit, und so in gewissem Sinn zu den erkenntnistheoretischen Wurzeln des Wahrscheinlichkeitsbegriffes. Insoferne bildet die Begriffsgruppe von Wahrscheinlichkeit und Unabhängigkeit ein im Sinne der Erkenntnistheorie thematisches Ganzes. Dieser Einheit versuchten wir in diesem Referat gerecht zu werden.

Angesichts der kurzen, zur Verfügung stehenden Zeit und der thematischen Beschränkung kann nicht dargelegt werden, daß das vorgelegte Konzept auch in den folgenden Kapiteln zum Verteilungs- und Korrelationsbegriff von großem Nutzen ist. Es würde mich jedoch freuen, wenn die doch sehr komprimierte und in ihrem theoretischen Teil notwendigerweise auf Andeutungen beschränkte Darstellung Sie zur Diskussion, vielleicht sogar zum Ausprobieren, angeregt haben sollte.

Aufgaben: 1) Diskutiere die Tabelle und die angegebenen Erläuterungen unter Verwendung der Begriffe Arbeit und Leistung.

2) Berechne die Werte $P(B/A)$ und $P(B/\bar{A})$, und vergleiche die Begriffe Bedingte Wahrscheinlichkeit und Leistung !

3) Diskutiere in analoger Weise die nachfolgende Tafel :

	1980 A	1981 $\bar{A}$	$\sum$
Kauf B	5	10	15
Kein Kauf $\bar{B}$	35	70	105
Verkaufsgesp.	40	80	120

Lösung: Im ersten Fall ist $P(B/A) = \frac{1}{8}$ und $P(B/\bar{A}) = \frac{1}{4}$

dh., mißt man die Leistung des Verkäufers anhand des Anteils der Käufer an der Menge der Interessenten, so hat er seine Leistung verdoppelt. Diese Art der Leistungsmessung ist durchaus diskussionswürdig !

Im zweiten Fall ist $P(B/A)=P(B/\bar{A})=P(B/E)=P(B)= \frac{1}{8}$

dh., der Vertreter arbeitet doppelt soviel, leistet aber genauso viel (im obigen Sinn des Verhältnisses von Aufwand und Erfolg). Die 1980 und 1981 angewandten Werbe- und Verkaufsmethoden unterscheiden sich im Arbeitsaufwand, nicht im Erfolg. Die Kaufbereitschaft ist demgemäß unabhängig von der Verkaufsmethode !

Definition: Das Ereignis B ist vom Ereignis A unabhängig, wenn gilt:
$P(B/A) = P(B)$

Damit lautet der Multiplikationssatz für unabhängige Ereignisse gemäß (**): $P(A \cap B)= P(A).P(B)$

dh., für unabhängige Ereignisse benötigt man bloß die unbedingten Wahrscheinlichkeiten zur Berechnung der Wahrscheinlichkeit ihres gemeinsamen Eintretens, für abhängige Ereignisse die bedingten Wahrscheinlichkeiten !

LITERATURVERZEICHNIS:

L 1: W.ARNOLD/A.KIRSCHENHOFER
Aufgabensammlung zur Wahrscheinlichkeitsrechnung mit did.
Beiträgen I + II
Universität Linz, Inst. f. Mathematik : Hft 9/77/78, Hft 10/78/79

L 2: H.ATHEN/H.GRIESEL (Hrsg.)
Mathematik heute - Grundkurs Stochastik
Verl. Schroedel/Schöningh, 1979

L 3: H.BASLER
Grundbegriffe der Wahrscheinlichkeitsrechnung und
statistischen Methodenlehre
Verl. physika, 1974[5]

L 4: J.CIGLER
Wahrscheinlichkeitsrechnung
Skriptum zur Lehrerfortbildung, Heft 10, BMfUK

L 5: G.CLAUSS/H.EBNER
Grundlagen der Statistik
Verl. Harri Deutsch, 1975

L 6: H.DINGES
Schwierigkeiten mit der Bayesschen Regel
Mathematisch-Physikalische Semesterberichte XXV, S 113-156,1978

L 7: W.EBERL
Einführung in die Wahrscheinlichkeitsrechnung und Statistik
Vorlesungsskriptum der TU Wien, 1969

L 8: A.ENGEL
Wahrscheinlichkeitsrechnung und Statistik I, II
Verl. Klett, Bd.I, 1973, Bd.II, 1976

L 9: A. ENGEL u.a.
Zufall oder Strategie
Verl. Klett, 1974

L10: F.J.FRITZ u.a.
Stochastische Matrizen
Verl. Springer, 1979

L11: B.W.GNEDENKO/A.J.CHINTSCHIN
Elementare Einführung in die Wahrscheinlichkeitsrechnung
VEB - Deutscher Verlag für Wissenschaften, 1973.

L12: P.HARFF/M.STICKMANN
Wirtschaftsstatistik
Verl. Orac, 1977

L13: KITAIGORODSKI, A.
Unwahrscheinliches - möglich oder unmöglich?
Verl. Aulis, 1977

L14: R.LAMMERMAYER
Angewandte Beispiele zur Wahrscheinlichkeitsrechnung
ÖMG, Didaktik Reihe, Heft 5, 1980

L15: G.MAIBAUM
Wahrscheinlichkeitsrechnung
Verl. Volk und Wissen, 1975

L16: W.PESCHEK
Projektbericht AHS - Mathematik
UBW Klagenfurt, 1979

L17: G. SCHRAGE
Schwierigkeiten bei der stochastischen Modellbildung. 2 Beispiele aus der Praxis
Journal der Didaktik der Mathematik, 1/Heft 1/2, 1980

L18: G.SCHRÖPFER
Totale Wahrscheinlichkeit und Bayessches Theorem
ÖMG, Didaktik Reihe, Heft 7, 1981

L19: H.STEINBRING
Zur Entwicklung des Wahrscheinlichkeitsbegriffes: Bernoullis Theorem aus didaktischer Sicht
Journal der Didaktik der Mathematik, 1/Heft 3, 1980

L20: H.K.STRICK
Einführung in die Beurteilende Statistik
Verl. Schroedel, 1980

L21: W.A.WALLIS/H.V.ROBERTS
Methoden der Statistik
Verl. R.Haufe, 1956

L22: F.WALTER
Wider Münze und Würfel
MNU 33/5, 1980

L23: H.WINTER
Erfahrungen zur Stochastik
Didaktik der Mathematik, Jg.4/Heft 1, 1976