Europäische Hochschulschriften
Publications Universitaires Européennes
European University Studies

Reihe XX

Philosophie

Série XX Series XX
Philosophie
Philosophy

Bd./Vol. 77

PETER LANG
Frankfurt am Main · Bern

Jürgen Humburg

Grundzüge zu einem neuen Aufbau der Wahrscheinlichkeitstheorie

PETER LANG
Frankfurt am Main · Bern

CIP-Kurztitelaufnahme der Deutschen Bibliothek

Humburg, Jürgen:

Grundzüge zu einem neuen Aufbau der Wahrscheinlich=
keitstheorie / Jürgen Humburg. - Frankfurt am
Main ; Bern : Lang, 1981.
(Europäische Hochschulschriften : Reihe,20,
Philosophie ; Bd. 77)
ISBN 3-8204-5916-2
NE: Europäische Hochschulschriften / 20

ISBN 3-8204-5916-2

© Verlag Peter Lang GmbH, Frankfurt am Main 1981

Alle Rechte vorbehalten.
Nachdruck oder Vervielfältigung, auch auszugsweise, in allen Formen wie Mikrofilm, Xerographie, Mikrofiche, Mikrocard, Offset verboten.

Druck und Bindung: fotokop wilhelm weihert KG, darmstadt

INHALTSVERZEICHNIS

ZUSAMMENFASSUNG 1

BEZEICHNUNGEN 4

§1 DER BEGRIFF DER LOGISCHEN WAHRSCHEINLICHKEIT 6

(1.1) Ereignis und Ereigniskörper 6
(1.2) Die schlichte logische Wahrscheinlichkeit..... 6
(1.3) Die logische Wahrscheinlichkeit auf Grund einer Konzeption 7
(1.4) Die Maßfunktionen 10
(1.5) Die Gestalt der Maßfunktionen 10
(1.6) Die subjektivistische Deutung der Maßfunktionen als faire Wettquotienten 15
(1.7) Die Vervollständigung der komparativen Struktur..... 17
(1.8) Die quantitative logische Wahrscheinlichkeit..... 19
(1.9) Die quantitativen logischen Wahrscheinlichkeitsschranken..... 20
(1.10) Die bedingten Konzeptionen..... 20
(1.11) Die Vermeidbarkeit des Begriffs der subjektiven Wahrscheinlichkeit 22
(1.12) Inhaltliche Bemerkungen zum Begriff der Wette 25

§2 DER BEGRIFF DES EXPERIMENTS UND DER STATISTISCHEN GESAMTHEIT 27

(2.1) Experiment und experimentelles Schema 27
(2.2) Verallgemeinerung der Konzeption der Gleichartigkeit.. 28
(2.3) Die Maßfunktionen 30

§3 DAS STATISTISCHE VERHALTEN GLEICHARTIGER PHÄNOMENE 33

(3.1) Die Konvergenz der relativen Häufigkeiten 33
(3.2) Der Begriff der statistischen Wahrscheinlichkeit..... 37
(3.3) Der Additionssatz für die statistische Wahrscheinlichkeit..... 39
(3.4) Die Geschwindigkeit der Konvergenz 40
(3.5) DE FINETTI-Theorem 42
(3.6) Der Begriff der Unterkonzeption 43
(3.7) Die Konzeptionen K_p 44
(3.8) Die Konzeptionen K^* 46
(3.9) Schlußbemerkung 51

§4 DER BEGRIFF DER STATISTISCHEN UNABHÄNGIGKEIT 52

(4.1) Definiton der statistischen Unabhängigkeit auf Grund von K 52
(4.2) Äquivalente Bedingungen für die statistische Unabhängigkeit..... 53
(4.3) Minimalbedingungen für die statistische Unabhängigkeit 54

§5 Die Überprüfung von Konzeptionen 55
(5.1) Die Überprüfung der Konzeption K der Gleichartigkeit 55
(5.2) Beispiel für die Aufgabe der Konzeption der Gleichartigkeit 56
(5.3) Überprüfung der Konzeptionen K_p 58

§6 Schlußbemerkung zum Begriff der Konzeption 58

§7 Die Anwendung der Theorie auf die Quantenmechanik 60

§8 Die Anwendung der Theorie auf die Klassische Physik 62
(8.1) Charakterisierung der unmittelbaren Klassischen Situation 62
(8.2) Die Zurückführung auf die Quantenmechanik 62

§9 Vergleich mit den klassischen Systemen der Wahrscheinlichkeitstheorie 63
(9.1) Vergleich mit dem System von de Finetti 63
(9.2) Vergleich mit dem System von Carnap 64
(9.3) Vergleich mit dem System von Mises 65
(9.4) Vergleich mit dem System von Kolmogoroff 68
(9.5) Vergleich mit dem System von Richter 69
(9.6) Schlußbemerkung zum System der Wahrscheinlichkeitsrechnung von Pascal , Laplace 70

§10 Vergleich mit einigen moderneren Systemen der Wahrscheinlichkeitstheorie 71
(10.1) Vergleich mit dem System von Hacking 71
A) Ziel des Systems von Hacking 71
B) Der Begriff der joint proposition und des statistischen Datums 71
C) Likelihoodgesetz 72
D) Statistische Tests. Der Likelihood-Test 72
E) Das statistische Datum D_B der Binomialverteilung .. 73
F) Die theoretische Beziehung zwischen D_B und der Konzeption K 73
G) Die praktische Beziehung zwischen D_B und K . Der Konfidenz-Test 74
H) Die Beziehung zwischen D_B und K* 76
I) Der subjektivistische Konfidenz-Test 77
J) Der Konfidenz-Test als Likelihood-Test 80
K) Die Überprüfung von D_B und K 81

(10.2) Vergleich mit dem System von Kyburg 82
A) Das formale Sprachsystem L 82
B) Der Rationale Korpus 82
C) Die statistische Aussage 83
D) Randomness 84
E) Wahrscheinlichkeit 85
F) Kritik des Systems von Kyburg 85

(10.3) Vergleich mit dem System von Schnorr 87
A) Ziel des Systems von Schnorr 87
B) Die bisherige Theorie 87
C) Hyperzufällige Folgen 89
D) Zufällige Folgen 90
E) Das Prinzip vom ausgeschlossenen Spielsystem 91
F) Allgemeines Maß μ 92
G) Kritik am System von Schnorr 93
H) Schnorrs Kritik am System von Kolmogoroff 94
I) Zusammenhang mit unserem System 96

§11 Schlußbemerkung 99
(11.1) Die σ-Additivität von P 100
(11.2) Genauer Bezug zu anderen Wahrscheinlichkeitssystemen 100
(11.3) Untersuchungen zu nicht gleichartigen Konzeptionen .. 100
(11.4) Umkehrsatz zur statistischen Wahrscheinlichkeit..... 100
(11.5) Spezialuntersuchungen zur Theorie von Carnap 101
(11.6) Spezialuntersuchungen zur Quantenmechanik 102
(11.7) Untersuchungen zum Hume'schen Problem 102

Liste der in der Arbeit eingeführten Symbole 108

Literaturverzeichnis 112

Lebenslauf 114

ERRATA

Die auf Seite 8 befindliche Definition muß wie folgt lauten:

Definition: $A \sim_K B$ genau dann, wenn $A \underset{\sim K}{<} B$ und $B \underset{\sim K}{<} A$.

ZUSAMMENFASSUNG

Das Ziel der vorliegenden Arbeit ist die inhaltliche Klärung des Wahrscheinlichkeitsbegriffs, bzw. der verschiedenen Wahrscheinlichkeitsbegriffe. Gestützt auf einen gegenüber CARNAP modifizierten Begriff der logischen Wahrscheinlichkeit gelingt es mit Hilfe der mathematischen Resultate von DE FINETTI, den Begriff der statistischen Wahrscheinlichkeit zu definieren.

Wir gehen dabei von der für jede empirische Wissenschaft grundlegenden Vorstellung aus, daß gewisse Phänomen von gleicher Art seien, daß sich ein bestimmter Vorgang wiederholen lasse, daß gewisse Experimente "identisch" seien. Wir führen für diese Vorstellung den Begriff, Konzeption der Gleichartigkeit, ein.

Aus der Konzeption der Gleichartigkeit ergeben sich rein logisch wahrscheinlichkeitstheoretische Folgerungen. Sind nämlich die Ereignisse $E(\lambda)$ gleichartig auf Grund einer solchen Konzeption K, so gilt, daß die Durchschnitte von je n dieser Ereignisse gleichwahrscheinlich sind auf Grund der Konzeption K; in Formeln:

$E(\lambda_1)\ldots E(\lambda_n) \sim_K E(\lambda_1')\ldots E(\lambda_n')$; $\lambda_i \neq \lambda_j$, $\lambda_i' \neq \lambda_j'$ für $i \neq j$.

Auf Grund weiterer Axiome ergibt sich aus K eine komparative Wahrscheinlichkeitsstruktur, die den Ausgangspunkt unserer Betrachtungen bildet und die wir logische Wahrscheinlichkeit auf Grund von K nennen.

Wir untersuchen eine quantitative Metrisierung der komparativen Struktur, dh. σ-additive Funktionen m_K, die mit der komparativen Struktur verträglich sind; wir nennen diese Funktionen Maßfunktionen zu K.

Die Maßfunktionen lassen sich als die subjektive Wahrscheinlichkeit eines Individuums deuten, das die Konzeption K akzeptiert.

Es gilt nun: Für jede subjektive Interpretation der komparativen Struktur, dh. für jede Maßfunktion, existiert mit Maß Eins der Limes der relativen Häufigkeiten.- Bei solchen Ereignissen, für die alle Maßfunktionen übereinstimmen, sprechen wir von einer

quantitativen logischen Wahrscheinlichkeit.- Die Betrachtung der Maßfunktionen zeigt insbesondere, daß die Aussage, daß die relativen Häufigkeiten der Ereignisse $E(\lambda)$ konvergieren, ein maximales Element in der komparativen Struktur darstellt. Auf Grund dieser Betrachtung ist es zulässig, auf Grund der Konzeption K den Limes der relativen Häufigkeiten einzuführen und als statistische Wahrscheinlichkeit zu bezeichnen.

Mit Hilfe der Maßfunktionen läßt sich weiterhin die Geschwindigkeit der genannten Konvergenz abschätzen. Für jede Maßfunktion m_K zu K gilt nämlich das Analogon zur BERNOULLI'schen Ungleichung:

$m_K(|h_n - \lim_{n\to\infty} h_n| \leq \varepsilon) \geq 1 - 1/4n\varepsilon^2$; h_n bezeichnet dabei die relative Häufigkeit der $E(\lambda)$.

Weiterhin gelingt es in der Arbeit, Relationen für den Begriff der statistischen Unabhängigkeit zu gewinnen, die in Termini der komparativen Wahrscheinlichkeit gehalten sind.

Besondere Bedeutung besitzt die vorliegende Arbeit für die Quantenmechanik: Die im Bereich der Quantenmechanik gegebene Gleichartigkeit der Phänomene erklärt ihr statistisches Verhalten.

Die übliche Mathematische Statistik wird in dieser Arbeit begründet. Allerdings erscheint es auf Grund dieser neuen Theorie angemessener, neben dem Begriff der statistischen Wahrscheinlichkeit auch den Begriff der subjektiven, bzw., da wir den Begriff der subjektiven Wahrscheinlichkeit nur als Hilfsbegriff zulassen, den Begriff der logischen Wahrscheinlichkeit zu verwenden.

Bemerkenswert ist der folgende Dualsimus: Das statistische Verhalten gleichartiger Phänomene läßt sich einmal nach dem Vorbild der Klassischen Wahrscheinlichkeitsrechnung mit Hilfe einer Zahl, der statistischen Wahrscheinlichkeit, beschreiben; zum andern lassen sich aber alle Formeln für die statistische Wahrscheinlichkeit auch mit Hilfe einer Funktion, genannt statistische Wahrscheinlichkeitsfunktion, ausdrücken. Diese Funktion ist dabei definiert als Grenzwert der relativen Häufigkeiten, abhängig von dem jeweiligen Ergebnis ω .

Die statistische Wahrscheinlichkeitsfunktion ist dabei der primäre Begriff; der Begriff der statistischen Wahrscheinlichkeit wird daraus abgeleitet; er ist definiert als Wert der statistischen Wahrscheinlichkeitsfunktion für das unbekannte, wahre, tatsächlich in der Natur vorliegende Ergebnis $\hat{\omega}$.

Für das HUME'sche Problem, das induktive Schließen, scheint die vorliegende Arbeit beispielhaft eine gewisse Lösung zu bieten.

Die entwickelten Begriffe, wie Konzeption, Maßfunktion, logische Wahrscheinlichkeit, etc. erscheinen über den Begriff der Konzeption der Gleichartigkeit hinaus auch allgemeiner von Bedeutung zu sein.

BEZEICHNUNGEN

1) Wir benutzen - als Bestandteil der Umgangssprache - die folgenden logischen Zeichen:

Logische Implikation: $A \Rightarrow B$: Aus A folgt B .

Logische Äquivalenz: $A \Leftrightarrow B$: A ist äquivalent zu B .

Definierender Doppelpunkt: $a := b$: a ist definitionsgemäß gleich b .

$A :\Leftrightarrow B$: A ist defintionsgemäß äquivalent zu B .

All- und Existenzoperatoren: $\forall$, $\exists$.

2) Wir verwenden den folgenden maßtheoretischen Formalismus:

$\emptyset$ bezeichnet die leere Menge.

Ω bezeichnet die zu Grunde gelegte Ausgangsmenge, ω die diesbezügliche Variable für die Elemente von Ω .

$A \dot{+} B$ bezeichnet die Vereinigung der Mengen A , B .

$A.B$ bezeichnet den Durchschnitt von A , B .

$\overline{A}$ bezeichnet das Komplement von A bezüglich der Grundmenge Ω .

$\Sigma^{\cdot} A_i$ bezeichnet die Vereinigung der A_i .

$\Pi^{\cdot} A_i$ bezeichnet den Durchschnitt der A_i .

$A + B := A \dot{+} B$, sofern $A.B = \emptyset$.

$\Sigma A_i := \Sigma^{\cdot} A_i$, sofern $A_i.A_j = \emptyset$ für alle i , j mit $i \neq j$.

Sei $\mathcal{G}$ eine Teilmenge der Potenzmenge $\mathcal{P}(\Omega)$ von Ω ; dann bezeichne:

$K_{\mathcal{G}}$ den kleinsten Mengenkörper, der $\mathcal{G}$ enthält.

$B_{\mathcal{G}}$ die Borel'sche Erweiterung von $\mathcal{G}$, dh. den kleinsten σ-Körper, der $\mathcal{G}$ enthält.

$\Pi' A(\lambda)$ bezeichne das Kartesische Produkt der $A(\lambda)$.

$\Pi^{\times} \mathcal{K}(\lambda)$ bezeichne den Produkt-σ-Körper der σ-Körper $\mathcal{K}(\lambda)$.

Der im Allgemeinen zu Grunde gelegte Meßraum wir mit $(\Omega, \mathcal{K})$ bezeichnet.

Eine $\mathcal{K}$-meßbare Funktion (Zufallsgröße) schreiben wir in der folgenden Schriftart: $f(\omega)$, $h(\omega)$, $p(\omega)$, $K(\omega)$, etc.

Zu dem tatsächlichen, wahren , normalerweise unbekannten Element $\hat{\omega}$ aus Ω führen wir zu einer derartigen Funktion f das das folgende Symbol ein (Fettschrift) : $f := f(\hat{\omega})$.

Wir gebrauchen häufig die folgende Abkürzung:

$\{f = g\} := \{\omega : f(\omega) = g(\omega)\}$.

$d\nu = f d\mu$ bedeutet, daß ν die Radon-Nikodym-Dichte f bezüglich μ besitzt.

3) Als Symbol für die natürlichen, rationalen, bzw. reellen Zahlen gebrauchen wir die folgenden Buchstaben in Fettschrift: $\mathbf{N}$, $\mathbf{Q}$, $\mathbf{R}$.

4) Manchmal setzen wir voraus, daß sich gewisse unserer Überlegungen in einer formalen Sprache L darstellen lassen; diese Sprache soll dabei nicht weiter spezialisiert werden und beliebig reich sein; sie soll insbesondere die folgenden logische Zeichen enthalten: $\vee$, $\wedge$, als Zeichen für die Disjunktion und Konjunktion; $\rightarrow$, $\leftrightarrow$ als Zeichen für die materiale Implikation und Äquivalenz, sowie die folgenden Quantoren $\bigvee$, $\bigwedge$.

Ist A ein Satz von L , so bedeute

$\vdash A$, A ist logisch wahr in L .

5) Für die in dieser Arbeit speziell definierten Symbole folgt am Ende der Arbeit, auf Seite 108 ein weiteres Verzeichnis.

6) Buchstaben in eckigen Klammern beziehen sich auf das Literaturverzeichnis, etwa [C] .

§1 Der Begriff der logischen Wahrscheinlichkeit

(1.1) Ereignis und Ereigniskörper

Der von uns zu entwickelnde Wahrscheinlichkeitsbegriff bezieht sich auf mögliche Ereignisse, denkbare Situationen der Wirklichkeit, kurz Ereignisse genannt.
Wir bezeichnen in dieser Arbeit Ereignisse mit großen lateinischen Buchstaben: E , A , B, C ,... .
Wir setzen weiterhin voraus, daß die von uns betrachteten Ereignisse aus einem σ-Körper $\mathcal{K}$ über der Grundmenge Ω stammen. $\mathcal{K}$ heißt der unseren Betrachtungen zu Grunde gelegte Ereigniskörper. Mathematisch gesprochen, stellt $(\Omega,\mathcal{K})$ ein Meßraum dar.

(1.2) Die schlichte logische Wahrscheinlichkeit

Aus der logischen Struktur gewisser Ereignisse A , B ergibt sich die folgende Aussage:
$A \lesssim B$; in Worten: A ist aus logischen Gründen höchstens so wahrscheinlich wie B .
Diese Aussage läßt sich dabei wie folgt verdeutliche: Eine Wette auf A ist aus logischen Gründen höchstens so gut wie eine Wette auf B ; dabei soll die Wette auf A dieselben Einsätze haben wie die Wette auf B ; genaueres zum Begriff der Wette siehe (1.6) .
Beispiel: $A \lesssim A \dot{+} B$, $A.B \lesssim A$.

Diese Beispiele sind ein Spezialfall des folgenden Axioms, mit dem wir zu einer ersten präzisen Eingrenzung des bis jetzt nur intuitiv gefaßten Begriffs der logischen Wahrscheinlichkeit "$\lesssim$" kommen:

Axiom 0
$\vdash A \subseteq B \quad \rightleftharpoons \quad A \lesssim B$.

Die Begründung dieses Axioms ist sehr einfach: Wenn wir aus dem Eintritt von A auf den Eintritt von B schließen können,

so ist mit jeder Wette auf A auch die analoge Wette auf B gewonnen. Wir werden also die Wette auf A der Wette auf B nicht vorziehen.
Wir wollen hier nicht untersuchen, ob es weitere Axiome für den obigen Begriff der logischen Wahrscheinlichkeit "$\lesssim$" gibt, sondern uns dem im Folgenden entwickelten Begriff der logischen Wahrscheinlichkeit auf Grund einer Konzeption zuwenden.

(1.3) Die logische Wahrscheinlichkeit auf Grund einer Konzeption

Unter einer Konzeption K verstehen wir theoretische Vorstellungen T und Informationen I , die wir zu K zusammenfassen: $K = T \wedge I$.

Es muß dabei darauf hingewiesen werden, daß es mitunter schwerfällt, theoretische Vorstellungen und Informationen exakt auseinanderzuhalten, da in Informationen im Allgemeinen auch theoretische Begriffe eingehen. Dennoch wollen wir das Schema $K = T \wedge I$ im Sinne eines heuristischen Ansatzes beibehalten.

Wie wir im Folgenden sehen werden, ergibt sich für gewisse Ereignisse A , B aus der logischen Struktur von K , A , B die folgende Aussage:
$A \lesssim_K B$; in Worten: Auf Grund der Konzeption K ist A höchstens so wahrscheinlich wie B .

Unter der logischen Wahrscheinlichkeit "$\lesssim_K$" verstehen wir somit eine komparative Bewertung gewisser Ereignisse auf Grund einer Konzeption K ; die Aussage $A \lesssim_K B$ ergibt sich dabei allein aus der logischen Struktur von K , A , B gemäß den zu formulierenden Axiomen.
Zunächst läßt sich die folgende Erweiterung von Axiom 0 formulieren:

Axiom 1
Wenn K logisch impliziert, daß A in B enthalten ist, In Formeln $\vdash K \rightarrow A \subseteq B$, dann gilt: $A \lesssim_K B$.

Das wichtigste Axiom bei dieser Begriffsbildung ergibt sich bei einer Konzeption K , auf Grund derer gewisse Ereignisse $E(\lambda)$, $\lambda \in A$ als gleichartig erscheinen.

Der intuitive Begriff gleichartig läßt sich präzisieren: Die Merkmale, um die sich die Ereignisse $E(\lambda)$ und $E(\lambda')$ unterscheiden, sollen auf Grund der Konzeption K ohne Einfluß auf den Eintritt dieser Ereignisse sein. Eine weitere Präzisierung des Begriffs "gleichartig" ergibt sich dann, wenn man sich die Konzeption K und die Ereignisse $E(\lambda)$ in einer formalen Sprache L gegeben denkt, so daß sich die Irrelevanz der unterscheidenden Merkmale in der formalen Struktur von K und $E(\lambda)$ widerspiegelt. Der Begriff gleichartig ist dabei so zu verstehen, daß auf Grund von K auch Durchschnitte der Länge n aus den $E(\lambda)$ gleichartig sind; insbesondere soll die innere Struktur von A, etwa eine auf A gegebenene Topologie, irrelevant sein für den simultanen Eintritt von n verschiedenen Ereignissen $E(\lambda)$.
Damit ergibt sich das folgende Axiom:
Zunächst eine
Definition: $A \sim_K B \;:\Leftrightarrow\; A \sim_K B$ und $B \sim_K A$.

Axiom 2

Aus der Gleichartigkeit der Ereignisse $E(\lambda)$, $\lambda \in \Lambda$ auf Grund einer Konzeption K ergibt sich die folgende Wahrscheinlichkeitsstruktur:

$$E(\lambda_1)\ldots E(\lambda_n) \sim_K E(\lambda_1')\ldots E(\lambda_n') \quad ; \quad \lambda_i \neq \lambda_j \;,\; \lambda_i' \neq \lambda_j' \quad \text{für } i \neq j \;.$$

Ein konkretes Beispiel für eine gleichartige Ereignisgesamtheit bildet eine Folge gleichartig polarsierter Photonen, die auf ein Polarisisationsfilter trifft; $E(i)$:= i-tes Photon passiert das Filter. Nach den Vorstellungen T der Quantenmechanik spielt die zeitliche Reihenfolge der Photonen keine Rolle, was den Durchgang durch das Filter anbelangt; also gilt Axiom 2.

Ein weiteres Beispiel bildet der sukzessive Wurf eines Würfels unter der Zusatzannahme, daß sich der Würfel im Verlauf des Würfelns nicht ändert.

Ein Beispiel dagegen für eine nicht gleichartige Gesamtheit bildet der wiederholte Wurf eines Würfels unter der zusätzlichen Information, daß der Würfel eine Uhr enthält, die den Schwerpunkt des Würfels im Laufe der Zeit verändert.

Am Beispiel des sukzessiven Wurf eines Würfels ließe sich die Möglichkeit den Begriff "gleichartig" formal zu präzisieren, etwa wie folgt dartun: Wir interessieren uns für die Ereignisse $E(t)$, daß der zum Zeitpunkt t geworfene Würfel die Augenzahl 6 zeigt.
Die theoretischen Vorstellungen T wären in diesem Fall die Gesetze der Newton'schen Mechanik. Die Informationen I beständen unter anderem in der genauen Beschreibung der Bedingungen, unter denen der Würfel zum Zeitpunkt t geworfen wird; etwa, daß der Würfel in einem Würfelbecher liegt, dieser Becher mit der Hand eine halbe Minute geschüttelt wird und dann über einem Tisch mit homogener Oberfläche umgedreht wird; wir bezeichnen diese Bedingungen mit $B(t)$. $W(t)$ bedeute, daß der Würfel zum Zeitpunkt t geworfen wird. I ließe sich dann wie folgt charakterisieren: $I := \bigwedge t(W(t) \to B(t)) \wedge I_o$; I_o beschreibe die Information über den Würfel dahingehend, daß es sich etwa um einen aus normalem Holz hergestellten Würfel handle, dh. daß sich der Würfel nicht im Verlauf des Würfelns ändere, dh. daß er etwa keine Uhr enthalte.

Der nächste Schritt wäre die Darstellung von T, $B(t)$, I_o, $E(t)$ in einer formalen Sprache L. Dann müßten gewisse Prinzipien formuliert werden, bezüglich derer der Begriff "gleichartig($\{E(t)\}$)" aus T, $B(t)$, I_o zu erschließen ist. Solche Prinzipien sind etwa: Die Gesetze der Newton'schen Mechanik ändern sich nicht abhängig von t. Die Bedingungen $B(t)$ haben für jeden Wert von t dieselbe formale Gestalt. Aus diesen beiden Eigenschaften von T und $B(t)$, sowie I_o ergibt sich, daß der Paraneter t im Hinblick auf K irrelevant ist, für die Frage, ob der Durchschnitt $E(t_1)\ldots E(t_n)$ eintritt.
Es erscheint plausibel, daß sich der intuitive Begriff gleichartig, wie angedeutet, formal präzisieren läßt. Die volle, explizite, formale Präzisierung dieses Begriffs wäre sicherlich sehr umständlich und wahrscheinlich auch schwierig und liegt nicht im Bereich dieser Arbeit.
Eine formale Präzisierung des Begriffs "gleichartig" bezüglich einer nicht spezialisierten Konzeption K erscheint unmöglich.

Zu diesem Zweck müßte nämlich ein allgemeines Schema jeder jemals nur denkbaren Theorie entwickelt werden und bezüglich eines solchen Schemas Prinzipien angegeben werden, nach denen sich auf Grund eben so allgemein konzipierter Informationen I die Gleichartigkeit der $E(\lambda)$ ableiten läßt.
Für die in dieser Arbeit entwickelten Thesen ist eine weitere Präzisierung des Begriffs "gleichartig" sowieso nicht erforderlich, da diese Arbeit nur in Angabe der Kosequenzen aus Axiom 2 besteht und der Begriff gleichartig im Hinblick auf eine wissenschaftliche Konzeption K intuitiv so klar erscheint, daß die Anwendbarkeit dieses Begriffes in praktisch allen Fällen entscheidbar sein dürfte; im andern Fall sind natürlich formale Untersuchungen nicht zu umgehen.

(1.4) Die Maßfunktionen

Wie schon in (1.1) bemerkt, setzen wir voraus, daß die betrachteten Ereignisse aus einem σ-Körper $\mathcal{K}$ über Ω stammen. Durch die obigen Axiome ist die komparative Struktur "$\underset{\sim}{<}_K$" definiert für gewisse geordnete Ereignispaare (A,B) ; die Gesamtheit dieser (A,B) heiße $\mathcal{H}_K$.

Definiton

m_K heißt eine Maßfunktion zu der gegebenen komparativen Struktur "$\underset{\sim}{<}_K$" , wenn gilt:

a) m_K ist ein Maß auf $\mathcal{K}$.
b) $m_K(\Omega) = 1$.
c) $m_K(A) \leq m_K(B)$ für alle (A,B) aus $\mathcal{H}_K$.

(1.5) Die Gestalt der Maßfunktionen

Der Ereigniskörper $\mathcal{K}$ mit der Grundmenge Ω sei die Borel'sche Erweiterung der auf Grund von K gleichartigen Ereignisse $E(\lambda)$, $\lambda \in \Lambda$, dh.: $\mathcal{K} := {}^{B}\{E(\lambda) : \lambda \in \Lambda\}$.

Satz

Eine Maßfunktion m_K auf $\mathcal{K}$ ist durch die folgenden Zahlen w_n eindeutig festgelegt: $w_n := m_K(E(\lambda_1)\dots E(\lambda_n))$, alle λ_i verschieden; nach Axiom 2 hängt $m_K(E(\lambda_1)\dots E(\lambda_n))$ nur von n , nicht aber von den speziellen Parametern λ_i ab.

Beweis

Sei $\mathcal{G} := {}^{K}\{E(\lambda) : \lambda \in \Lambda\}$ der durch die $E(\lambda)$ erzeugte Mengenkörper. Dann ist $\mathcal{G}$ bekannlich die Menge aller endlichen Vereinigungen von endlichen Durchschnitten gewisser $E(\lambda)$ und gewisser $\overline{E}(\mu)$. Sei A ein beliebiges Element aus $\mathcal{G}$; dann gilt also:

$$A = \underset{\text{endlich}}{\Sigma^{\cdot}} \Big(\underset{\text{endlich}}{\Pi^{\cdot}} E(\lambda) \Big) . \Big(\underset{\text{endlich}}{\Pi^{\cdot}} \overline{E}(\mu) \Big) .$$

$\Lambda(A)$ sei die Menge aller λ , μ , die in der obigen Darstellung von A vorkommen. Dann gilt: $|\Lambda(A)| =: N < \infty$. Weiterhin gilt: $A \in {}^{K}\{E(\lambda) : \lambda \in \Lambda(A)\} =: \mathcal{G}(A)$. Der endliche Mengenkörper $\mathcal{G}(A)$ hat bekanntlich die Atome

$$E^{\alpha_1}(\sigma_1)\dots E^{\alpha_N}(\sigma_N) \quad \text{mit} \quad \{\sigma_1,\dots,\sigma_N\} = \Lambda(A) \quad \text{und} \quad E^{\alpha}(\lambda) := \begin{cases} E(\lambda) & \alpha=0 \\ \overline{E}(\lambda) & \text{für } \alpha=1 \end{cases}$$

A stellt dann die disjunkte Vereinigung gewisser Atome von $\mathcal{G}(A)$ dar:

$$A = \underset{\text{gewisse } (\alpha_1,\dots,\alpha_N)}{\Sigma} E^{\alpha_1}(\sigma_1)\dots E^{\alpha_N}(\sigma_N) .$$

Es gilt nun (siehe [F]) :

$$m_K(E(\lambda_1)\dots E(\lambda_n).\overline{E}(\mu_1)\dots\overline{E}(\mu_r)) = \underset{k \geq 0}{\Sigma} \binom{r}{k}(-1)^k \cdot w_{n+k} =: w_{n,r} \quad , \quad w_o := 1 .$$

Daraus folgt, daß sich $m_K(A)$ als Summe gewisser $w_{n,r}$ darstellen läßt.
Damit ist m_K auf $\mathcal{G}$ durch die w_n eindeutig festgelegt und somit auch auf der Borelschen Erweiterung $\mathcal{K}$ von $\mathcal{G}$.

Nicht jedes System von Zahlen w_n führt jedoch zu einem Maß m_K auf $\mathcal{K}$. So müssen die w_n etwa die folgende Bedingung erfüllen: $0 \leq \underset{k \geq 0}{\Sigma} \binom{r}{k}(-1)^k w_{n+k} \leq 1$.

Eine durchsichtige notwendige und hinreichende Bedingung dafür, daß die w_n ein Maß m_K definieren, liefert das folgende Theorem:

DE FINETTI - Theorem 1

Voraussetzung: $|\Lambda| \geq \aleph_0$.

a) Erfüllt ein normiertes Maß m auf dem durch die Ereignisse $E(\lambda)$, $\lambda \in \Lambda$, erzeugten σ-Körper $\mathcal{K}$ die Bedingung $m(E(\lambda_1)\ldots E(\lambda_n)) = w_n$ für $\lambda_i \neq \lambda_j$ bei $i \neq j$, so gibt es ein normiertes Maß ϕ auf dem σ-Körper der Borelschen Mengen über dem Einheitsintervall mit der Eigenschaft $w_n = \int x^n d\phi$. Für ein allgemeines Ereignis $A \in \mathcal{K}$ gilt dabei $m(A) = \int p_x(A)\, d\phi$; dabei ist $p_x(A)$ die nach der üblichen Wahrscheinlichkeitsrechnung gebildete "Wahrscheinlichkeit" von A unter der Voraussetzung, daß die Ereignisse $E(\lambda)$ unabhängig sind mit der festen "Wahrscheinlichkeit" $x \in [0,1]$.

b) Stellt umgekehrt ϕ ein normiertes Maß auf den Borelschen Mengen des Einheitsintervalls dar, so führt die Definition $m(A) := \int p_x(A)\, d\phi$ zu einer Maßfunktion auf $\mathcal{K}$, die die Bedingungen von a) erfüllt.

Das obige Theorem stellt eine Erweiterung der Resultate von DE FINETTI dar, die wie folgt lauten (siehe [F]):

DE FINETTI - Theorem 1'

Voraussetzung: $|\Lambda| = \aleph_0$.

a) Erfüllt ein normierter Inhalt m auf dem durch die Ereignisse $E(\lambda)$, $\lambda \in \Lambda$, erzeugten Mengenkörper $\mathcal{G}$ die Bedingung $m(E(\lambda_1)\ldots E(\lambda_n)) = w_n$ für $\lambda_i \neq \lambda_j$ bei $i \neq j$, so gibt es ein normiertes Maß ϕ auf den Borelschen Mengen über dem Einheitsintervall mit der Eigenschaft $m(A) = \int p_x(A)\, d\phi$; dabei ist p_x definiert wie oben.

b) Stellt umgekehrt Φ ein normiertes Maß auf den Borelschen Mengen des Einheitsintervalls dar, so führt die Definiton $m(A) := \int p_x(A)\, d\phi$ zu einem Inhalt auf $\mathcal{G}$, der die Bedingung von a) für die Ereignisse aus $\mathcal{G}$ erfüllt.

Bemerkung

Die Voraussetzung, $|\Lambda| \geq \aleph_0$, im DE FINETTI-Teorem ist notwendig; ein Gegenbeispiel für endliches Λ findet sich in [H] .

Beweis des DE FINETTI-Theorems 1

1) Der Beweis von 1' findet sich in [F] und wird hier nicht wiederholt. Wir setzen also Theorem 1' im Weiteren als bewiesen voraus.

2) Wir beweisen zunächst Theorem 1 unter der Voraussetzung $|\Lambda| = \aleph_0$.

α) Um Teil a) der Behauptung zu zeigen, muß bewiesen werden, daß die Formel $m(A) = \int p_x(A)\, d\phi$ auch für die Ereignisse $A \in \mathcal{K}\text{-}\mathcal{G}$ gilt. Dieser Beweis findet sich in [R''] Seite 38-47 und wird hier nicht wiederholt. Es sei jedoch soviel bemerkt, daß sich dieser Beweis mit Hilfe von Dynkin-Systemen erbringen läßt. Dabei muß zunächst gezeigt werden, daß die Funktion $p_x(A)$ bei festem A eine Baire'sche Funktion von x ist, sonst ist das obige Integral nicht definiert.

β) Ist bewiesen, daß $p_x(A)$ eine Baire'sche Funktion ist, so ist sehr einfach einzusehen, daß die Definiton $m(A) := \int p_x(A)\, d\phi$ zu einem Maß auf $\mathcal{K}$ führt, das Teil b) der Behauptung erfüllt. Es gilt nämlich:

(1) Da $p_x(A) \geq 0$, $m(A) = 0$.

(2) Da $p_x(\Omega) = 1$, $m(\Omega) = 1$.

(3) Da $p_x(\Sigma E_i) = \Sigma p_x(E_i) \leq 1$, ist nach dem Satz von der majorisierten Konvergenz die Summenbildung mit dem Integral vertauschbar und folglich $m(\Sigma E_i) = \Sigma m(E_i)$. Die Bedingung der Gleichartigkeit ist trivialerweise erfüllt. Damit ist Teil b) der Behauptung für abzählbares Λ gezeigt.

3) Sei $|\Lambda| > \aleph_0$.

α) Beweis von Teil a) der Behauptung: Sei m ein Maß auf $\mathcal{K}$, das die Voraussetzungen von Theorem 1 erfüllt. Λ' sei eine abzählbare Teilmenge von Λ . $\mathcal{K}(\Lambda') := {}^{B}\{E(\lambda) : \lambda \in \Lambda'\}$.

$m'|\mathcal{K}(\Lambda')$ sei die Beschränkung von $m|\mathcal{K}$ auf $\mathcal{K}(\Lambda')$. Dann gilt nach Beweisteil 2) : Es gibt ein ϕ :
$w_n = m(E(\lambda_1')\ldots E(\lambda_n')) = m'(E(\lambda_1')\ldots E(\lambda_n')) = \int x^n \, d\phi$, $\lambda_i' \in \Lambda'$;
sowie $m'(A) = \int p_x(A) \, d\phi$ für $A \in \mathcal{K}(\Lambda')$. Dabei kann vorausgesetzt werden, daß für eine andere abzählbare Menge $\Lambda^* \subseteq \Lambda$ dasselbe Maß ϕ die Gleichung $m^*(A) = \int p_x(A) \, d\phi$ für $A \in \mathcal{K}(\Lambda^*)$ erfüllt, da die Zahlen w_n für m' , m^* übereinstimmen. Es bleibt zu zeigen, daß es zu jedem $A \in \mathcal{K}$ ein Λ' gibt mit $A \in \mathcal{K}(\Lambda')$. Es muß also gezeigt werden, daß gilt:

Behauptung: $\mathcal{K} \subseteq \sum_{\Lambda' \subseteq \Lambda, |\Lambda'| = \aleph_0}^{*} \mathcal{K}(\Lambda') =: \mathcal{K}'$.

Beweis:
(1) $\mathcal{K}' \supseteq \mathcal{G}$. Sei nämlich $A \in \mathcal{G}$, so gilt:
$A = \underset{\text{endlich}}{\Sigma^*} \; \underset{\text{endlich}}{\Pi^*} \; E(\lambda) \; \Pi^* \; \overline{E}(\lambda)$. Daraus folgt: A liegt in einem geeignet gewählten $\mathcal{K}(\Lambda')$ mit abzählbarem Λ' .
(2) $\mathcal{K}'$ ist ein σ-Körper, denn:
(i) $A \in \mathcal{K}' \rightarrow \exists \Lambda' : A \in \mathcal{K}(\Lambda') \rightarrow \overline{A} \in \mathcal{K}(\Lambda') \rightarrow \overline{A} \in \mathcal{K}'$.
(ii) $A_\nu \in \mathcal{K}' \rightarrow \exists \Lambda_\nu \subseteq \Lambda : |\Lambda_\nu| = \aleph_0$, $A_\nu \in \mathcal{K}(\lambda_\nu)$
$\rightarrow A_\nu \in \mathcal{K}(\Sigma^* \Lambda_j)$ für alle ν
$\rightarrow \Sigma^* A_\nu \in \mathcal{K}(\Sigma^* \Lambda_j)$, $|\Sigma^* \Lambda_j| = \aleph_0 \rightarrow \Sigma^* A_j \in \mathcal{K}'$.

Damit ist Teil a) von Theorem 1 bewiesen.

β) Beweis von Teil b) von Theorem 1 : Wie eben gezeigt, ist jedes $A \in \mathcal{K}$ in einem geeigneten $\mathcal{K}(\Lambda')$ enthalten und folglich ist nach Beweisteil 2α) die Funktion $p_x(A)$ Baire'sch. Alles Weitere ergibt sich wie im Beweisteil 2β).- Damit ist Theorem 1 bewiesen.

Die Schar der Maßfunktionen auf $\mathcal{K}$ ist also gegeben durch die Menge aller normierten Maße über dem Einheitsintervall.

Bemerkung: Man könnte zunächst meinen, das DE FINETTI - Theorem 1 stelle insofern eine schwächere Aussage als das Theorem 1' dar, als in Theorem 1 die stärkere Voraussetzung gemacht wird, daß m ein Maß auf $\mathcal{K}$ ist, während in Theorem 1' nur verlangt wird, daß m ein Inhalt auf $\mathcal{G}$ ist. Dies ist jedoch nicht der Fall, da sich jeder Inhalt, der die Voraussetzungen von Theorem 1' erfüllt, zu einem Maß auf $\mathcal{K}$ erweitern läßt. Nach Theorem 1' hat nämlich m die Gestalt: $m(A) = \int p_x(A) \, d\phi$; nach Theorem 1 führt diese Formel zu einem Maß auf $\mathcal{K}$ und damit ist gezeigt, daß sich m zu einem Maß erweitern läßt.

(1.6) Die subjektivistische Deutung der Maßfunktionen als faire Wettquotienten

Wir betrachten ein Individuum J , das die Konzeption K akzeptiert und "Wetten" auf Ereignisse abschließt gemäß der folgenden

Definition

Eine Wette W auf das Ereignis A ist eine Abmachung zwischen J und einer "Bank", die besagt, daß J beim Eintritt von A den Geldbetrag $\overline{s} \geq 0$ von der Bank erhält, während es beim Eintritt von $\overline{A}$ den Geldbetrag $s \geq 0$ an die Bank bezahlen muß. s , $\overline{s}$ heißen Einsätze, $g := s + \overline{s}$ heißt Gesamteinsatz; es wird stets vorausgesetzt $g > 0$. $q := s/(s + \overline{s})$ heißt Wettquotient bei dieser Wette.

Wir nehmen an, daß J mehrere Wetten $W_1, \ldots, W_r$ dieser Art eingehen kann. $(W_1, \ldots, W_r)$ nennen wir ein Wettsystem; $r = 1,2,3,\ldots$ oder $r = \infty$. Einzelwetten zählen als Wettsysteme mit $r = 1$.
Wir nehmen im Folgenden an, daß J für gewisse Wettsysteme anzugeben weiß, ob es ein System von Wetten $(W_1, \ldots, W_r)$ einem anderen System $(W_1', \ldots, W_k')$ vorzieht oder nicht, dh. ob es lieber simultan die Wetten $W_1, \ldots, W_r$ oder lieber simultan die Wetten $W_1', \ldots, W_k'$ abschließt, in Formeln: $(W_1, \ldots, W_r) \blacktriangleleft (W_1', \ldots, W_k')$; gesprochen: $(W_1', \ldots, W_k')$ wird $(W_1, \ldots, W_r)$ vorgezogen.

Weiterhin kann es möglich sein, daß J zwie Wettsysteme $\mathcal{W}$ und $\mathcal{W}'$ für gleich gut erachtet, in Formeln: $\mathcal{W} \equiv \mathcal{W}'$.

Wir sprechen dann von einer Präferenzstruktur von J für Wettsysteme. Diese Präferenzstruktur soll dabei einer Reihe von Forderungen genügen, die in [H'] dargestellt sind, derart, daß wir von einer rationalen Präferenzstruktur des Individuums J sprechen können.

Wir benötigen die folgende

Definition

$\overline{W}$ heißt inverse Wette zu W, wenn J beim Eintritt von $\overline{A}$ den Geldbetrag s erhält, während es beim Eintritt von A den Geldbetrag $\overline{s}$ bezahlen muß. Die inverse Wette $\overline{W}$ entsteht also aus W dadurch, daß die Rollen von Spieler und Bank vertauscht werden. $\overline{W}$ hat den Wettquotienten $\overline{q} := 1 - q$.

Definition

J betrachtet eine Wette W als fair, wenn gilt: $W \equiv \overline{W}$.
Der Wettquotient zu einer fairen Wette heißt fairer Wettquotient.

Auf Grund der in [H'] dargestellten einfachen Axiome über die Präferenzstruktur gelten nun die folgenden Sätze:

Satz 1

Zu einem beliebigen Ereignis A gibt es genau einen fairen Wettquotienten q(A) , abhängig von J .

Satz 2

$0 = q(\emptyset) \leq q(A) \leq q(\Omega) = 1$.

Satz 3

$q(\Sigma A_\nu) = \Sigma q(A_\nu)$; $\nu \in \mathbb{N}$.

Der durch die rationale Präferenzstruktur von J eindeutig bestimmte faire Wettquotient q(A) stellt ein Maß für die subjektive Wahrscheinlichkeit dar, die J dem Ereignis A zumißt.

Zwischen den Relationen $\lesssim_K$, $\trianglelefteq$, $\equiv$ formulieren wir den bereits in (1.2) angedeuteten Zusammenhang:
Zunächst eine

Definition

$W \trianglelefteq W' \quad :\Leftrightarrow \quad W \trianglelefteq W'$ oder $W \equiv W'$.

Prinzip P

$A \lesssim_K A' \Rightarrow$ { W sei eine Wette auf A , W' eine Wette auf A' , W und W' besitzen dieselben Einsätze s , $\overline{s}$. Dann gilt: Für jedes rationale Individuum J , das die Konzeption K akzeptiert, ist bei beliebigen Einsätzen s , $\overline{s}$: $W \trianglelefteq W'$.

Satz 4

W , W' , A , A' , s , $\overline{s}$ seien definiert wie im Prinzip P . Dann gilt für den fairen Wettquotienten q(A) eines rationalen Individuums J :
$W \trianglelefteq W'$ bei beliebigen Einsätzen s , $\overline{s} \Rightarrow \quad q(A) \leq q(A')$.

Beweis

Zum Beweis von Satz 4 benötigen wir die Theorie der Utility von NEUMANN, MORGENSTERN (siehe [N]) gemäß der Bemerkung am Schluß von [H'] . Danach gilt nämlich:

$W \trianglelefteq W' \Leftrightarrow \quad U(\overline{s})q + (1-q)U(-s) \leq U(\overline{s})q' + (1-q')U(-s)$,

$q := q(A)$, $q' := q(A')$

$\Leftrightarrow \quad (U(\overline{s}) - U(-s))(q - q') \leq 0 \Leftrightarrow \quad q \leq q'$, da $U(\overline{s}) > U(-s)$ für $g = s+\overline{s} > 0$;die Utility eines Gewinns ist stets größer als die Utility eines Verlusts, vorausgesetzt, daß nicht beide, Gewinn und Verlust,gleich Null sind.

Aus Prinzip P und Satz 4 ergibt sich für jedes Individuum, das die Konzeption K der Gleichartigkeit akzeptiert:
$q(E(\lambda_1)\ldots E(\lambda_n)) = q(E(\lambda_1')\ldots E(\lambda_n'))$.

Auf Grund der vorangegangenen Sätze lassen sich die Maßfunktionen m_K demnach als die fairen Wettquotienten,- die subjektiven Wahrscheinlichkeiten - , eines Individuums deuten, das die Konzeption K akzeptiert.

(1.7) Die Vervollständigung der komparativen Struktur

Aus Prinzip P und Satz 4 des vorangegangenen Abschnitts folgt:
Sei q_J der faire Wettquotient des Individuums J ; dann gilt:
$A \lesssim_K A' \Leftrightarrow \quad q_J(A) \leq q_J(A')$ für jedes rationale Individuum J , das die Konzeption K akzeptiert.

Das führt zu dem folgenden Axiom:

Axiom 3
Gilt für alle Maßfunktionen der Konzeption K : $m_K(E) \leq m_K(E')$, so sei $E \lesssim_K E'$.

Bemerkung: Wegen (1.4c) kann diese Erweiterung der komparativen Struktur, $\lesssim_K$, zu keinem Widerspruch führen.

In der Wahrscheinlichkeitstheorie sind zusätzliche Axiome für eine komparative Wahrscheinlichkeitsstruktur bekannt. Diese Axiome ergeben sich hier als triviale Folgerungen aus Axiom 3 :

Satz 1
Die Struktur, $\lesssim_K$, ist transitiv.

Satz 2
Gilt für alle $i \in N$: $A_i \lesssim_K B_i$, dann ist $\Sigma A_i \lesssim_K \Sigma B_i$.

Zu der in dieser Arbeit ungewöhnlichen Beziehung zwischen komparativer Struktur $\lesssim_K$ und metrisierenden Funktionen m_K sind noch einige Bemerkungen nötig. Üblicherweise sieht diese Beziehung so aus, daß man von einer komparativen Struktur ausgeht und von dieser Struktur Axiom 1 und die in den Sätzen 1 , 2 angegebenen Eigenschaften als weitere Axiome verlangt. Man zeigt dann, daß sich diese Struktur durch σ-additive Funktionen metrisieren läßt.
Hier gehen wir zunächst von gewissen komparativen Grundaussagen $\mathcal{H}_K$ aus und metrisieren diese Struktur $\mathcal{H}_K$ durch σ-additive Funktionen. Dann erweitern wir diese Struktur gemäß Axiom 3 . Da wir die σ-Additivtät der metrisierenden Funktionen dadurch begründet haben, daß sich die Maßfunktionen als faire Wettquotienten interpretieren lassen, ist Axiom 3 gemäß Prinzip P und Satz 4 von (1.6) gut begründet und die in den Sätzen 1 , 2 ausgesprochenen Eigenschaften der komparativen Struktur sind zwingend. Hätten wir wie üblich, von Anfang an Satz 2 als Axiom aufgestellt, so wären wir in der schwierigen Situation, dieses Axiom zu motivieren, das sich hier ganz natürlich als Folge von Axiom 3 ergibt.

(1.8) Die quantitative logische Wahrscheinlichkeit

Sei A ein Ereignis mit der Eigenschaft, daß bei gegebener Konzeption K jede Maßfunktion zu dieser Konzeption für das Ereignis A denselben Wert besitzt. Dann führen wir für das Ereignis A die folgende Bezeichnung ein:
$l_K(A) := m_K(A)$; $l_K(A)$ heißt die quantitative logische Wahrscheinlichkeit von A bei der Konzeption K .

Die Funktion $l_K(A)$ ist also nur für spezielle Ereignisse definiert; ein Beispiel außer den trivialen Beispielen $A=\emptyset$ und $A=\Omega$ werden wir in (3.1) kennenlernen.
Weitere Beispiele ergeben sich in der CARNAP'schen Theorie der logischen Wahrscheinlichkeit (siehe [C]), wo die Familie der zulässigen Maßfunktionenen erheblich eingeschränkt wird.

Man verifiziert leicht, daß der Definitionsbereich von l_K ein σ-Körper $\mathcal{L}_K \subseteq K$ darstellt und l_K ein Maß auf $\mathcal{L}_K$ ist; genauer: $(\Omega, \mathcal{L}_K, l_K)$ ist ein Wahrscheinlichkeitsfeld im üblichen mathematischen Sinne.

Die Definition der quantitativen logischen Wahrscheinlichkeit entspricht dem Vorbild der deduktiven Logik: Ein Satz heißt logisch wahr, wenn er bei jeder Interpretation den Wert,wahr, erhält. Entsprechend hat A die quantitative logische Wahrscheinlichkeit $l_K(A)$, wenn bei jeder subjektiven Wahrscheinlichkeitsbewertung das Ereignis A denselben Wert $l_K(A)$ erhält.

Aus den Überlegungen von (1.6) ergibt sich das folgende

Anwendungsprinzip 1
Das Ereignis A besitze eine quantitative logische Wahrscheinlichkeit $l_K(A)$. W sei eine Wette auf A . Dann stellt $l_K(A)$ für jedes die Konzeption K akzeptierendes Individuum den Wettquotienten dar, bei dem die Wette W auf Grund von K fair ist.

(1.9) Die quantitativen logischen Wahrscheinlichkeitsschranken

In (3.4) werden wir auf ein Ereignis A stoßen mit der Eigenschaft, daß für jede Maßfunktion m_K zu K gilt: $m_K(A) \geq s > 0$, obwohl die Werte $m_K(A)$ alle verschieden sind und somit eine quantitative logische Wahrscheinlichkeit für A nicht existiert. Gemäß den Betrachtungen von (1.6) läßt sich diese Situation wie folgt deuten:

Anwendungsprinzip 2

Es sei für jede Maßfunktion m_K zu K : $s_1 \leq m_K(A) \leq s_2$.

W sei eine Wette auf A . Dann gilt: Ist der Wettquotient von W kleiner als s_1 , so ist die Wette W der Wette $\overline{W}$ vorzuziehen. Ist der Wettquotient größer als s_2 , so ist umgekehrt die Wette $\overline{W}$ der Wette W vorzuziehen.
Die Schranken s_i heißen quantitative logische Wahrscheinlichkeitsschranken.

(1.10) Die bedingten Konzeptionen

Zu der Konzeption K trete die neue Information $\hat{A}$ hinzu, die besagt, daß das Ereignis $A \in \mathcal{K}$ eingetreten ist.
Dann wollen wir zunächst voraussetzen, daß wir auf Grund der Information $\hat{A}$ zu der Konzeption $K' := K \wedge \hat{A}$ übergehen.

Wir werden später in (5.2) den Fall kennenlernen, daß wir durch die Information $\hat{A}$ die Konzeption K für erschüttert halten und zu einer gänzlich neuen Konzeption K_{neu} übergehen. In den Überlegungen dieses Abschnitts sei aber angenommen, daß die Konzeption K' in ihren theoretischen Vorstellungen mit K übereinstimmt und sich von K nur durch die zusätzliche Information $\hat{A}$ unterscheidet.

Multiplikationsprinzip

$q(E)$ sei der faire Wettquotient des Individuums J bei der Konzeption K , $q'(E)$ sei der faire Wettquotient von J bei $K' := K \wedge \hat{A}$. Dann gilt für beliebiges $E \in \mathcal{K}$:
$q'(E).q(A) = q(E.A)$.

Zum Beweis dieses Prinzips siehe [H'] .

Entgegen unserem ursprünglichen Vorgehen charakterisieren wir die wahrscheinlichkeitstheoretischen Folgerungen aus K' nicht durch die aus K' folgende komparative Struktur, sondern durch die auf Grund von K' zugelassenen Maßfunktionen $m_{K'}$:

Axiom 4

Die Konzeption $K' = K \wedge \hat{A}$ besitzt die folgenden Maßfunktionen:

$$m_{K'}(E) = m_K(E.A)/m_K(A) \quad .$$

Maßfunktionen mit der Eigenschaft $m_K(A) = 0$ liefern beim Übergang zu der Konzeption K' keinen Beitrag.

Auf Grund der Konzeption K' spielen nämlich nur noch Ereignisse $E \subseteq A$ eine Rolle; eine Maßfunktion m_K , so daß $m_K(A) = 0$, besagt nichts über Ereignisse $E \subseteq A$, da für all diese Ereignisse unterschiedslos gilt: $m_K(E) = 0$. Daher liefern derartige Maßfunktionen beim Übergang zu K' keinen Beitrag.

Die aus K' folgende komparative Struktur ist durch Axiom 3 festgelegt:

Satz

Sei $K' = K \wedge \hat{A}$; dann ergibt sich die komparative Wahrscheinlichkeitsstruktur $\lesssim_{K'}$, wie folgt aus der Struktur $\lesssim_K$:

$$E_1 \lesssim_{K'} E_2 \quad :\models \quad E_1.A \lesssim_K E_2.A \quad .$$

Beweis

$E_1 \lesssim_{K'} E_2 \;:\models\; m_{K'}(E_1) \leq m_{K'}(E_2)$ für alle Maßfunktionen der Konzeption K' (gemäß Axiom 3)
$:\models\; m_K(E_1.A)/m_K(A) \leq m_K(E_2.A)/m_K(A)$ für alle Maßfunktionen der Konzeption K mit $m_K(A) \neq 0$.
$:\models\; m_K(E_1.A) \leq m_K(E_2.A)$ für alle Maßfunktionen der Konzeption K.
$:\models\; E_1.A \lesssim_K E_2.A$ (gemäß Axiom 3) .

(1.11) Die Vermeidbarkeit des Begriffs der subjektiven Wahrscheinlichkeit

a) (1.6) führt an, daß sich die Maßfunktionen als subjektive Wahrscheinlichkeit eines Individuums J, das die Konzeption K akzeptiert, deuten lassen. Aus der speziellen Deutung der subjektiven Wahrscheinlichkeit als faire Wettquotienten, ließ sich gemäß [H'] ableiten, daß die Maßfunktionen σ-additiv und multiplikativ sind.
Die inhaltliche Deutung der Maßfunktionen als subjektive Wahrscheinlichkeiten, bzw. als faire Wettquotienten eines Individuums J wurde benutzt, um die Anwendungsprinzipien 1, 2 und die Axiome 3, 4 zu erlangen.
Allerdings ergibt sich daraus die Konsequenz:
Der objektive Charakter der gewonnenen Aussagen ist im Sinn von intersubjektiv zu sehen. Dieses Ergebnis erscheint genauso inadäquat, wie wenn man die Gültigkeit der quantorenlogischen Theoreme mit dem Prädikat intersubjektiv versehen würde.

b) Außerdem wird hier die Ansicht vertreten, daß der Begriff der subjektiven Wahrscheinlichkeit zwar eine sehr hilfreiche Vorstellung, aber letzten Endes höchst fragwürdig ist.
Wir wollen zunächst mehrere Deutungen des Begriffs der subjektiven Wahrscheinlichkeit q_J eines Individuums J geben:

(1) Es gibt ein reales Individuum J, aus dessen empirisch getesteten Wettverhaltens sich die Funktion q_J ergibt.

Die Interpretation (1) führt also in den Bereich der empirischen Psychologie und scheint völlig belanglos für die Gültigkeit der vorliegenden Ergebnisse zu sein.

(2) Man konstruiert zu einer vorgegebenen Funktion q einen Roboter, der Wetten gemäß dieser Funktion abschließt.

Auch dieser Standpunkt, daß man einen derartigen Roboter konstruieren könnte, scheint nicht sehr belangvoll für die vor liegenden Fragen zu sein.

(3) Man fingiert ein Individuum J, das über die Konzeption K_J verfügt, aus der sich eindeutig die Maßfunktion $m_{K_J} = q_J$ ergibt.

In dieser Version wäre der Begriff der subjektiven Wahrscheinlichkeit als eigenständiger Begriff aufgehoben und fiele mit dem CARNAP'schen Begriff der logischen Wahrscheinlichkeit zusammen, was sich etwa wie folgt formalisieren ließe: $c(E,K_J) = q_J(E)$.

Allerdings ist dies Deutung als äußerst utopisch anzusehen und ebenfalls nicht von uns intendiert. Es ist CARNAP nämlich nicht gelungen, selbst für das einfache Modell der m_λ-Funktionen, Voraussetzungen zu finden, aus denen sich der Wert von λ eindeutig festlegen ließe; siehe [C].

Der Begriff der subjektiven Wahrscheinlichkeit erweist sich also unter den verschiedensten Gesichtspunkten als entweder unerheblich für unsere Betrachtungen oder als äußerst problematisch. Außerdem,- dies ist der entscheidende Gesichtspunkt -, wird hier die Ansicht vertreten, daß der Begriff der subjektiven Wahrscheinlichkeit, auch wenn er ein hilfreiches Bild darstellt, doch grundsätzlich vermieden werden kann.

b) Die Vermeidung des Begriffs der subjektiven Wahrscheinlichkeit läßt sich etwa wie folgt dartun:
Ein Wissenschaftler J_o betrachte die Konzeption K und die voraussetzungsgemäß daraus folgende Menge $\mathcal{A}_K$ gewisser komparativer Aussagen $A \lesssim_K B$. Wie in (1.2) dargestellt, bedeuten diese Aussagen Anleitungen, wie gewisse Entscheidungen zwischen Wetten auf Grund von K zu treffen sind. J_o frage sich nun, ob sich aus K auch gewisse quantitative Wettanleitungen folgern lassen.
J_o ziehe dazu irgend ein System von fairen Wettquotienten auf $\mathcal{A}$ in Betracht, das mit K verträglich und außerdem in sich konsistent, dh. den Forderungen von [H'] genügen muß und folglich σ-additiv und multiplikativ sein muß.
J_o ziehe in dieser Weise alle Maßfunktionen m_K in Betracht.

J_o erkennt schließlich, daß alle Maßfunktionen auf $\mathcal{L}_K$ übereinstimmen und demgemäß l_K auf $\mathcal{L}_K$ ein sich eindeutig und notwendig auf Grund von K ergebendes System fairer Wettquotienten ist; er gelangt damit zu dem Anwendungsprinzip 1 .
In ähnlicher Weise läßt sich die Ableitung der Axiome 3 , 4 und des Anwendungsprinzips 2 darlegen.
Der Kommentar, daß die gewonnenen Aussagen objektiv im Sinn von intersubjektiv sind, erscheint dabei überflüssig, da das Wort subjektiv gar nicht mehr vorkommt.

Anstelle der zwar hilfreichen,aber grundsätzlich zu entbehrenden Formulierung, die Maßfunktionen seien subjektive Wahrscheinlichkeiten, empfehlen sich die folgenden Verbalisierungen:

Die Maßfunktionen m_K sind die auf Grund von K möglichen, in sich konsistenten Systeme von fairen Wettquotienten.

l_K ist das auf Grund von K notwendige System von fairen Wettquotienten.
Nach diesen Ausführungen wird auch klar, warum wir von Anfang an , das neutrale Wort "Maßfunktion m_K" gewählt haben.

Abschließend wollen wir noch bemerken, daß sich die in (1.8) bemerkte Analogie des Begriffs l_K zum Begriff der logischen Wahrheit der deduktiven Logik, ernsthaft parallelisieren läßt und sich insbesondere zeigen läßt, daß sich alle in dieser Arbeit gewonnenen Aussagen A , so deuten lassen: Nach Konstruktion eines geeigneten Begriffs der Interpretation In gilt: Jede Interpretation In , die die aus K folgenden komparativen Aussagen $\mathcal{H}_K$ wahr macht, macht die Aussagen A wahr.
Eine Interpretation In ist dabei definiert, als eine gewisse Menge von Präferenzaussagen über gewisse Wettsysteme.
Die Demonstration dieser Möglichkeit führt jedoch über den Rahmen dieser Arbeit hinaus.

Trotz dieser grundsätzlichen Ablehnung des Begriffs der subjektiven Wahrscheinlichkeit, wollen wir uns auch in der Folge, mitunter des eben hilfreichen Bilds der subjektiven Wahrscheinlichkeit bedienen.

(1.12) Inhaltliche Bemerkungen zum Begriff der Wette

Der Wahrscheinlichkeitsbegriff, ganz gleich ob komparativer oder quantitativer Art, soll als "Lebensweiser" dienen, oder technisch ausgedrückt, RATIONALE ENTSCHEIDUNGEN ermöglichen.
Eine der einfachsten Entscheidungssituationen ist die Präferenz zwischen Wetten,bzw. Wettsystemen.
Wie in [H'] dargestellt, läßt sich bereits aus diesem einfachen Spezialfall der Entscheidungstheorie der σ-Additionssatz und der Multiplikatinssatz für faire Wettquotienten ableiten, sowie außerdem zeigen, daß die Interpretation von fairen Wettquotienten als Wahrscheinlichkeit mit dem allgemeinen Ansatz der Entscheidungstheorie konsistent ist.
Die Anwendungsprinzipien 1 , 2 sollten die formal aus den Begriffen m_K abgeleiteten Begriffe der quantitativen logischen Wahrscheinlichkeit, bzw. der quantitativen logischen Wahrscheinlichkeitsschranken, inhaltlich dahingehend erläutern, als die praktischen Konsequenzen angegeben wurden, die diese Begriffe für den Spezialfall von Entscheidungssituationen, nämlich die Präferenz zwischen Wetten haben.
Aus diesen Spezialfällen soll natürlich die Anwendbarkeit dieser Begriffe auf allgemeine Situationen der Entscheidungstheorie nahegelegt werden.
Insbesondere soll das Anwendungsprinzip 1 das allgemeinere Anwendungsprinzip 1* suggerieren:

Anwendungsprinzip 1*
Seien δ_i Handlungen, die nur von den Ereignissen $E \in \mathcal{L}_K$ abhängen, dh.: δ_i seien Abbildungen von Ω in einen Folgenraum F , derart daß für jedes $f \in F$ die Menge $\{\delta_i = f\} \in \mathcal{L}_K$.
Auf F sei eine Utility-Funktion $U(f)$ gegeben. Dann gilt:
$U(\delta_i) := \int U(\delta_i(\omega))\, dl_K(\omega)$ gibt die Präferenzstruktur zwischen den δ_i wieder.

In ähnlicher Weise ließe sich eine Verallgemeinerung von Anwendungsprinzip 2 formulieren.
Es ist ein Ziel der Erweiterung dieser Arbeit , die Ableitung von

von σ-Additionssatz und Multiplikationssatz auf eine allgemeine Theorie der Präferenz zwischen Handlungen zu stützen. Es wird in diesem Zusammenhang auf das System von SAVAGE [S] verwiesen; das von uns avisierte System soll sich dabei von dem System von SAVAGE dahingehend unterscheiden, daß die Ableitung des σ-Additionssatzes mit einer größeren Zwangsläufigkeit als bei SAVAGE erfolgt.
Abschließend sei bemerkt, daß auch bei dieser avisierten allgemeineren Ableitung von σ-Additionssatz und Multiplikationssatz die spezielle Deutung von Wahrscheinlichkeitswerten als faire Wettquotienten erhalten bleibt, da Wetten eben spezielle Handlungen darstellen.

§2 DER BEGRIFF DES EXPERIMENTS UND DER STATISTISCHEN GESAMTHEIT

(2.1) Experiment und experimentelles Schema

Es ist eine grundlegende Vorstellung jeder Wissenschaft, daß sich ein gewisses Experiment wiederholen lasse; dh. daß gewisse Phänomene von gleicher Art seien.
Wir versuchen diese Vorstellung wie folgt zu erfassen: Eine Klasse von gleichartigen Experimenten ist gegeben durch eine Versuchsvorschrift B innerhalb einer Sprache L, der jedes dieser Experimente zu gehorchen hat. Genauer gesagt: Wir wollen annehmen, daß die Versuchsbeschreibung B einen Parameter λ enthält, durch dessen Spezialisierung genau ein Experiment gegeben ist; es gebe nicht zwei Experimente, die der Beschreibung $B(\lambda)$ genügen.
Wir nennen dementsprechend:

$B(\lambda)$ ein konkretes Experiment

$B = \{B(\lambda) : \lambda \in \Lambda\}$ ein experimentelles Schema .

Als Beispiel betrachten wir das von Michelson durchgeführte Experiment $M(x,t)$; die Parameter (x,t) geben Ort und Zeit für die Durchführung des Experiments an.

Zu der Versuchsbeschreibung B gehört die Angabe, welche Beobachtungsergebnisse für möglich gehalten werden; wir beschränken uns hier auf den Fall, daß jedes der Experimente $B(\lambda)$ genau k diskrete mögliche Ergebnisse $e_1(\lambda), \ldots, e_k(\lambda)$ besitzt.
Das Ereignis, daß bei dem Versuch $B(\lambda)$ das Ergebnis $e_i(\lambda)$ eingetreten ist, bezeichnen wir mit $E_i(\lambda)$.
Die Ereignisse $E_i(\lambda)$ sollen dabei eine vollständige Ereignisdisjunktion bilden.

Zur mengentheoretischen Konstruktion des avisierten Meßraums $(\Omega,\mathcal{K})$ sei das Folgende bemerkt:
$\Omega := \Pi' \Omega(\lambda)$, $\Omega(\lambda) := \{e_1(\lambda), \ldots, e_k(\lambda)\}$

$\Omega \ni \omega = \{e_{\nu_\lambda}(\lambda) : \lambda \in \Lambda\}$, wobei ν_λ eine jedem $\omega \in \Omega$ eineindeutig zugeordnete Abbildung $\nu_\lambda : \Lambda \to \{1,\ldots,k\}$ ist.

$K := \prod^{\times} \mathcal{P}(\Omega(\lambda))$ Produkt-σ-Körper der Potenzmengen von $\Omega(\lambda)$.
$E_i(\lambda_o) := Z(e_i(\lambda_o))$:= Zylinder zu $e_i(\lambda_o)$ in Ω
$:= \{\omega = \{e_{\nu_\lambda}(\lambda) : \lambda \in \Lambda\} : \nu_{\lambda_o} = i\}$.

$\mathfrak{E}$ heißt ein Ereignisschema von $\mathfrak{B}$, wenn gilt:
$\mathfrak{E} = \{E(\lambda) : \lambda \in \Lambda\}$ mit $E(\lambda) = E_{i_1}(\lambda) + \ldots + E_{i_r}(\lambda)$.

Genaugenommen sollten wir statt $\mathfrak{E}$ und $E(\lambda)$
$\mathfrak{E}|\mathfrak{B}$ und $E|B(\lambda)$ schreiben; davon sehen wir aber in dieser Arbeit ab.

(2.2) Verallgemeinerung der Konzeption der Gleichartigkeit

Die Vorstellung, daß die Experimente $B(\lambda)$ von gleicher Art seien, beruht bei uns auf einer Konzeption K .
Wir haben in (1.3) dargestellt, wie der Begriff der Gleichartigkeit auf Grund einer Konzeption K zu erfassen ist.
Wir erweitern diesen Begriff der Konzeption der Gleichartigkeit nunmehr auf die Gesamtheit $\mathfrak{B} = \{B(\lambda) : \lambda \in \Lambda\}$:
Die Experimente $B(\lambda)$ heißen dabei gleichartig auf Grund von K , wenn auf Grund von K der Parameter λ irrelevant für den Eintritt der Ereignisse $E_i(\lambda)$ ist.
Analog zu (1.3) formulieren wir das folgende Axiom:

Axiom 5

Aus der Gleichartigkeit der Experimente $B(\lambda)$ auf Grund einer Konzeption K ergibt sich:
$E_{i_1}(\lambda_1)\ldots E_{i_n}(\lambda_n) \sim_K E_{i_1}(\mu_1)\ldots E_{i_n}(\mu_n)$, wobei $\lambda_i \neq \lambda_j$
$\mu_i \neq \mu_j$ für $i \neq j$.

Bemerkung: Axiom 2 ist ein Spezialfall von Axiom 5 ; man setze $k = 2$ und $E_1(\lambda) = E(\lambda)$, $E_2(\lambda) = \bar{E}(\lambda)$.

Satz 1

$\mathfrak{B} = \{B(\lambda) : \lambda \in \Lambda\}$ sei gleichartig auf Grund von K .
$\mathfrak{E} = \{E(\lambda) : \lambda \in \Lambda\}$ sei ein beliebiges Ereignisschema von $\mathfrak{B}$,dh.:
$E(\lambda) = E_{i_1}(\lambda) + \ldots + E_{i_r}(\lambda)$. Dann gilt: Die $E(\lambda)$ sind gleichartig auf Grund von K , dh.:
$E(\lambda_1)\ldots E(\lambda_n) \sim_K E(\mu_1)\ldots E(\mu_n)$, $\lambda_i \neq \lambda_j$, $\mu_i \neq \mu_j$ für $i \neq j$.

Beweis

$$E(\lambda_1)\ldots E(\lambda_n) = \sum_{1 \leq \nu_j \leq r} E_{i_{\nu_1}}(\lambda_1)\ldots E_{i_{\nu_n}}(\lambda_n) .$$

Die analoge Darstellung gilt für $E(\mu_1)\ldots E(\mu_n)$.
Nach Axiom 5 gilt:
$E_{i_{\nu_1}}(\lambda_1)\ldots E_{i_{\nu_n}}(\lambda_n) \sim_K E_{i_{\nu_1}}(\mu_1)\ldots E_{i_{\nu_n}}(\mu_n)$.

Damit folgt die Behauptung gemäß Satz 2 von (1.7) .

Satz 2

$A_i = \{A_i(\lambda) : \lambda \in \Lambda\}$ seien beliebige Ereignisschemata zu $\mathfrak{B}$.
$\mathfrak{B}$ sei gleichartig auf Grund von K . Dann gilt: Die Ereignisse $A_i(\lambda)$ sind gleichartig auf Grund von K , dh.:
$A_{i_1}(\lambda_1)\ldots A_{i_n}(\lambda_n) \sim_K A_{i_1}(\mu_1)\ldots A_{i_n}(\mu_n)$, $\lambda_i \neq \lambda_j$, $\mu_i \neq \mu_j$ für $i \neq j$.

Beweis

Der Durchschnitt $A_{i_1}(\lambda_1)\ldots A_{i_n}(\lambda_n)$ wird in gleicher Weise aus den Ereignissen $E_i(\lambda_j)$ aufgebaut wie der Durchschnitt $A_{i_1}(\mu_1)\ldots A_{i_n}(\mu_n)$ aus den Ereignissen $E_i(\mu_j)$. Analog wie in Satz 1 ergibt sich daher aus Axiom 5 und Satz 2 von (1.7) die gewünschte Behauptung.

Definition

Folgt aus der Konzeption K die Gleichartigkeit des Schemas $\mathfrak{B} = \{B(\lambda) : \lambda \in \Lambda\}$ gemäß Axiom 5 , so nennen wir die Klasse $\mathfrak{B}$ eine statistische Gesamtheit auf Grund von K .

(2.3) Die Maßfunktionen

Die $B(\lambda)$ seien gleichartig auf Grund von K . $\mathcal{K}$ sei der von den Ereignissen $E_i(\lambda)$ erzeugte σ-Körper:
$\mathcal{K} = {}^{B}\{E_i(\lambda) : \lambda \in \Lambda\ ,\ 1 \leq i \leq k\}$.
Dann gilt analog zu (1.5) :

Satz 1
Eine für K zulässige Maßfunktion m_K auf $\mathcal{K}$ ist durch die folgenden Zahlen eindeutig festgelegt:

$$m_K(E_1(\lambda_1^1)\ldots E_1(\lambda_{s_1}^1).E_2(\lambda_1^2)\ldots E_2(\lambda_{s_2}^2)\ldots E_k(\lambda_1^k)\ldots E_k(\lambda_{s_k}^k))$$

$$=:w_{s_1,\ldots,s_k}\quad ,\ \text{alle Parameter } \lambda_j^i \text{ verschieden .}$$

Nach Axiom 5 hängt das Maß m_K für den obigen Durchschnitt nur davon ab, wie oft die k verschiedenen Ereignissorten in diesem Durchschnitt vorkommen, nicht jedoch von den speziellen λ_j^i .

Beweis
Analog zu (1.5) ergibt sich:
Jedes Ereignis $A \in \mathcal{G} := {}^{K}\{E_i(\lambda) : \lambda \in \Lambda\ ,\ 1 \leq i \leq k\}$ liegt in einem Mengenkörper $\mathcal{G}(A)$, der aus endlich vielen, etwa N , $E_i(\lambda)$ erzeugt wird. $\mathcal{G}(A)$ hat also die Atome:
$E_{i_1}(\sigma_1)\ldots E_{i_N}(\sigma_N)$. $m_K(A)$ stellt also die Summe gewisser Zahlen $w_{s_1,\ldots,s_k}$ dar. Also ist m_K auf $\mathcal{G}$ durch die Zahlen $w_{s_1,\ldots,s_k}$ eindeutig festgelegt und somit auch auf der Borel'schen Erweiterung $\mathcal{K}$ von $\mathcal{G}$.

Satz 2
Seien $A_1,\ldots,A_t$ beliebige Ereignisschemata von B , dann gilt nach Satz 2 von (2.2) analog wie für die Ereignisse $E_i(\lambda)$:
$m_K(A_1(\lambda_1^1)\ldots A_1(\lambda_{s_1}^1)\ldots A_t(\lambda_1^t)\ldots A_t(\lambda_{s_t}^t)) =: v_{s_1,\ldots,s_t}$

hängt nur von $s_1,\ldots,s_t$ ab, nicht jedoch von den speziellen Parametern λ_j^i , vorausgesetzt, daß alle λ_j^i verschieden sind.

De Finetti - Theorem 2
Voraussetzung: $|\Lambda| \geq \aleph_0$.

a) Sei m ein normiertes Maß auf dem von den Ereignissen $A_i(\lambda)$, $1 \leq i \leq t$, erzeugten σ-Körper $\mathcal{K}'$ mit der Eigenschaft, daß für beliebige Parameter λ_j^i (alle λ_j^i verschieden) gilt:

$$m(A_1(\lambda_1^1)\ldots A_1(\lambda_{s_1}^1)\ldots A_t(\lambda_1^t)\ldots A_t(\lambda_{s_t}^t)) = v_{s_1,\ldots,s_t} \quad .$$

Dann gibt es ein normiertes Maß ϕ auf den Borel'schen Mengen des Würfels $W := \{(x_1,\ldots,x_t) : \text{für alle } i,\ x_i \in R\ ,\ 0 \leq x_i \leq 1\}$, mit der Eigenschaft:
$v_{s_1,\ldots,s_t} = \int x_1^{s_1}\ldots x_t^{s_t}\, d\phi$.

b) Auf Grund der Relation $\sum_{i=1}^{k} E_i(\lambda) = \Omega$ ergibt sich speziell für die Ereignisse $E_i(\lambda)$ ein Maß ψ auf den Borelschen Mengen von $D := \{(x_1,\ldots,x_k)$: alle $x_i \geq 0$ sowie $\Sigma x_i = 1\}$, so daß:
$w_{s_1,\ldots,s_k} = \int x_1^{s_1}\ldots x_k^{s_k}\, d\psi$. Für ein allgemeines Ereignis A aus $\mathcal{K}$ gilt dabei wie in (1.5):
$m(A) = \int p_{x_1,\ldots,x_k}(A)\, d\psi$; dabei stellt $p_{x_1,\ldots,x_k}(A)$ die "Wahrscheinlichkeit" von A dar unter der Voraussetzung, daß die Ereignisse $E_i(\lambda)$ unabhängig sind mit den "Wahrscheinlichkeiten" x_i , $\Sigma x_i = 1$.

c) Stellt umgekehrt ψ ein normiertes Maß auf den Borelschen Mengen über D dar, so führt die Definition
$m(A) := \int p_{x_1,\ldots,x_k}(A)\, d\psi$ zu einem Maß auf $\mathcal{K}$, das der Bedingung der Gleichartigkeit genügt.

Zum Beweis des De Finetti - Theorems 2
Theorem 2 stellt analog wie in (1.5) die Erweiterung eines von De Finetti in F' bewiesenen Theorems 2' dar. Der Beweis von Theorem 2 auf Grund von Theorem 2' ergibt sich analog wie in (1.5) . Theorem 2' macht wie in (1.5) die stärkere Voraussetzung $|\Lambda| = \aleph_0$ und beweist die Formel
$m(A) = \int p_{x_1,\ldots,x_k}(A)\, d\psi$ nur für A aus dem Mengenkörper $\mathcal{G}$.

Bemerkung

Das Maß ϕ ist durch die Zahlen $v_{s_1,\ldots,s_t}$ eindeutig bestimmt. Es gilt nämlich der folgende

Hilfssatz (siehe etwa [R'] V (6.46)): Eine Verteilungsfunktion $F(y)$ ist durch ihre Momente

$\mu'_{r_1,\ldots,r_k} := \int y_1^{r_1} \ldots y_k^{r_k}\, dF(y)$ eindeutig bestimmt, sofern

gilt: Die Summme $\sum_{r_i \in N} \frac{\mu'_{r_1,\ldots,r_k}}{r_1!\ldots r_k!} z_1^{r_1} \ldots z_k^{r_k}$ konvergiert

bei geeignetem $U > 0$ für alle z mit der Eigenschaft $|z_i| < U$ für alle i .

Da im DE FINETTI - Theorem alle Momente beschränkt sind, läßt sich dieser Hilfssatz anwenden.

§3 Das statistische Verhalten gleichartiger Phänomene

(3.1) Die Konvergenz der relativen Häufigkeiten

Der Ausgangspunkt unserer Betrachtungen bildet eine statistische Gesamtheit $B = \{B(\lambda) : \lambda \in \Lambda\}$ auf Grund einer Konzeption K, sowie der σ-Körper $\mathcal{K} := {}^{B}\{E_i(\lambda): \lambda \in \Lambda , 1 \leq i \leq k\}$ über der Grundmenge Ω mit den Elementen $\omega \in \Omega$, sowie ein beliebiges Ereignisschema $\mathfrak{E}$ von B .
Wir wollen nun das Folgende beweisen: Auf Grund der Konzeption K hat mit quantitativer logischer Wahrscheinlichkeit Eins jede Realisierungsfolge von B denselben Limes der relativen Häufigkeiten für den Eintritt der Ereignisse $E(\lambda)$ in dieser Folge.

Zum Beweis betrachten wir eine fest vorgegebene Realisierungsfolge $B(\lambda_i)$.
$h_n(E,\omega)$ seien die relativen Häufigkeiten des Ereignisschemas E in dieser Folge. Dann definieren wir die folgende Funktion:

$$p(E,\omega) := \begin{cases} \lim\limits_{n\to\infty} h_n(E,\omega) & \text{, sofern dieser Limes existiert} \\ 0 & \text{sonst} \end{cases}$$

Definition

$p(E,\omega)$ heißt die statistische Wahrscheinlichkeitsfunktion von E .

Bemerkung: Gemäß seiner Definition ist $p(E,\omega)$ abhängig von der Folge $\{\lambda_i\}$; in Satz 2 wird gezeigt, daß $p(E,\omega)$ bis auf eine 1_K-Nullmenge unabhängig ist von der speziellen Folge $\{\lambda_i\}$.

Sei $\Omega = \Omega^* + \Omega'$ mit $\Omega^* := \{\omega: \lim\limits_{n\to\infty} h_n(E,\omega) = p(E,\omega)\}$.

Satz 1

$1_K(\Omega^*) = 1$ oder ausführlich $1_K(\{\omega: h_n(E,\omega) \to p(E,\omega)\}) = 1$.

Satz 1 ist mathematisch gesehen nichts anderes als das Gesetz der starken Konvergenz der großen Zahlen für symmetrische Wahrscheinlichkeitsfunktionen und findet sich im Prinzip schon bei DE FINETTI (siehe [F]).
Die Formulierung $l_K(\{\omega : h_n(E,\omega) \to p(E,\omega)\}) = 1$ soll dabei nur sinnfällig machen, daß die Konvergenz der relativen Häufigkeiten mit praktischer Sicherheit eine objektive Konsequenz aus der Konzeption der Gleichartigkeit darstellt.
Zu DE FINETTI bestehen allerdings auch einige mathematische Unterschiede. Da DE FINETTI nur über die endliche Additivität der subjektiven Wahrscheinlichkeit verfügt und die σ-Additivität ablehnt, kann er das obige Gesetz nicht in dieser Form darstellen; er gebraucht ein Konvergenzkriterium. Da in [H'] die σ-Additivität für die subjektive Wahrscheinlichkeit abgeleitet wurde, sind wir hier nicht an die Beschränkung DE FINETTI's gebunden.
Wie wir gleich sehen werden , ist der Beweis des obigen Satzes im Rahmen unserer Arbeit wesentlich einfacher als die diesbezügliche Beweisskizze DE FINETTI's für das von ihm formulierte Konvergenzkriterium für die starke Konvergenz.

Die Aussage $l_K(\Omega^*) = 1$, läßt sich auch wie folgt ausdrücken: $\Omega^* \sim_K \Omega$. Das heißt, das Ereignis der Konvergenz der relativen Häufigkeiten ist ein maximales Element, gleichwahrscheinlich zur logischen Gewißheit, in der aus der Konzeption K resultierenden komparativen Struktur. - In dieser Form tritt die Tatsache, daß diese Aussage eine objektive Folgerung aus der Konzeption der Gleichartigkeit ist, besonders deutlich hervor, da in dieser Form jeder Bezug auf subjektive Wahrscheinlichkeiten, bzw. die Deutung von objektiv im Sinn von intersubjektiv vermieden wird.
Ob diese Aussage ohne quantitative Hilfsmittel allein unter Ausnützung der angegebenen Axiome 1 , 2 und der Sätze 1 , 2 von (1.7) bezüglich der komparativen Struktur bewiesen werden kann, ist eine reizvolle mathematische Aufgabe.
Diese Aufgabe liegt nicht im Bereich dieser Arbeit, da das unumgänglich notwendige Ergebnis (3.4) auf jeden Fall quantitative Begriffe erforderlich macht.

Beweis von Satz 1
Der Beweis ergibt sich aus dem üblichen starken Gesetz der großen Zahlen, in dem man das DE FINETTI - Theorem 1 von (1.5) anwendet. Nach dem üblichen starken Gesetz gilt nämlich mit den Bezeichnungen von (1.5) :

$p_x(\{\omega : h_n(E,\omega) \to x\}) = 1$.

Da $\{\omega : h_n(E,\omega) \to x\} \subseteq \{\omega: h_n(E,\omega) \to p(E,\omega)\}$ folgt:

$p_x(\{\omega: h_n(E,\omega) \to p(E,\omega)\}) = 1$; die Integration dieser Gleichung nach ϕ liefert: $m_K(\{\omega : h_n(E,\omega) \to p(E,\omega)\}) = 1$.
Dies gilt für jede Maßfunktion m_K von K ; daraus folgt die Behauptung.

Satz 2
Je zwei Realisierungsfolgen haben mit quantitativer logischer Wahrscheinlichkeit Eins denselben Limes der relativen Häufigkeiten: $l_K(\{\lim_{n\to\infty} h_n = \lim_{n\to\infty} h_n'\}) = 1$.

Dabei sei $h_n(E,\omega)$ die relative Häufigkeit von E in der Folge $B(\lambda_i)$ und $h_n'(E,\omega)$ in $B(\lambda_i')$.

Beweis
Wie bereits ausgeführt gilt: $p_x(\{\lim h_n = x\}) = p_x(\{\lim h_n' = x\}) = 1$ und folglich $p_x(\{\lim h_n = \lim h_n'\}) = 1$. Sei nämlich:

$\Omega_x := \{\omega: h_n(E,\omega) \to x\}$, $\Omega_x' := \{\omega: h_n'(E,\omega) \to x\}$; dann gilt:

$p_x(\Omega_x) = p_x(\Omega_x') = 1$ und folglich $p_x(\Omega_x.\Omega_x') = 1$, denn

$p_x(\overline{\Omega_x.\Omega_x'}) \leq p_x(\overline{\Omega_x}) + p_x(\overline{\Omega_x'}) = 0$. Weiter gilt:

$\Omega_x.\Omega_x' \subseteq \{\omega : \lim h_n(E,\omega) = \lim h_n'(E,\omega)\}$. Also ist

$1 = p_x(\Omega_x.\Omega_x') \leq p_x(\{\lim h_n = \lim h_n'\})$. Durch Integration nach ϕ folgt die Behauptung.

Satz 3

Jede abzählbare Klasse F von Realisierungsfolgen hat mit quantitativer logischer Wahrscheinlichkeit Eins denselben Limes der relativen Häufigkeiten:

$$l_K\Big(\prod'_{\{\lambda_i\}\in F}\{\omega: h_n^{\{\lambda_i\}}(E,\omega) \to p(E,\omega)\}\Big) = 1 .$$

$h_n^{\{\lambda_i\}}$ ist dabei die relative Häufigkeit in der Folge $\{\lambda_i\}$.

Der Beweis erfolgt analog wie der Beweis zu Satz 2 , da die abzählbare Vereinigung von Nullmengen wieder eine Nullmenge ist.

Wir benötigen den folgenden

Hilfssatz

Sei (Ω,K,μ) ein Wahrscheinlichkeitsfeld und $A_i \in K$, $i\in I$ eine überabzählbare Familie von Mengen. Dann ist durch die folgenden Bedingungen μ-fast eindeutig eine Menge A^* festgelegt:

(1) $A^* \subseteq A_i$ μ-fast, für jedes i .

(2) $A' \subseteq A_i$ μ-fast, für jedes i $\Rightarrow$ $A' \subseteq A^*$ μ-fast .

A^* heißt der μ-essentielle Durchschnitt der A_i und wird wie folgt bezeichnet: $A^* = \mu\text{-}\prod'_i A_i$.

Bemerkung: A^* ist derjenige abzählbare Durchschnitt der A_i , der von minimalem Maß ist unter allen abzählbaren Durchschnitten aus den A_i .

Mit den Begriffsbildungen dieses Hilfssatzes läßt sich Satz 3 wie folgt formulieren:

Satz 4

Es besteht quantitative logische Wahrscheinlichkeit Eins, daß l_K-essentiell alle Realisierungsfolgen des Schemas B zu demselben Limes der relativen Häufigkeiten führen:

$$l_K\Big(l_K\text{-}\prod'_{\{\lambda_i\}}\{\omega: h_n^{\{\lambda_i\}}(E,\omega) \to p(E,\omega)\}\Big) = 1 .$$

(3.2) Der Begriff der statistischen Wahrscheinlichkeit

Um die Klassische Wahrscheinlichkeitsrechnung auf der Grundlage unseres Ansatzes zu rekonstruieren, müssen wir jedem Ereignisschema E einer statistischen Gesamtheit B eine Zahl, die statistische Wahrscheinlichkeit von E , zuordnen. In Anlehnung an MISES (siehe [M]) definieren wir die statistische Wahrscheinlichkeit als Limes der tatsächlich auftretenden relativen Häufigkeiten; dieser Limes existiert bei uns mit quantitativer logischer Wahrscheinlichkeit Eins.

Genauer: Sei $\hat{\omega}$ das wahre, uns normalerweise unbekannte, in der Natur tatsächlich vorliegende Ergebnis aus Ω .
Dann nennen wir die folgende Zahl
$P(E) := p(E,\hat{\omega})$
die statistische Wahrscheinlichkeit von E .

Jedem Ereignis A aus K entspricht eine Aussage $\hat{A}$, die besagt, daß A eingetreten ist. Wir bezeichnen die Funktionen m_K , l_K, übertragen auf die Menge der Aussagen, mit m_K , l_K : $m_K(\hat{A}) := m_K(A)$, $l_K(\hat{A}) := l_K(A)$; zu bemerken ist dabei, daß logisch äquivalente Aussagen dieselben Werte m_K , l_K erhalten.

Die Aussage $\hat{A}$ ist logisch äquivalent zu der Aussage $\hat{\omega} \in A$.
Damit ergibt sich aus (3.1) der Satz: $l_K(\hat{\omega} \in \Omega^*) = 1$.
Die Aussage $\hat{\omega} \in \Omega^*$ ist logisch äquivalent zu der Aussage $h_n(E) \to P(E)$, wobei $h_n(E) := h_n(E,\hat{\omega})$. Folglich gilt:

$$\boxed{l_K(h_n(E) \to P(E)) = 1 .}$$

Der Dualismus der Formeln für die statistische Wahrscheinlichkeitsfunktion und die statistische Wahrscheinlichkeit gilt für alle in der Folge von uns abgeleiteten Formeln.
Zum Verhältnis der Größen $p(E,\omega)$ und $P(E)$ sind noch einige Erläuterungen nötig: Zunächst ergibt sich nach unserem Vorgehen

eine mathematisch völlig neue Situation, nämlich die Beschreibung statistischer Gesetzmäßigkeiten durch die Funktion $p(E,\omega)$, von uns statistische Wahrscheinlichkeitsfunktion genannt. Im Gegensatz zur Klassischen Wahrscheinlichkeitsrechnung, wo mit einer Zahl gearbeitet wird, ergibt sich bei uns die Wahrscheinlichkeitsrechnung mit der Funktion $p(E,\omega)$. Es frägt sich nun, wie der Zusammenhang dieser neuartigen Wahrscheinlichkeitsrechnung mit der Klassischen Wahrscheinlichkeitsrechnung aussieht. Dieser Zusammenhang ergibt sich durch die obige Definition der statistischen Wahrscheinlichkeit: $P(E) = p(E,\hat{\omega})$.
Wie wir im Folgenden sehen werden, lassen sich alle Formeln bezüglich der statistischen Wahrscheinlichkeitsfunktion in äquivalente Formeln bezüglich der statistischen Wahrscheinlichkeit umformen. Damit ist gezeigt, daß die Wahrscheinlichkeitsrechnung mit der Funktion $p(E,\omega)$ äquivalent ist zu der Klassischen Wahrscheinlichkeitsrechnung mit der Zahl $P(E)$.
Allerdings ist $p(E,\omega)$ der primäre Begriff, während $P(E)$ abgeleitet ist.

Bemerkung

Während sich der Begriff der quantitativen logischen Wahrscheinlichkeit auf ein konkretes Ereignis bezog, bezieht sich der Begriff der statistischen Wahrscheinlichkeit, bzw. - Funktion, auf ein Ereignisschema E und besagt im Allgemeinen nichts über ein spezielles $E(\lambda_o)$ aus E.
Ein spezielles $E(\lambda_o)$ bestimmt nämlich im Allgemeinen das Ereignisschema E nicht eindeutig. Es könnte nämlich sein, daß das spezielle $E(\lambda_o)$ innerhalb der Sprache L von der Gestalt ist:
$E(\lambda_o) = E^*(\mu_1,\lambda_o)$ und etwa das Ereignisschema
$E^* := \{E^*(\mu,\lambda) : (\mu,\lambda) \in \Gamma\}$ gar keine gleichartige Gesamtheit darstellt, insbesondere, daß sich etwa die wie folgt zu charakterisierende Situation ergibt:
Das Ereignisschema:
$E_{\mu_1} := \{E^*(\mu_1,\lambda) : \lambda \in \Lambda\}$ ist gleichartig auf Grund von K.

Das Ereignisschema:
$E_{\lambda_1} := \{E^*(\mu,\lambda_1) : \mu \in M\}$ ist gleichartig auf Grund von K.

Falls $\lambda_1 \in \Lambda$ und $\mu_1 \in M$ folgt daraus:
$E^*(\mu_1,\lambda_1) \sim_K E^*(\lambda,\mu_1)$, sowie $E^*(\lambda_1,\mu_1) \sim_K E^*(\lambda_1,\mu)$
und folglich: $E^*(\lambda,\mu) \sim_K E^*(\lambda',\mu')$ für alle $(\lambda,\mu) \in \Lambda\times M$.
Es kann aber sein, daß nicht mehr gilt:
$E^*(\mu,\lambda).E^*(\mu',\lambda') \sim_K E^*(\overline{\mu},\overline{\lambda}).E^*(\mu'',\lambda'')$.
Es ließe sich dann nicht mehr beweisen:
$l_K(P(E_{\mu_1}) = P(E_{\lambda_1})) = 1$ oder ausführlich:
$l_K(\{\omega : p(E_{\mu_1},\omega) = p(E_{\lambda_1},\omega)\}) = 1$.

Deshalb muß die statistische Wahrscheinlichkeit auf ein Ereignisschema bezogen werden, da ein spezielles Ereignis im Allgemeinen verschiedene Ereignisschemata zuläßt, zu denen es gehören kann und deshalb, wie ausgeführt, Mehrdeutigkeiten entstehen könnten.

(3.3) Der Additonssatz für die statistische Wahrscheinlichkeit

Seien E, E' zwei Ereignisschemata von B mit der Eigenschaft $E(\lambda).E'(\lambda) = \emptyset$ für alle λ; dann bezeichnen wir mit $E + E'$ das folgende Ereignisschema:
$E + E' := \{E(\lambda) + E'(\lambda) : \lambda \in \Lambda\}$.
Mit diesen Bezeichnungen gilt:
$l_K(\{\omega : p(E+E',\omega) = p(E,\omega) + p(E',\omega)\}) = 1$.

$$\boxed{l_K(P(E+E') = P(E) + P(E')) = 1.}$$

Der Beweis ergibt sich auf Grund von (3.1) aus der Additivität der relativen Häufigkeiten.

(3.4) Die Geschwindigkeit der Konvergenz

Die Geschwindigkeit der genannten Konvergenz der relativen Häufigkeiten läßt sich mit quantitativer logischer Wahrscheinlichkeit abschätzen. Es gilt das Analogon zur BERNOULLI'schen Ungleichung:

a) $m_K(\{\omega : |h_n(E,\omega) - p(E,\omega)| \leq \varepsilon\}) \geq 1 - 1/4n\varepsilon^2$

$$\boxed{m_K(|h_n(E) - P(E)| \leq \varepsilon) \geq 1 - 1/4n\varepsilon^2}$$

b) $m_K(\{\omega : |h_n(E,\omega) - h_r(E,\omega)| \leq \varepsilon) \geq 1 - |\frac{1}{r} - \frac{1}{n}|\frac{1}{4\varepsilon^2}$

$$\boxed{m_K(|h_n(E) - h_r(E)| \leq \varepsilon) \geq 1 - |\frac{1}{r} - \frac{1}{n}|\cdot\frac{1}{4\varepsilon^2}}$$

Bemerkung: Diese Abschätzungen gelten für jede Maßfunktion; die obigen Zahlenwerte haben also die in (1.9) formulierte Bedeutung von quantitativen logischen Wahrscheinlichkeitsschranken.

Beweis

1) Wir beweisen zuerst die Abschätzung a), die wir auf die übliche BERNOULLI'sche Ungleichung zurückführen. Diese Ungleichung lautet mit den Bezeichnungen von (1.5):

(α) $p_x(\{\omega : |h_n(E,\omega) - x| \geq \varepsilon\}) \leq \frac{x(1-x)}{n\cdot\varepsilon^2} \leq 1/4n\varepsilon^2$.

Andereseits gilt:

(β) $p_x(\{\omega : p(E,\omega) = x\}) = 1$.

Daraus folgt: $p_x(\{\omega : |h_n(E,\omega) - p(E,\omega)| \geq \varepsilon\}) \leq 1/4n\varepsilon^2$.

Durch Integration dieser Gleichung nach ϕ und Anwendung des Additionssatzes folgt die Behauptung a).

2) Beweis von Abschätzung b): Sei a eine K-meßbare Funktion auf Ω; dann bezeichnet $Ea := \int a \, dm_K$ den Erwartungswert von a bezüglich der Maßfunktion m_K.

Seien a_i die Indikatorfunktionen zu der vorausgesetzten Ereignisfolge $E(\lambda_i)$. Dann gilt:

α) $E(h_n - h_r)^2 = E(\frac{1}{n}\sum_{i=1}^{n} a_i - \frac{1}{r}\sum_{j=1}^{r} a_j)^2$;

Rechnet man diesen Ausdruck aus, so ergibt sich mit $Ea_i^2 = w_1$, $Ea_ia_j = w_2$ für $i \neq j$, die folgende Gleichung:

$$E(h_n - h_r)^2 = \frac{|n - r|}{nr}(w_1 - w_2) .$$

Beweis: Ohne Einschränkung der Allgemeinheit kann man voraussetzen $n \geq r$. Aus der Linearitat der Erwartungswertbildung ergibt sich nun:

$$E(h_n - h_r)^2 = \frac{1}{n^2}\sum_{\substack{1\leq i\leq n\\ 1\leq j\leq n}} Ea_ia_j + \frac{1}{r^2}\sum_{\substack{1\leq i\leq r\\ 1\leq j\leq r}} Ea_ia_j - \frac{2}{nr}\sum_{\substack{1\leq i\leq n\\ 1\leq j\leq r}} Ea_ia_j$$

Indem man berechnet wie oft in den obigen Summen die Indizes i, j übereinstimmen, bzw. verschieden sind, ergibt sich:

$$E(h_n - h_r)^2 = \frac{nw_1 + (n^2 - n)w_2}{n^2} + \frac{rw_1 + (r^2 - r)w_2}{r^2}$$

$$+ 2\,\frac{rw_1 + (nr - r)w_2}{nr} = (1/r - 1/n)(w_1 - w_2) .$$

β) Andererseits gilt:

$$E(h_n - h_r)^2 = \int (h_n - h_r)^2 \, dm_K \geq \int_{\{|h_n - h_r| \geq \varepsilon\}} (h_n - h_r)^2 \, dm_K$$

$$\geq \varepsilon^2 m_K(\{|h_n - h_r| \geq \varepsilon\}) .$$

Aus α) und β) folgt:

γ) $m_K(\{|h_n - h_r| \geq \varepsilon\}) \leq |1/n - 1/r|(w_1 - w_2)/\varepsilon^2$.

Nach (1.5) gilt: $w_1 - w_2 = \int x(1-x)\, d\phi \leq 1/4$.
Mit Hilfe des Additionssatzes folgt die Abschätzung b).

3) Es sei noch bemerkt, daß sich die Abschätzung a) auch aus der Abschätzung b) beweisen läßt, in dem man die Dreiecksungleichung anwendet:
$|h_n(E,\omega) - p(E,\omega)| \leq |h_n(E,\omega) - h_r(E,\omega)| + |h_r(E,\omega) - p(E,\omega)|$; nach (3.1) konvergiert der letzte Ausdruck bis auf eine m_K- Nullmenge gegen 0 für r gegen Unendlich; wollten wir diesen Beweis genau ausführen, müßten wir allerdings von dem Satz Gebrauch machen, daß die Konvergenz fast überall die gleichmäßige Konvergenz bis auf eine Menge beliebig klein wählbaren Maßes zur Folge hat. -Wir wollen uns mit dieser Andeutung einer weiteren Beweismöglichkeit für Abschätzung a) begnügen.

(3.5) De Finetti-Theorem

a) Von de Finetti (siehe[F]) wurde das folgende Resultat bewiesen:
Sei $B(\lambda_i)$ eine feste Realisierungsfolge von B ;
$A_1,\ldots,A_r$ seien beliebige Ereignisschemata von B ;
$h_n(A_\nu,\omega)$ seien die relativen Häufigkeiten der Ereingisse $A_\nu(\lambda_i)$ in der Folge $B(\lambda_i)$.
m_K sei ein normiertes Maß mit der Eigenschaft, daß für beliebige Parameter λ_j^i (alle λ_j^i verschieden) gilt:

$$m_K(A_1(\lambda_1^1)\ldots A_1(\lambda_{s_1}^1)\ldots A_r(\lambda_1^r)\ldots A_r(\lambda_{s_r}^r)) = v_{s_1,\ldots,s_r} \quad .$$

Dann gilt: Die Verteilungsfunktionen

$$F_n(x_1,\ldots,x_r) := m_K(\{\omega : h_n(A_1,\omega) \leq x_1 ,\ldots, h_n(A_r,\omega) \leq x_r \})$$

streben im Sinne der Verteilungskonvergenz gegen die Verteilungsfunktion Φ mit der Eigenschaft:

$$v_{s_1,\ldots,s_r} = \int x_1^{s_1}\ldots x_r^{s_r}\, d\Phi \quad .$$

b) Mit Hilfe der Funktionen $p(A_\nu,\omega)$ läßt sich das de Finetti-Theorem wie folgt schreiben:

$$v_{s_1,\ldots,s_r} = \int p(A_1,\omega)^{s_1}\ldots p(A_r,\omega)^{s_r}\, dm_K \quad .$$

c) $\Phi(x_1,\ldots,x_r) = m_K(\{\omega: p(A_1,\omega) \leq x_1 ,\ldots, p(A_r,\omega) \leq x_r \})$.

Beweis von c)
Nach (3.1) gilt: $h_n(A_\nu,\omega) \to p(A_\nu,\omega)$ m_K-fast; daraus folgt, daß die Verteilungsfunktionen F_n im Sinne der Verteilungskonvergenz gegen die Funktion $m_K(\{\omega: p(A_1,\omega) \leq x_1 ,\ldots, p(A_r,\omega) \leq x_r\})$ streben (siehe etwa [R'] VII (1.21)) . Daraus ergibt sich zusammen mit a) die Behauptung.

Beweis von b)
Wir benötigen den folgenden Hilfssatz:
Sei μ ein normiertes Maß auf $\mathcal{K}$ über Ω . $\alpha(\omega)$ sei ein Zufallsvektor auf Ω , $g(x)$ eine Baire'sche Funktion, $F_\alpha(x) := \mu(\{\omega: \alpha(\omega) \leq x\})$ die Verteilungsfunktion von α .
Dann gilt: $\int g(\alpha(\omega))\, d\mu = \int g(x)\, dF_\alpha$.

Es sei nun $\mu := m_K$, $\alpha(\omega) := (p(A_1,\omega),\ldots, p(A_r,\omega))$ und folglich nach c) $F_\alpha(x) = m_K(\{\omega: \alpha(\omega) \leq x\}) = \Phi(x)$; weiter sei $g(x) := x_1^{s_1}\ldots x_r^{s_r}$. Damit ergibt sich aus dem Hilfssatz

$$\int p(A_1,\omega)^{s_1}\ldots p(A_r,\omega)^{s_r}\, dm_K = \int x_1^{s_1}\ldots x_r^{s_r}\, d\Phi \quad .$$

Damit ist Teil b) der Behauptung bewiesen.

(3.6) Der Begriff der Unterkonzeption

Definition
$\hat{K}$ heißt eine Unterkonzeption von K , wenn die auf Grund von $\hat{K}$ zugelassenen Maßfunktionen eine Teilmenge der auf Grund von K zugelassenen Maßfunktionen bilden.

Daraus folgt nach Axiom 3 : $A \leq_K B \;\vdash\; A \leq_{\hat{K}} B$.

(3.7) Die Konzeptionen K_p

a)

Definition

Sei E ein beliebiges Ereignisschema von B ; sei $p \in [0,1]$

Dann definieren wir gemäß (1.10) :

$K_p := K \wedge \hat{A}$ mit $A := \{\omega: p(E,\omega) = p\}$. $\leftarrow \quad \hat{A} \leftrightarrow (P(E) = p)$.

Gemäß (1.10) hat die Konzeption K_p die folgenden Maßfunktionen:

$$m_{K_p}(C) = \frac{m_K(C.A)}{m_K(A)} \quad , \text{ sofern } m_K(A) \neq 0 .$$

Wir berechnen zunächst $m_K(A)$:

$m_K(A) = \int p_x(A)\, d\phi$ gemäß (1.5) . Statt p_x gebrauchen wir das Symbol μ_x . Es gilt nunmehr:

$$\mu_x(A) := p_x(A) = p_x(\{\omega: p(E,\omega) = p\}) = \begin{cases} 0 & \text{für } x \neq p \\ 1 & \text{für } x = p \end{cases}$$

Folglich: $m_K(A) = \phi(\{p\})$ und $m_K(C.A) = \mu_p(C)\cdot\phi(\{p\})$.

Damit ergibt sich: $m_{K_p}(C) = \mu_p(C)$.

Es gibt also auf Grund von K_p nur eine Maßfunktion auf $\mathcal{K}(E) :=^B \{E(\lambda) : \lambda \in \Lambda\}$, nämlich die Funktion μ_p .
K_p ist eine Unterkonzeption von K .

Satz

$$l_{K_p}(E(\lambda)) = m_{K_p}(E(\lambda)) = \mu_p(E(\lambda)) = p = P(E) \overset{l_{K_p}\text{-fast}}{=} p(E,\omega) .$$

Die quantitative logische Wahrscheinlichkeit von $E(\lambda)$ ist also auf Grund von K_p gleich der statistischen Wahrscheinlichkeit, bzw. l_{K_p} - fast gleich der statistischen Wahrscheinlichkeitsfunktion.
Damit ist motiviert, warum wir die Größen $p(E,\omega)$, $P(E)$ mit dem Wort Wahrscheinlichkeit in Zusammenhang brachten, das wir zu Beginn unserer Arbeit im Sinne eines fairen Wettquotienten konzipiert hatten.
Unter der Voraussetzung K_p stellt $P(E)$ den fairen Wettquotienten auf $E(\lambda)$ dar.

b) Als Beispiel betrachten wir den sukzessiven Wurf einer Münze. E sei das Ereignisschema, daß die Münze auf den Kopf fällt. Dann werden wir zunächst die Konzeption K annehmen, die die Gleichartigkeit der verschiedenen Münzwürfe besagt.
Die Münze habe sich nun bei genauer Messung als völlig symmetrisch erwiesen. Dann läßt sich aus Symmetriegründen die Konzeption $K_{1/2}$ begründen.

c) Die Aufgabe der Mathematischen und Angewandten Statistik ist es den Wert $P(E)$ möglichst zuverlässig und genau zu ermitteln. Eine theoretische Basis dafür bildet etwa die "BERNOULLI'sche Ungleichung" (3.4), auf Grund derer durch entsprechende Beobachtungszahl der Wert $P(E)$ beliebig genau und sicher ermittelt werden kann.
Die obigen Ausführungen zeigen, welche Konklusionen sich ergeben, wenn dieser Wert als "ermittelt" anzusehen ist, dh. der Konzeption K die Aussage $P(E) = p$ hinzugefügt wird.

d) Die Tatsache, daß die statistische Wahrscheinlichkeit $P(E)$ nach ihrer konkreten Ermittelung $P(E) = p$ zu der praktischen Konsequenz führt $l_{K_p}(E(\lambda_o)) = P(E)$ steht in einem gewissen Widerspruch zu den Ausführungen von Seite 38/39 , wo hervorgehoben wird, daß die statistische Wahrscheinlichkeit eine Eigenschaft von Ereignisschemata und nicht von Ereignissen ist, obwohl die obige Relation eben zeigt, daß unter der Konzeption K_p die statistische Wahrscheinlichkeit eben doch dem konkreten Ereignis $E(\lambda_o)$ in gewisser Weise zugeschrieben werden kann.
Das Wesen dieser gewissen Widersprüchlichkeit wird offenbar, wenn wir die bekannte Situation S zweier einander widersprechender statistischer Informationen betrachten:
Seien also E^1 , E^2 gleichartige Ereignisschemata auf Grund von K , die beide das Ereignis $E(\lambda_o)$ enthalten; außerdem gelte: $P(E^1) = p_1 \neq p_2 = P(E^2)$.

Die gewisse Widersprüchlichkeit der Situation S läßt sich durch die gleichzeitige Gültigkeit der folgenden Aussagen darstellen:

$$l_{K_1}(E(\lambda_o) = p_1 \neq p_2 = l_{K_2}(E(\lambda_o) \quad , \quad K_i := K \wedge (P(E^i) = p_i) \quad .$$

Dieser Widerspruch ist jedoch nur scheinbar, da in der Situation S weder die Konzeption K_1 , noch die Konzeption K_2 vorliegt, sondern die folgende Konzeption K' (requirement of total evidence) :

$K' := K \wedge (P(E^1) = p_1) \wedge (P(E^2) = p_2) \wedge (E(\lambda_o) \in E^1.E^2)$.

Die Situation S ist voll gemeistert, wenn zu $E(\lambda_o)$ eine quantitative logische Wahrscheinlichkeit $l_{K'}(E(\lambda_o))$ existiert. Von einer Lösung ist aber auch in dem bescheideneren Fall zu sprechen, wenn für das Ereignis $E(\lambda_o)$ quantitative logische Wahrscheinlichkeitsschranken s_i existieren, dh.:

$0 < s_1 \leq m_{K'}(E(\lambda_o)) \leq s_2 < 1$ für alle $m_{K'}$ zu K' .

Weitere Untersuchungen zu diesem Problem führen über den Rahmen dieser Arbeit hinaus; man beachte dazu (10.2D) und (10.2Fe) .

e) Abschließend bemerken wir noch:
P(E) ist die Eigenschaft eines Ereignisschemas,die im Allgemeinen nichts über $E(\lambda_o) \in E$ besagt.
In der speziellen Situation K_p gilt allerdings: P(E) ist der in dieser Situation eindeutig bestimmte faire Wettquotient von $E(\lambda_o)$, also gewissermaßen eine Eigenschaft von $(K_p, E(\lambda_o))$.
Bereits in der duch eine Zusatzinformation aus K_p hervorgehenden Situation K' ist dies nicht mehr der Fall.

(3.8) Die Konzeptionen K*

Definition
Sei E ein beliebiges Ereignisschema von B . Dann ist die Konzeption K* wie folgt charakterisiert: Es sind nur solche Maßfunktionen m_{K*} zugelassen, so daß für das Ereignisschema E gilt: Das gemäß (1.5) zugeordnete Maß ϕ^* besitzt eine stetige positive Dichte auf dem Einheitsintervall: $d\phi^* = f(x)\,dx$.
K* ist also eine Unterkonzeption von K .

Eine Konzeption K* ist genau dann gegeben, wenn die bestehenden theoretischen Vorstellungen und Informationen keine wahrscheinlichkeitstheoretischen Singularitäten bezüglich der Beschaffenheit von P(E) nahelegen.

Unter der Voraussetzung, daß ein K* vorliegt, sei $B(\lambda_i)$ eine beliebige Realisierungsfolge von B . Nach n Versuchen sei r_n-mal $E(\lambda)$ und $(n-r_n)$-mal $\overline{E}(\lambda)$ aufgetreten, so daß $h_n = r_n/n$ die relative Häufigkeit von $E(\lambda)$ in dieser Folge ist.

Dann gilt für die der Konzeption $K^*_{n,r_n} := K^* \wedge (h_n = r_n/n)$ zugeordneten Maßfunktionen nach dem Multiplikationssatz, sofern $\mu_i \in \Lambda\text{-}\{\lambda_1,\ldots,\lambda_n\}$:

$$m_{K^*_{n,r_n}}(E(\mu_1)\ldots E(\mu_s)) = \frac{m_{K^*}(\{h_n = r_n/n\}\cdot E(\mu_1)\ldots E(\mu_s))}{m_{K^*}(\{h_n = r_n/n\})}$$

$$= \frac{\int \binom{n}{r_n} x^{r_n}(1-x)^{n-r_n} x^s \, d\phi^*}{\int \binom{n}{r_n} x^{r_n}(1-x)^{n-r_n} \, d\phi^*} = \int x^s \, d\phi^*_{n,r_n} \qquad \text{mit}$$

$$d\phi^*_{n,r_n} := \frac{x^{r_n}(1-x)^{n-r_n}}{\int x^{r_n}(1-x)^{n-r_n} \, d\phi^*} \, d\phi^* \quad .$$

Die Konzeption K^*_{n,r_n} ist also gleichartig bezüglich des Ereigniskörpers $K_n(E) := \mathfrak{B}\{E(\lambda) : \lambda \in \Lambda\text{-}\{\lambda_1,\ldots,\lambda_n\}\}$, jedoch nicht für K ; sei nämlich etwa $E(\lambda_1)$ und $\overline{E}(\lambda_2)$ eingetreten, dann gilt $m_{K^*_{n,r_n}}(E(\lambda_1)) = 1$, $m_{K^*_{n,r_n}}(E(\lambda_2)) = 0$, im Gegensatz zur Gleichartigkeit.

Satz 1

Sei ϕ^* ein Maß der Konzeption K^* ,
$\Phi^*_{n,r_n}(x) := m_{K^*_{n,r_n}}(\{\omega: p(E,\omega) \leq x\})$ die dem Maß ϕ^*_{n,r_n} zugeordnete Verteilungsfunktion.
Konvergieren die relativen Häufigkeiten r_n/n gegen p für n gegen Unendlich, so konvergieren die Verteilungsfunktionen Φ^*_{n,r_n} gegen $D(x-p)$; $D(x)$ bezeichnet dabei die Dirichlet'sche Sprungfunktion, dh. $D(x)$ schreibt der Menge $\{0\}$ das Maß Eins zu und dem restlichen Einheitsintervall das Maß Null.

Die Konvergenz der Verteilungsfunktionen versteht sich dabei hier wie im Folgenden stets im Sinne der Verteilungskonvergenz.

Beweis

Wir gebrauchen den folgenden

Hilfssatz 1

Die Verteilungsfunktionen Ψ_o , Ψ_n haben die Momente

μ_o^s , μ_n^s , $s \in N$ mit $\mu_n^s \to \mu_o^s$ für $n \to \infty$ und zwar für alle $s \in N$; $|\mu_o^s|, |\mu_n^s| < C$, $C < \infty$.

Dann konvergieren die Verteilungsfunktionen Ψ_n im Sinne der Verteilungskonvergenz gegen Ψ_o .

Der Beweis des Hilfssatzes ergibt sich durch Betrachtung der charakteristischen Funktionen der Ψ_v ; siehe etwa [R'] V (6.15),(6.32),(7.7) .

Gemäß diesem Hilfssatz ist also zu beweisen:

$$r_n/n \to p \quad \Rightarrow \quad \int x^s \, d\phi^*_{n,r_n} \to p^s \ .$$

1) Sei zunächst speziell $d\phi^* = dx$. Dann gilt:

$$\int x^s \, d\phi^*_{n,r_n} = \frac{\int x^{r_n+s}(1-x)^{n-r_n}\,dx}{\int x^{r_n}(1-x)^{n-r_n}\,dx} = \frac{B(r_n+s+1,n-r_n+1)}{B(r_n+1,n-r_n+1)}$$

$B(x,y)$ bezeichnet dabei die Euler'sche Betafunktion.

Folglich gilt:

$$\int x^s \, d\phi^*_{n,r_n} = \frac{(r_n+s)!(n-r_n)!/(s+n+1)!}{r_n!(n-r_n)!/(n+1)!}$$

$$= \frac{(r_n+1)(r_n+2)\ldots(r_n+s)}{(n+2)\ldots(n+s+1)} \to p^s \ , \text{ sofern } r_n/n \to p \text{ für } n \to \infty \ .$$

2) Wir gebrauchen den folgenden Hilfssatz (siehe etwa [R'] V (7.4)):

Hilfssatz 2

Die Verteilungsfunktionen $\Psi_n(x)$ mögen im Sinne der Verteilungskonvergenz gegen $\Psi_o(x)$ konvergieren. Ist $g(x)$ stetig für alle $x \in R$, so gilt:

$$\lim_{n\to\infty} \int_{(|x| \leq a)} g(x)\, d\Psi_n = \int_{(|x| \leq a)} g(x)\, d\Psi_o$$

für jedes endliche $a > 0$, für das $\pm a$ keine Unstetigkeitskoordinate von $\Psi_o(x)$ ist.

Es gelte nun $r_n/n \to p$ und es sei $d\phi^* = f(x)\,dx$; $f(x)$ positiv und stetig auf dem Einheitsintervall. Gemäß Hilfssatz 1 ist zu beweisen:

$$Q := \frac{\int x^{r_n+s}(1-x)^{n-r_n} f(x)\,dx}{\int x^{r_n}(1-x)^{n-r_n} f(x)\,dx} \to p^s \quad \text{für } n \to \infty \ .$$

In dem Ausdruck Q dividieren wir nun Zähler und Nenner durch das Integral $\int x^{r_n}(1-x)^{n-r_n}\,dx$. Damit ergibt sich:

$$Q = Z/N \quad \text{mit} \quad Z := \frac{\int x^{s+r_n}(1-x)^{n-r_n} f(x)\,dx}{\int x^{r_n}(1-x)^{n-r_n}\,dx} \ ,$$

$$N := \frac{\int x^{r_n}(1-x)^{n-r_n} f(x)\,dx}{\int x^{r_n}(1-x)^{n-r_n}\,dx} \ .$$

Aus Hilfssatz 2 folgt für $\Psi_n := L_{n,r_n}$ mit $dL = dx$, $\Psi_o := D(x-p)$, $g(x) := x^s f(x)$, $a := 2$, da die Funktionen Ψ_n nach Beweisteil 1) gegen Ψ_o konvergieren:

$Z \to f(p)p^s$, $N \to f(p)$ für $n \to \infty$.

Der Faktor $f(p)$ kürzt sich weg; das liefert die Behauptung.

Satz 2

Φ^*_{n,r_n} sei eine der oben eingeführten Verteilungsfunktionen zur Konzeption K^*_{n,r_n} . $\Phi_{n,r_n(\omega)}$ symbolisiere die Abhängigkeit der Verteilungsfunktion Φ^*_{n,r_n} von dem auftretenden ω , da ja $r_n = r_n(\omega)$ von ω abhängt. Wir schreiben dafür kurz: $\Phi^*_n(\omega) := \Phi^*_{n,r_n(\omega)}$. Dann gilt:

$$l_{K^*}(\{\omega: \Phi^*_n(\omega) \to D(x-p(E,\omega))\}) = 1 .$$

Der Beweis ergibt sich aus (3.1) und dem vorangegangenen Satz 1 .

Anschaulich läßt sich dieser Satz wie folgt deuten: Die Konzeption K^* konvergiert bei zunehmender Information mit quantitativer logischer Wahrscheinlichkeit Eins gegen die gemäß (3.7) definierte Konzeption $K_{p(E,\omega)}$, in Formeln:

$$l_{K^*}(\{\omega: K^*_n(\omega) \to K_{p(E,\omega)}\}) = 1 \text{ , mit } K^*_n(\omega) := K^*_{n,r_n(\omega)} .$$

Mit $K^*_n := K^*_n(\hat{\omega})$ und $K_{P(E)} := K_{p(E,\hat{\omega})}$ ergibt sich:

$$\boxed{l_{K^*}(K^*_n \to K_{P(E)}) = 1 .}$$

Satz 3

$$m_n(\omega) := m_{K^*_{n,r_n(\omega)}}(E(\lambda_{n+1})) = \frac{m_{K^*}(E(\lambda_{n+1}) \cdot \{\omega': h_n(E,\omega') = r_n(\omega)/n\})}{m_{K^*}(\{\omega': h_n(E,\omega') = r_n(\omega)/n\})}$$

ist die bedingte subjektive Wahrscheinlichkeit von $E(\lambda_{n+1})$.
Es gilt:

$$l_{K^*}(\{\omega: m_n(\omega) \to p(E,\omega)\}) = 1 .$$

$m_n := m_n(\hat{\omega})$ ist die im Verlauf der Beobachtungen tatsächlich auftretende bedingte subjektive Wahrscheinlichkeit. Damit gilt:

$$\boxed{l_{K^*}(m_n \to P(E)) = 1 .}$$

Beweis

Der Beweis ergibt sich aus Satz 2 und Hilfssatz 2 :

$$m_n(\omega) = m_{K^*_{n,r_n(\omega)}}(E(\lambda_{n+1})) = \int x \, d\Phi^*_{n,r_n(\omega)} \underset{l_{K^*}\text{-fast}}{\to} \int x \, dD(x-p(E,\omega))$$

$$= p(E,\omega) .$$

Dieser Satz motiviert nochmals, warum wir $p(E,\omega)$ und $P(E)$ mit dem Begriff Wahrscheinlichkeit in Zusammenhang brachten. Außerdem ist motiviert, warum wir die statistische Wahrscheinlichkeit $P(E)$ auch als objektive Wahrscheinlichkeit bezeichnen; $P(E)$ ist nämlich nach der obigen Formel der von jeder Information und subjektiven Willkür unabhängige Grenzwert der bedingten Wahrscheinlichkeiten jeder möglichen subjektiven Wahrscheinlichkeitsbewertung der Konzeption K^* .

(3.9) Schlußbemerkung

Die übliche objektivistische Wahrscheinlichkeitsrechnung ist hiermit im folgenden Sinne gerechtfertigt: Die Konzeption $K_{P(E)} = \{\mu_{P(E)}\}$, die die übliche objektivistische Wahrscheinlichkeitsrechnung charakterisiert, ergibt sich bei uns als Grenzkonzeption bei hinreichend vielen Beobachtungen. Der Standpunkt,

diese Konzeption auf Grund der Gleichartigkeit mit einem festen, aber unbekannten $P(E)$ von vornherein vorauszusetzen und sich etwa auf Grund der BERNOULLI'schen Ungleichung um die möglichst sichere und genaue Ermittlung von $P(E)$ zu bemühen, ist also nach unserer Theorie ein voll begründetes Verfahren.- Wie schon in der Zusammenfassung bemerkt, erscheint es dabei adäquater, zusätzlich zu dem von uns begründeten Begriff $P(E)$ und der bezüglich der Konzeption $K_{P(E)}$ daraus abgeleiteten Wahrscheinlichkeitsrechnung, die von uns bezüglich der Maßfunktionen m_K entwickelten Begriffe der quantitativen logischen Wahrscheinlichkeit, bzw. der quantitativen logischen Wahrscheinlichkeitsschranke zu benutzen. - Diese Verwendung der Begriffe m_K und l_K scheint dabei nicht nur von theoretischem Wert,- zur Begründung der objektivistischen Wahrscheinlichkeitsrechnung - , sondern auch von praktischem Wert zu sein.
Die Darlegung dieser auch praktischen Bedeutung unserer Begriffe liegt jedoch nicht im Rahmen dieser Arbeit.

§4 DER BEGRIFF DER STATISTISCHEN UNABHÄNGIGKEIT

(4.1) Definiton der statistischen Unabhängigkeit auf Grund von K

Seien A_1 und A_2 zwei Ereignisschemata, dann bezeichnen wir mit $A_1.A_2$ das folgende Ereignisschema:

$A_1.A_2 := \{A_1(\lambda).A_2(\lambda) : \lambda \in \Lambda\}$.

Zwei Ereignisschemata A_1 und A_2 einer statistischen Gesamtheit B heißen auf Grund von K statistisch unabhängig, wenn gilt:

$l_K(\{\omega: p(A_1.A_2,\omega) = p(A_1,\omega).p(A_2,\omega)\}) = 1$, bzw.

$$\boxed{l_K(P(A_1.A_2) = P(A_1).P(A_2)) = 1 .}$$

(4.2) Äquivalente Bedingungen für die statistische Unabhängigkeit

A_1 und A_2 sind auf Grund von K genau dann statistisch unabhängig, wenn gilt:

$$A_1(\lambda_1)\ldots A_1(\lambda_n).A_2(\rho_1)\ldots A_2(\rho_r) \sim_K A_1(\mu_1)\ldots A_1(\mu_n).A_2(\delta_1)\ldots A_2(\delta_r) ;$$

dabei muß vorausgesetzt werden: $\lambda_i \neq \lambda_j$, $\rho_i \neq \rho_j$, $\mu_i \neq \mu_j$, $\delta_i \neq \delta_j$ für $i \neq j$; es ist jedoch zugelassen, daß gewisse der λ_i mit gewissen der ρ_j übereinstimmen.

Beweis
Seien A_1 und A_2 auf Grund von K statistisch unabhängig; dann ist zu zeigen, daß $m_K(A_1(\lambda_1)\ldots A_1(\lambda_n).A_2(\rho_1)\ldots A_2(\rho_r)$ nur von n , r abhängt.
Es gelte: $\lambda_{i_\nu} = \rho_{j_\nu}$ für $\nu = 1,\ldots,s$; $\lambda_i \neq \rho_j$ für $i \neq i_1,\ldots,i_s$ $j \neq j_1,\ldots,j_s$.

Durch Umnummerieren können wir erreichen, daß gilt:
$\lambda_i = \rho_i$ für $i = 1,\ldots,s$;
$\lambda_i \neq \rho_j$ für $i , j > s$.

Wir betrachten die folgenden drei Ereignisschemata:
A_1 , A_2 , $A_3 := A_1.A_2$.
Es gilt:
$A_1(\lambda_1)\ldots A_1(\lambda_n).A_2(\rho_1)\ldots A_2(\rho_r)$
$= A_3(\lambda_1)\ldots A_3(\lambda_s).A_1(\lambda_{s+1})\ldots A_1(\lambda_n).A_2(\rho_{s+1})\ldots A_2(\rho_r) =: A$

Aus (3.5) folgt:
$m_K(A) = \int p(A_1,\omega)^{n-s} p(A_2,\omega)^{r-s} p(A_3,\omega)^s \, dm_K$.
Es gilt nach Voraussetzung: $p(A_3,\omega) = p(A_1,\omega)p(A_2,\omega)$ m_K-fast.
Damit ergibt sich:
$m_K(A) = \int p(A_1,\omega)^n p(A_2,\omega)^r \, dm_K$; das war zu zeigen.
Die Umkehrung ergibt sich aus dem folgenden Satz:

(4.3) Minimalbedingungen für die statistische Unabhängigkeit

A_1 und A_2 sind auf Grund von K statistisch unabhängig, sofern:

$A_1(\lambda).A_2(\lambda) \sim_K A_1(\mu).A_2(\rho)$,

$A_1(\lambda).A_1(\mu).A_2(\lambda) \sim_K A_1(\sigma).A_1(\delta).A_2(\rho)$ mit $\lambda\neq\mu$, $\sigma\neq\delta$,

$A_1(\lambda).A_2(\lambda).A_2(\mu) \sim_K A_1(\sigma).A_2(\rho).A_2(\delta)$ mit $\lambda\neq\mu$, $\rho\neq\delta$,

$A_1(\lambda).A_1(\sigma).A_2(\lambda).A_2(\sigma) \sim_K A_1(\mu).A_1(\rho).A_2(\delta).A_2(\kappa)$ mit $\lambda\neq\sigma$, $\mu\neq\rho$, $\delta\neq\kappa$.

Beweis

Nach Voraussetzung gilt:

$m_K(A_1(\lambda).A_2(\mu)) =: v_{1,1}$ hängt nicht von λ , μ ab (auch für $\lambda = \mu$) .

$m_K(A_1(\lambda)A_1(\mu).A_2(\rho)) =: v_{2,1}$ hängt nicht von λ , μ , ρ ab, $\lambda \neq \mu$ (auch für $\rho = \mu$ oder $\lambda = \rho$) .

$m_K(A_1(\lambda).A_2(\mu).A_2(\rho)) =: v_{1,2}$ hängt nicht von λ , μ , ρ ab, $\mu \neq \rho$ (auch für $\lambda = \mu$ oder $\lambda = \rho$).

$m_K(A_1(\lambda).A_1(\rho).A_2(\mu).A_2(\delta)) =: v_{2,2}$ hängt nicht von λ , ρ , μ , δ ab, $\lambda \neq \rho$, $\mu \neq \delta$ (auch für $\lambda = \mu$ und $\rho = \delta$) .

Wir betrachten eine beliebige Folge von Experimenten $B(\lambda_i)$; $A_1(\lambda_i)$, $A_2(\lambda_i)$ haben die Indikatorfunktionen a_i , b_i und $A_1(\lambda_i).A_2(\lambda_i)$ folglich die Indikatorfunktion a_ib_i .

In der Folge $B(\lambda_i)$ seien h_n^j die relativen Häufigkeiten von $A_j(\lambda_i)$, h_n die relative Häufigkeit von $A_1(\lambda_i).A_2(\lambda_i)$.

Dann gilt:

$$E(h_n^1h_n^2 - h_n)^2 = E(1/n^2(\sum_{1\leq i,j\leq n} a_ib_j) - 1/n(\sum_{1\leq r\leq n} a_rb_r))^2$$

$$= 1/n^4(\sum_{1\leq i,j,r,v\leq n} Ea_ib_ja_rb_v) - 2/n^3(\sum_{1\leq i,j,r\leq n} Ea_ib_ja_rb_r)$$

$$+ 1/n^2(\sum_{1\leq r,j\leq n} Ea_rb_ra_jb_j) .$$

Es gilt nun:

$Ea_ia_rb_jb_v = v_{2,2}$, sofern $i \neq r$, $j \neq v$.

$Ea_ia_rb_j^2 = v_{2,1}$, sofern $i \neq r$.

$Ea_i^2b_jb_v = v_{1,2}$, sofern $j \neq v$.

$Ea_i^2b_j^2 = v_{1,1}$.

Indem man berechnet, wie oft in den obigen Summen gewisse Indizes übereinstimmen, bzw. verschieden sind, ergibt sich:

$$E(h_n^1h_n^2 - h)^2 = 1/n^4[\, n^2v_{1,1} + (n^3 - n^2)v_{2,1} + (n^3 - n^2)v_{1,2}$$

$$+ (n^4 - 2n^3 + n^2)v_{2,2}\,] - 2/n^3[\, nv_{1,1} + (n^2 - n)v_{2,1} + (n^2 - n)v_{1,2}$$

$$+ (n^3 - 2n + n)v_{2,2}\,] + 1/n^2[\, nv_{1,1} + (n^2 - n)v_{2,2}\,]$$

$$= (1/n - 1/n^2)(v_{1,1} - v_{1,2} - v_{2,1} - v_{2,2}) .$$

Daraus folgt, daß $h_n^1h_n^2 - h_n$ im Quadratmittel gegen 0 konvergiert. Daraus folgt, daß diese Folge auch nach Maß gegen 0 konvergiert (siehe etwa [R'] IV (3.7)) . Nach [R'] IV (1.13) gibt es dann eine Teilfolge von $h_n^1h_n^2 - h_n$, die m_K-fast gegen 0 konvergiert. Nach (3.1) konvergiert diese Teilfolge jedoch m_K-fast gegen $p(A_1,\omega)p(A_2,\omega) - p(A_1.A_2,\omega)$. Damit ergibt sich: $p(A_1,\omega)p(A_2,\omega) - p(A_1.A_2,\omega) = 0$ m_K-fast.
Das war zu beweisen.

§5 Die Überprüfung von Konzeptionen

(5.1) Die Überprüfung der Konzeption K der Gleichartigkeit

Die Konzeption der Gleichartigkeit von Experimenten hat die Konvergenz der relativen Häufigkeiten mit quantitativer logischer Wahrscheinlichkeit Eins zur Folge.
Stellen wir in der Natur ein statistisches Verhalten der Phänomene fest, - praktische Konstanz der relativen Häufigkeiten - , so wird diese Konzeption bestätigt.
Die Konzeption K wird erschüttert, wenn die relativen Häufigkeiten keine praktische Konstanz zeigen.
Wird eine Konzeption $\overline{K}$ sichtbar, die "besser" mit der Erfahrung in Einklang steht, als die Konzeption K , so wird die Konzeption K zu Gunsten der Konzeption $\overline{K}$ verworfen.
Spielt bei der Entscheidung zwischen K und $\overline{K}$ nur die Konstanz der relativen Häufigkeiten eine Rolle, so lassen sich aus der

unter (3.4b) abgeleiteten "BERNOULLI'schen Ungleichung" logische Irrtumswahrscheinlichkeiten für die Verwerfung der Konzeption K berechnen; das heißt, auf Grund der Relation (3.4b) läßt sich eine quantitative logische Wahrscheinlichkeitsschranke dafür angeben, daß die Konzeption K verworfen wird, obwohl sie richtig ist.
Im Allgemeinen liegen jedoch die Verhältnisse beim Vergleich zwischen K und $\overline{K}$ komplizierter; Genaueres dazu siehe (10.1K) .

(5.2) Beispiel für die Aufgabe der Konzeption der Gleichartigkeit

Machen wir die Beobachtung, daß beim sukzessiven Wurf eines Würfels nach jedem 9-ten Wurf die 6 erscheint, während sonst die 1 auftritt, so werden wir die Konzeption K der Gleichartigkeit dieser Würfe fallen lassen und zu einer Konzeption K_{neu} übergehen, die besagt, daß das Wurfexperiment einen Mechanismus enthält, der die beobachtete Regelmäßigkeit liefert.
Innerhalb unseres Formalismus' ist die Konzeption K_{neu} dadurch charakterisiert, daß sie nur die Maßfunktion m_{neu} enthält, die der obigen Regelmäßigkeit das Maß Eins zuordnet.

Zu diesem Beispiel ist zu bemerken, daß die Konzeption K_{neu} ebenfalls die Konvergenz der relativen Häufigkeiten zur Folge hat, so daß die in (5.1) formulierten Bedingungen für die Verwerfung von K nicht vorliegen. Die formale Erfassung dieses Beispiels ergibt sich ebenfalls nach (10.1K) .

Es wäre einfach eine Konzeption K'_{neu} zu konstruieren mit der Maßfunktion m'_{neu} , die einer Ergebnisfolge das Maß Eins zuordnet, in der die relativen Häufigkeiten nicht konvergieren. Das Entscheidende an diesen Beispielen ist jedoch nicht die Konvergenz oder Nichtkonvergenz der relativen Häufigkeiten, sondern die Tatsache, daß einer bestimmten Ergebnisfolge die logische Wahrscheinlichkeit Eins zugeordnet wird, während auf Grund der Konzeption K das Maß dieser Ergebnisfolge für jede Maßfunktion m_K unter einer quantitativen logischen Wahrscheinlichkeitsschranke liegt, die bei entsprechend langer Beobachtungsdauer beliebig klein gemacht werden kann.

Im Gegensatz zu diesen deterministischen Konzeptionen, lassen sich auch Konzeptionen konstruieren, wie etwa die eines Würfels mit zeitlich variierendem Schwerpunkt, wo die Überlegungen von (5.1) anwendbar sind; wie schon bemerkt, werden diese Überlegungen in (10.1K) formal präzisiert.

Der Begriff der statistischen Wahrscheinlichkeit ist nur innerhalb der Konzeption der Gleichartigkeit erklärt; wird diese Konzeption aufgegeben, so ist die statistische Wahrscheinlichkeit nicht mehr definiert. Dies ist insbesondere dann ganz natürlich, wenn die Phänomene kein statistisches Verhalten zeigen und deshalb die Konzeption der Gleichartigkeit aufgegeben werden muß.

Es gibt natürlich auch nicht-gleichartige Konzeptionen, in denen trotzdem der Limes der relativen Häufigkeiten mit logischer Wahrscheinlichkeit Eins existiert. Die Untersuchung solcher Konzeptionen führt jedoch über den Rahmen dieser Arbeit hinaus.

Es wird jedoch vermutet, daß man unter solchen Konzeptionen den Limes der relativen Häufigkeiten nicht mehr als statistische Wahrscheinlichkeit bezeichnen kann, da gewisse der folgenden Eigenschaften zu fehlen scheinen:

(1) Die Axiome von MISES sind mit quantitativer logischer Wahrscheinlichkeit Eins erfüllt; siehe dazu (10.3I) .

(2) Wenn bekannt ist, daß der Limes der relativen Häufigkeiten den Wert p besitzt, ist p die quantitative logische Wahrscheinlichkeit eines Ereignisses $E(\lambda)$; siehe dazu (3.7).

(3) Für die Unterkonzeption K* von K konvergieren die die bedingten subjektiven Wahrscheinlichkeiten gegen den Limes der relativen Häufigkeiten; siehe dazu (3.8) .

Wie schon bemerkt,führen genauere Untersuchungen zu diesen Fragen über den Rahmen dieser Arbeit hinaus.

(5.3) Die Überprüfung der Konzeptionen K_p

Die Konzeption K_p besagt, daß die relativen Häufigkeiten mit quantitativer logischer Wahrscheinlichkeit Eins gegen p konvergieren.
Zeigen die relativen Häufigkeiten eine praktische Konstanz um den Wert p, so ist die Konzeption K_p bestätigt.
Differieren die relativen Häufigkeiten bei langer Beobachtungsdauer erheblich von dem Wert p, so ist die Konzeption K_p erschüttert.
Für den allgemeinen Fall der Verwerfung der Konzeption K_p zu Gunsten einer Konzeption $\overline{K}$ verweisen wir wieder auf (10.1K).

Kommt es bei der Entscheidung zwischen K_p und $\overline{K}$ nur auf Annäherung der relativen Häufigkeiten an den Wert p an, so läßt sich aus der BERNOULLI'schen Ungleichung (3.4a) eine logische Irrtumswahrscheinlichkeit für die Verwerfung der Konzeption K_p berechnen. Dies ist insbesondere dann der Fall, wenn von vornherein feststeht, daß die Konzeption K vorliegt, dh. wenn $\overline{K}$ ebenfalls eine Unterkonzeption von K ist.

§6 SCHLUßBEMERKUNG ZUM BEGRIFF DER KONZEPTION

Der Begriff der Konzeption in der vorliegenden Theorie stellt einen Grundbegriff dar, der für den formalen Aufbau der Theorie nicht weiter,- als durch die angegebenen Axione -, präzisiert werden muß.
Für die Anwendung und die erkenntnistheoretische Fundierung des Systems sind jedoch noch einige Erläuterungen zur inhaltlichen Beschaffenheit dieses Begriffs nötig.
In der vorliegenden Arbeit tritt der Begriff der Konzeption im Wesentlichen nur in der Begriffsbildung gleichartiger Experimente auf. Es lassen sich zwar auch andere Konzeptionen konstruieren, wie etwa in (5.2) skizziert; nähere Untersuchungen dazu wollen wir jedoch nicht vornehmen.
Da der statistische Wahrscheinlichkeitsbegriff erst auf der Basis der Konzeption gleichartiger Experimente eingeführt werden soll,

dürfen die theoretischen Vorstellungen T, die der Konzeption K zu Grunde liegen, den statistischen Wahrscheinlichkeitsbegriff noch nicht enthalten.
Wie schon dargelegt, ist ein Beispiel für die Konzeption K etwa dadurch gegeben, daß T die Vorstellungen der Newton'schen Mechanik und I gewisse Informationen über die Anfangsbedingungen beim Wurf eines Würfels sind.
Ein anderes Beispiel sind die Vorstellungen T der Quantenmechanik über die Möglichkeit, wie genau Elementarteilchen, etwa polarisierte Photonen, gemessen werden können.
Nach §5 besteht das folgende wesentliche Ergebnis:

DIE KONZEPTION DER GLEICHARTIGKEIT LÄßT SICH IM HINBLICK AUF IHRE WAHRSCHEINLICHKEITSTHEORETISCHEN FOLGERUNGEN DURCH BEOBACHTUNGEN ÜBERPRÜFEN: DAMIT IST DIESER BEGRIFF WEITER PRÄZISIERT.

Es bietet sich an, den Begriff der Konzeption in einem etwas engeren, oder besser gesagt, modifizierten Sinne zu gebrauchen, als es in dieser Arbeit der Fall ist.
In dieser Arbeit wurde der Begriff, Konzeption K, als Zusammenfassung gewisser theoretischer Vorstellungen T und gewisser Informationen I gebraucht, wofür wir auch die Bezeichnung wissenschaftliche Oberhypothese anstatt Konzeption gebrauchen könnten; es wurde dabei ausdrücklich hervorgehoben, daß diese Oberhypothese den statistischen Wahrscheinlichkeitsbegriff noch nicht enthalten soll.
Es bietet sich nun an, die etwa aus der wissenschaftlichen Oberhypothese der Gleichartigkeit folgende Wahrscheinlichkeitsstruktur oder ganz allgemein Wahrscheinlichkeitsstrukturen selbst, annalog der hier entwickelten Art, als Konzeptionen oder genauer als Wahrscheinlichkeitskonzeptionen zu bezeichnen.

In diesem Sinne scheint es auch adäquater zu sein, die Formulierung, Konzeption der Gleichartigkeit, auf die durch die Axiome 2 bzw. 5 beschriebene Wahrscheinlichkeitstruktur selbst und nicht die dieser Struktur vorausgehenden wissenschaftliche Oberhypothese zu beziehen.

Festhalten wollen wir nochmals, daß eine derartige Wahrscheinlichkeitskonzeption ein wissenschaftlicher Ansatz oder eine wissenschaftliche Hypothese darstellt, die durch Beobachtungen überprüfbar ist und sowohl erhärtet als auch verworfen werden kann.
Dieser engere Gebrauch des Worts Konzeption für die jeweilige Wahrscheinlichkeitskonzeption selbst, anstatt für die diese Konzeption implizierende wissenschaftliche Obehypothese, scheint insbesondere bei dem weiteren Ausbau der hier entwickelten Gedanken vorzuziehen zu sein.
Dabei ist eine Wahrscheinlichkeitskonzeption in unserem Sinne im Gegensatz zum üblichen Gebrauch nicht durch eine Wahrscheinlichkeitsfunktion charakterisiert, sondern durch eine Klasse von Maßfunktionen zusammen mit den zugeordneten Begriffen der quantitativen logischen Wahrscheinlichkeit, bzw. den quantitativen logischen Wahrscheinlichkeitsschranken.

§7 Die Anwendung der Theorie auf die Quantenmechanik

In der Quantenmechanik tritt die Konzeption der Gleichartigkeit mit besonderer Stringenz auf: Wir betrachten das zu Anfang gegebene Beispiel einer gleichartig polarisierten Folge von Photonen, die auf ein Polarisationsfilter treffen.
Das Auftreffen der einzelnen polarisierten Photonen auf das Filter sind zeitlich verschiedene Prozesse, die nach den Vorstellungen der Quantenmechanik gleichartig sind, da sie durch Messungen nicht weiter differenziert werden können:
Es gibt keine Messung an einem polarisierten Photon, die etwas darüber besagen würde, ob dieses Photon das Filter eher passiert als ein anderes in gleicher Weise polarisiertes Photon.
Nehmen wir weiterhin an, daß die einzelnen Photonen in großem zeitlichem Abstand, etwa jede Stunde ein Photon, auf das Filter treffen. Dann ist es klar, daß der Durchgang von n Photonen dieser Folge logisch gleichwahrscheinlich ist zu dem Durchgang von n anderen Photonen aus dieser Folge:

Der zeitliche Abstand der einzelnen Photonen in dieser Folge ist so groß, daß die zeitliche Topologie der Photonen keine Rolle mehr spielt, was den Durchgang durch das Filter anbelangt.

Treffen die polarisierten Photonen in enger zeitlicher Nachbarschaft auf das Filter, so ist es zwar a-priori nicht klar, daß der Durchgang eines Photons durch das Filter nicht vom Durchgang der anderen Photonen beeinflußt wird und müßte genau genommen erst empirisch erhärtet werden. -Wenn wir jedoch annehmen, daß das Auftreffen der Photonen auf das Filter isolierte physikalische Prozesse darstellen, so ergibt sich aus dieser Vorstellung ebenfalls, daß der Durchgang von n Photonen logisch gleichwahrscheinlich ist zum Durchgang von n anderen Photonen:
Die physikalische Beschaffenheit des Filters und damit der Durchgang durch das Filter wird nach dieser Vorstellung nicht dadurch beeinflußt, wie benachbart die einzelnen Photonen sind.
Nach den in dieser Weise skizzierten Vorstellungen T der Quantenmechanik stellt also der Durchgang polarisierter Photonen durch ein Filter eine Klasse gleichartiger Prozesse dar, die Axiom 2 von (1.3) erfüllen.
Aus den weiterhin formulierten Axiomen über die aus der Gleichartigkeit folgenden Wahrscheinlichkeitstruktur ergibt sich, daß diese Phänomene,da sie gleichartig sind, ein statistisches Verhalten zeigen müssen.
Die vorstehende Theorie stellt also einen logischen Zusammenhang her zwischen der in der Quantenmechanik gegebenen theoretischen Struktur der Gleichartigkeit der Phänomene und dem empirischen Befund ihres statistischen Verhaltens:

Das statistische Verhalten quantenmechanischer Phänomene erklärt sich aus ihrer Gleichartigkeit.
Umgekehrt ermöglicht das statistische Verhalten der Phänomene die in der Quantenmechanik bestehende Begriffsbildung der physikalischen Gleichartigkeit dieser Phänomene.

§8 Die Anwendung der Theorie auf die Klassische Physik

(8.1) Charakterisierung der unmittelbaren Klassischen Situation

Im Gegensatz zur Quantenmechanik ist die Konzeption der Gleichartigkeit von Experimenten in der Klassischen Physik kein absoluter Begriff, sondern hängt davon ab, wie genau die betreffenden Messungen vorgenommen werden.
Durch genaue Messungen beim Wurf eines Würfels läßt sich das jeweilige Ergebnis bei einem solchen Wurf genau voraussagen.
Die Vorstellung der Gleichartigkeit einer Klasse von Würfelwürfen beruht auf einer gleichartigen unvollständigen Information über die einzelnen Wurfexperimente.
Während in der Quantenmechanik das auftretende statistische Verhalten aus einer absoluten,in der Natur bestehenden Gleichartigkeit erklärt werden konnte, kann im Fall klassischer Experimente nur davon gesprochen werden, daß wir auf Grund einer gleichartigen Information mit einem statistischen Verhalten zu rechnen haben.
Die Gleichartigkeit der Experimente ist weniger eine Aussage über die Experimente selber, als vielmehr eine Aussage über unsere Information von diesen Experimenten.
Die vorstehende Theorie zeigt so gesehen nur die rationalen Konsequenzen auf, die aus einer gleichartigen Information zu ziehen sind.

(8.2) Die Zurückführung auf die Quantenmechanik

Es bietet sich an, die Gleichartigkeit makrophysikalischer Experimente auf die absolute Gleichartigkeit der Quantenmechanik zurückzuführen und damit ebenfalls als eine absolute,in der Natur bestehende Struktur aufzufassen.
Der unterschiedliche Ausgang beim Wurf eines Würfels hängt ab von mikrobiologischen Vorgängen, die den Wurf des Würfels mit der Hand steuern. Diese letztlichen mikrobiologischen Vorgänge, die das Ergebnis beim Wurf eines Würfels determinieren, stellen nach den Vorstellungen der Quantenmechanik eine in objektiver Weise gleichartige Gesamtheit dar. Damit ließe sich die Gesamtheit der Würfelwürfe ebenfalls als eine gleichartige Gesamtheit auf Grund einer objektiven,in der Natur bestehenden Struktur verstehen.

§9 Vergleich mit den klassischen Systemen der Wahrscheinlichkeitstheorie.

(9.1) Vergleich mit dem System von de Finetti

Der vorstehende Aufbau ähnelt in seiner mathematischen Struktur dem System von de Finetti [F] .
a) Allerdings wird hier der Begriff der subjektiven Wahrscheinlichkeit nur als Vorstellungshilfe zugelassen, die grundsätzlich entbehrt werden könnte und durch den Begriff der auf Grund von K möglichen Maßfunktion zu ersetzen ist; siehe (1.11) .
So wertvoll der Begriff der subjektiven Wahrscheinlichkeit unter einem heuristischen Standpunkt auch sein mag, kann ihm jedoch keine selbsständige Bedeutung zugestanden werden.
Wir behaupten nämlich, daß das System von de Finetti, so wie es von de Finetti konzipiert wurde, gar nicht anwendbar ist.
Betrachten wir zum Beispiel den radioaktiven Zerfall. Dann wird ein Wissenschaftler J die Konzeption K der Gleichartigkeit annehmen. Er wird weiterhin auch die von uns gewonnenen Gesetze annehmen, die bei jeder möglichen "subjektiven" Interpretation der komparativen Struktur gelten. Er wird jedoch im Allgemeinen keine spezielle subjektive Bewertung als für sich verbindlich angeben können. Mit welcher Begründung sollte er auch eine spezielle subjektive Bewertung auswählen. Er könnte sich für eine beliebige entscheiden, da unsere Theorie für jede denkbare de Finett-Bewertung gilt; aber das wäre reine Willkür.*)
Wir behaupten also, daß das de Finetti-System erst anwendbar wird durch die von uns durchgeführte Konstruktion, in der nur die Aussagen betrachtet werden, die für jede "subjektive" Bewertung gelten, ohne daß dabei behauptet wird, daß es reale Individuen gäbe, die eine derartige subjektive Bewertung hätten.
Wie in (1.11) dargelegt,ist der Begriff der subjektiven Bewertung bei uns eine reine Hilfsvorstellung, die zur Motivierung unserer Axiome nützlich ist, in denen jedoch der Begriff der sub-

) Zu dem Wort beliebig ist zu bemerken, daß sinnvollerweise nur solche Maßfunktionen gewählt werden könnten, die den Wert P(E) nicht einschränken, also insbesondere alle Maßfunktionen der Konzeption K ; im anderen Fall würde eine,die möglichen Werte von P(E) einschränkende Maßfunktion eben bedeuten, daß der Wissenschaftler J von vornherein gewisse Werte für P(E) ausschließt.

jektiven Bewertung letztlich eliminiert ist.
Auch wenn man den Begriff der subjektiven Wahrscheinlichkeit zuließe, wäre es unbefriedigend, wenn wir die statistische Gesetzmäßigkeit etwa beim radioaktiven Zerfall wie folgt zu beschreiben versuchten: Jeder Beobachter des Zerfallsexperiments hat eine DE FINETTI-Bewertung. Es ist befriedigender, wenn wir das Zerfallsexperiment im Rahmen unserer Theorie wie folgt beschreiben können: Die Zerfallsprozeße sind gleichartig,- es bestehen daher die von uns angegebenen, objektiv erfaßten Konsequenzen.
Während das DE FINETTI-System auf einen fiktiven Beobachter rekurrieren muß, um die statistische Gesetzmäßigkeit zu beschreiben, liefert unser System ein objektives begriffliches Vokabular, in dem diese Gesetzmäßigkeit erfaßt ist.

b) Die vorstehende Theorie unterscheidet sich scharf von dem System von DE FINETTI durch die in (1.10) theoretisch berücksichtigte und in (5.2) an einem Beispiel illustrierte Möglichkeit, bei entsprechender Information die Konzeption der Gleichartigkeit aufzugeben.
Dies ist im System von DE FINETTI unmöglich, da die Verwertung einer Information immer nur durch den Übergang zu den bedingten Wahrscheinlichkeiten erfolgt, einen Standpunkt, den wir gemäß (5.2) für unzureichend halten: Nach unserer Auffassung stellt sie Konzeption der Gleichartigkeit einen wissenschaftlichen Ansatz dar, der der Überprüfung durch Beobachtungen fähig ist, während diese Konzeption bei DE FINETTI als eine subjektive Gegebenheit erscheint, an der sich nichts mehr ändern läßt.

(9.2) Vergleich mit dem System von CARNAP

Die vorstehende Arbeit ähnelt im Begriff der logischen Wahrscheinlichkeit dem System von CARNAP [C].
a) Wie die vorstehende Arbeit zeigt, implizieren die Axiome [C] von CARNAP das statistische Verhalten der Phänomene.
Diese Axiome besitzen also einen empirischen Gehalt und können daher nicht als a-priorische Eigenschaft eines formalen Spachsystems

aufgefaßt werden, sondern müssen auf eine wissenschaftliche Konzeption gegründet werden.
Wie bei DE FINETTI ist es auch im System von CARNAP unmöglich, die Konzeption der Gleichartigkeit bei entsprechender Beobachtung aufzugeben. Wie in (9.1b) dargestellt halten wir diesen Standpunkt für unhaltbar.

b) CARNAP ist daran interessiert, die in (1.5) und (2.3) eingeführte Maßfunktion m_K eindeutig feszulegen.
Es gelingt ihm auch durch einsichtige Axiome, die Klasse der Maßfunktionen auf eine Klasse einzuschränken, die nach dem DE FINETTI-Theorem (2.3) durch die folgenden Maße auf dem Simplex D charakterisiert ist:

$$d\psi = c(x_1 \ldots x_k)^{\alpha} dL_{k-1}$$; dabei ist L_{k-1} das

(k-1)-dimensionale Lebesguemaß auf dem Simplex D , c ist die Normierungskonstante und α hängt mit dem CARNAP'schen λ wie folgt zusammen: $\alpha = \lambda/k - 1$; siehe [H_o] (6.4) .

Den Versuchen CARNAPS durch Spezialisierung von λ zu einer eindeutig bestimmten Maßfunktion zu gelangen, war bislang kein voller Erfolg beschieden.

(9.3) Vergleich mit dem System von MISES

Der vorstehende Aufbau ähnelt dem System von MISES [M] insofern, als die statistische Wahrscheinlichkeit ebenfalls als Limes der relativen Häufigkeiten eingeführt wird.
Im Gegensatz zu MISES gelten in unserem Aufbau die MISES'schen Axiome nicht in Strenge sondern nur mit quantitativer logischer Wahrscheinlichkeit Eins.*)
An dem System von MISES üben wir die folgende Kritik:

a) Im System von MISES wird das Phänomen statistischer Gesetzmäßigkeiten nur beschrieben, aber nicht wie in unserem System als Folgerung aus der Konzeption der Gleichartigkeit erklärt.
Im System von MISES wird keine Bedingung angegeben,- wie etwa

*) Genauer: Es gelten in unserem System mit quantitativer logischer Wahrscheinlichkeit Eins die von ihrem ursprünglichen, trivialen Widerspruch befreiten MISES-Axiome; siehe (10.3I) .

die Bedingung der Gleichartigkeit der Experimente -, um zu entscheiden, wann eine Experimentenfolge die Eigenschaft der Konvergenz im Unendlichen aufweist.

b) Der im System von MISES existierende Limes der relativen Häufigkeiten hat nicht automatisch die Bedeutung eines fairen Wettquotienten für jedes Glied der Folge. Diese Deutung ist jedoch für die praktische Anwendung der Wahrscheinlichkeitstheorie unabdingbar. Diesem Mangel könnte abgeholfen werden, in dem man das folgende zusätzliche Axiom dem System von MISES beifügt:

> Axiom: Ist p der Limes der relativen Häufigkeiten in einem MISES'schen Kollektiv, so stellt p für jedes Glied dieses Kollektivs den fairen Wettquotienten auf diese Glied dar.

Allerdings ist dazu zu bemerken, daß sich dieses Axiom nur rechtfertigen ließe, wenn man voraussetzt, daß die MISES'sche Ereignisfolge in irgend einem intuitiven Sinne gleichartig ist. Im anderen Falle wäre dieses Axiom nämlich geradzu falsch. Es ist nämlich leicht vorstellbar, daß eine Ereignisfolge die MISES-Axiome erfüllt, die ja nur etwas über das Verhalten im Unendlichen besagen, ohne daß diese Folge in einem intuitiven Sinne gleichartig ist, insbesondere ohne daß der Limes der relativen Häufigkeiten als fairer Wettquotient auf jedes dieser Ereignisse interpretiert werden kann.

c) Das System von MISES liefert keine Aussage über die Geschwindigkeit der Konvergenz der relativen Häufigkeiten und hat damit keinerlei praktische Relevanz, da allein das Verhalten im Endlichen von praktischer Bedeutung ist.
Wir bemerken dazu, daß natürlich auch im System von MISES die BERNOULLI'sche Ungleichung ableitbar ist. Im System von MISES besagt diese Ungleichung aber ebenfalls nur etwas über das Verhalten im Unendlichen eines aus dem Ausgangskollektiv abgeleiteten Kollektivs und gibt uns keinerlei Information über das Verhalten im Endlichen. Erst wenn wir, wie unter b) ausgeführt, dem System von MISES das Axiom der Deutung des Limes als fairen Wettquotienten jedes Ereignisses des Kollektivs anfügen, läßt die BERNOULLI'sche Ungleichung im System von MISES eine Interpretation zu, die zu einer Aussage über die Geschwindigkeit der Konvergenz führt.

d) In [M] charakterisiert MISES den Anwendungsbereich seiner Axiome zwar ausdrücklich durch den Gebrauch des Worts "gleichartige" Ereignisse, ist sich aber über die entscheidende Bedeutung dieses Begriffs nicht im klaren.
Dazu bemerken wir das Folgende:
Das System von MISES scheint sich wie folgt anwenden zu lassen: Wir beobachten in einer Folge von Ereignissen von einer gewissen Stelle an eine praktische Konstanz der relativen Häufigkeiten und nehmen daraufhin den induktiven Schluß vor, daß sich diese praktische Konstanz bei der weiteren Beobachtung dieser Ereignisfolge zeigen und sogar noch verstärken wird.
Dieser Schluß ist jedoch, wie intuitiv einsichtig ist, nur dann möglich, wenn wir die Ereignisfolge in irgend einem Sinne als "gleichartig" voraussetzen. Im anderen Fall lassen sich sofort Gegenbeispiele konstruieren in Übereinstimmung mit dem HUME'schen Satz, daß ein Schluß vom Beobachteten auf das Unbeobachtete nicht möglich ist; siehe dazu STEGMÜLLER [St] .
Das System von MISES muß also, um praktisch zu funktionieren, in irgend einer Weise von dem Begriff "gleichartig" Gebrauch machen. - Wie diese Arbeit zeigt, ist der Begriff gleichartig, in der von uns durchgeführten Präzisierung, ausreichend, um alles, was im Sinne von MISES gesagt werden muß, ableiten zu können, ohne daß darüberhinaus noch irgendwelche Axiome nötig wären.
Es kommen zwar auch in unserer Arbeit Axiome vor; diese Axiome und die beiden Anwendungsprinzipien sagen jedoch nichts über die Wirklichkeit, sondern sind so zu sehen, daß sie die Implizitdefinition des Begriffs der logischen Wahrscheinlichkeit darstellen, vergleichbar einem axiomatischen Kalkül der Aussagenlogik, dessen Axiome die Implizitdefinition der aussagenlogischen Junktoren darstellt.
Abschließend stellen wir nochmals fest: Das MISES'sche System ist erkenntnistheoretisch so unvollständig, daß es nicht anwendbar ist. Um es anwendbar zu machen, müßte in irgend einem Sinne der Begriff "gleichartig" der Charakterisierung des Kollektivbegriffs hinzugefügt werden. Allerdings ist dieser unverzichtbare Begriff in unserer Präzisierung ausreichend, um das ganze System der statistischen Wahrscheinlichkeitstheorie abzuleiten, ohne daß noch weitere Axiome nötig wären.
Zu unserer Kritik des Systems von MISES siehe auch (10.3H) .

(9.4) Vergleich mit dem System von KOLMOGOROFF

Das System von KOLMOGOROFF [K] führt den Begriff der Wahrscheinlichkeit als einen undefinierten Grundbegriff ein, der gewissen Axiomen gehorcht. Der Zusammenhang mit den relativen Häufigkeiten ergibt sich aus der BERNOULLI-Ungleichung:

$$p(\{|h_n(E) - p(E)| \geq \varepsilon\}) \leq 1/4n\varepsilon^2 \quad .$$

Dieser Aufbau ist in sich konsistent und stellt auch eine gute Beschreibung des Phänomens statistischer Gesetzmäßigkeiten dar. Allerdings enthält dieses System einen erkenntnistheoretischen Zirkel: Die geforderten Axiome werden damit motiviert, daß die Wahrscheinlichkeit den "Grenzwert" der relativen Häufigkeiten darstellen soll.

Man kann also nur davon sprechen, daß die KOLMOGOROFF-Theorie das Phänomen statistischer Gesetzmäßigkeiten beschreibt; sie liefert keine Erklärung dieser Gesetzmäßigkeit, sondern setzt diese Gesetzmäßigkeit, die sie in ihren Axiomen zu erfassen versucht, bereits voraus.

Die Zirkularität des Aufbaus von KOLMOGOROFF zeigt sich besonders anschaulich in der BERNOULLI'schen Ungleichung: Die zu erklärende Wahrscheinlichkeit $p(E)$ wird im Innern der Formel durch die relativen Häufigkeiten approximiert; das Maß der Approximation wird jedoch im Äußern der Formel ebenfalls durch das erst zu erklärende p dargestellt.

Im Gegensatz dazu wird in unserem Analogon zur BERNOULLI'schen Ungleichung der Grad der Approximation der statistischen Wahrscheinlichkeit durch die relativen Häufigkeiten mit der Maßfunktion m_K ausgedrückt, die bereits vorher, unabhängig vom statistischen Wahrscheinlichkeitsbegriff definiert war.

Wie in (3.9) dargestellt, führt unsere Theorie unter der Voraussssetzung der Konzeption der Gleichartigkeit zu einer erkenntnistheoretisch befriedigenden Begründung der KOLMOGOROFF-Theorie.

Wie das System von MISES gebraucht auch KOLMOGOROFF zur Charakterisierung des Anwendungsbereichs seiner Theorie ähnliche Worte wie das Wort gleichartig. - Wie schon bemerkt, ist das Ergebnis dieser Arbeit der Nachweis, daß außer diesem unvermeidbaren Begriff der Gleichartigkeit in der von uns gegebenen Präzisierung nichts mehr zur Begründung der objektiven Wahrscheinlichkeitstheorie vorausgesetzt werden muß.

Siehe dazu auch (10.3H) .

(9.5) Vergleich mit dem System von RICHTER

Das System von RICHTER [R] verbessert das System von KOLMOGOROFF insofern, als die Motivierung der Axiome nicht mehr das Phänomen statistischer Gesetzmäßigkeiten voraussetzt, sondern nur verlangt, daß die im Naturgeschehen als existent postulierte Wahrscheinlichkeit stetigen Rechengesetzen genügt.

Diese Arbeit zeigt, daß sich das von RICHTER eingeführte Postulat der objektiven Wahrscheinlichkeit aus dem Begriff der Gleichartigkeit von Experimenten begründen läßt.

Allerdings besteht bei RICHTER eine gewisse Vagheit, ob sich die objektive Wahrscheinlichkeit auf konkrete Ereignisse oder Ereignisschemata bezieht. Der genaue Bezug unserer Begriffe zu dem System von RICHTER setzt also die Präzisierung des RICHTER'schen Systems in diesem Punkt voraus.

Wie in meiner Diplomarbeit [H_0] gezeigt, ist das System von RICHTER erst dann anwendbar, wenn gewisse statistische A-priori-Hypothesen vorausgesetzt werden, wobei das System von RICHTER die Gründe für die Annahme solcher A-priori-Hypothesen grundsätzlich nicht erbringen kann.

Die einfachste A-priori-Hypothese, unter der das System von RICHTER funktioniert, ist die Hypothese, daß die Experimente $B(\lambda)$ unabhängige Wiederholungen sind.

Diese Hypothese ist in gewisser Weise äquivalent zur Konzeption der Gleichartigkeit *) , aus der allein sich dann, wie in dieser Arbeit gezeigt, der Begriff der statistischen Wahrscheinlichkeit

---

*) Zu der Formulierung, gewisse Äquivalenz zwischen Gleichartigkeit und unabhängiger Wiederholung, wird das Folgende bemerkt: Gemäß dem System von RICHTER ergeben sich unter der Voraussetzung der unabhängigen Wiederholung die DE FINETTI'schen Wahrscheinlichkeiten als Chancen im Sinn von RICHTER. Die unabhängige Wiederholung impliziert also sicher die Gleichartigkeit. Meines Erachtens zeigt die vorliegende Arbeit, daß auch die Umkehrung gilt, nämlich, daß die Konzeption der Gleichartigkeit im Hinblick auf die in der Grenze vorliegende Konzeption $K_{P(E)}$ die unabhängige Wiederholung impliziert; siehe (3.9) . Allerdings lassen sich symmetrische Wahrscheinlichkeitsfunktionen, aufgefaßt als Chancen, im System von RICHTER auch unter wesentlich allgemeineren Hypothesen oder auch gänzlich anderen Voraussetzungen gewinnen, wie etwa des Ziehens aus einer Urne ohne Zurücklegen; siehe §6 von [H_0] . - Deshalb kann nur von einer gewissen Äquivalenz zwischen Gleichartigkeit und unabhängiger Wiederholung gesprochen werden.

oder, falls die Deutung der statistischen Wahrscheinlichkeit als objektive Wahrscheinlichkeit im Sinn von RICHTER akzeptiert wird, der Begriff der objektiven Wahrscheinlichkeit einschließlich aller Rechenregeln ableiten läßt, ohne daß wie bei RICHTER noch zusätzlich zur Konzeption der Gleichartigkeit, das Postulat der objektiven Wahrscheinlichkeit, einschließlich gewisser Axiome, was die Rechenregeln der postulierten Größe betrifft, darüberhinaus gefordert werden müßten.

(9.6) Schlußbemerkung zum System der Wahrscheilichkeitsrechnung von PASCAL , LAPLACE

Die auf PASCAL zurückgehende von LAPLACE in [L] dargestellte Wahrscheinlichkeitsrechnung macht fundamental Gebrauch vom Begriff der Gleichwahrscheinlichkeit, allerdings bezüglich der Ergebnisse innerhalb eines festen Experiments.- In der vorliegenden Arbeit tritt der Begriff der Gleichwahrscheinlichkeit ebenfalls als fundamentaler Begriff auf, allerdings als "Gleichwahrscheinlichkeit" bezüglich der verschiedenen gleichartigen Experimente. Die vorliegende Arbeit scheint also dem ursprünglichen Vorgehen von PASCAL , LAPLACE rechtzugeben, daß der Begriff der Gleichwahrscheinlichkeit, allerdings in einem gegenüber PASCAL und LAPLACE modifizierten Sinne, ein fundamentaler, irreduzibler Begriff darstellt, auf dem das Gebäude der Wahrscheinlichkeitsrechnung ruht.

§10 VERGLEICH MIT EINIGEN MODERNEREN SYSTEMEN DER WAHRSCHEINLICHKEITSTHEORIE

(10.1) Vergleich mit dem System von HACKING [HA]

A) Ziel des Systems von HACKING

HACKING geht von dem System von KOLMOGOROFF aus. Er interessiert sich dabei vor allem für eine Theorie des statistischen Schließens, dh. für einen Begriff der Bestätigung und schließlichen Annahme oder Verwerfung statistischer Hypothesen. Er meint, die von ihm entwickelte Theorie des statistischen Schließens, die seiner Ansicht nach noch nicht in den KOLMOGOROFF-Axiomen enthalten ist, liefere eine weitere begriffliche Klärung des von KOLMOGOROFF intendierten Begriffs der long-run-Wahrscheinlichkeit, bei HACKING Chance genannt. *)

B) Der Begriff der joint proposition und des statistischen Datums

Ein zentraler Begriff HACKINGs ist der Begriff des statistischen Datums. Als statistisches Datum bezeichnet HACKING dabei eine sogenannte joint proposition. Eine joint proposition bezieht sich dabei auf eine experimentelle Anordnung X , Ereignisse von der Art K bei X , eine Klasse D von Verteilungsfunktionen bezüglich der Ereignisse von der Art K , sowie ein Ereignis A beim Versuch T von der Art K' . Als joint proposition bezeichnet HACKING dabei das folgende 6-Tupel: (X,K,D;T,K',A) .
Von einer einfachen joint proposition sprechen wir, wenn D nur ein Element enthält, ansonsten sprechen wir von einer komplexen joint proposition.
Zwischen dem Begriff des statistischen Datums und unserem Begriff der Konzeption besthet eine gewisse Analogie:

*) Es sei ausdrücklich darauf hingewiesen, daß der Begriff Chance von HACKING etwas anderes bedeutet als der Begriff Chance bei RICHTER: Bei HACKING bedeutet Chance die statistische oder objektive Wahrscheinlichkeit, während bei RICHTER die Chance subjektive Wahrscheinlichkeit bedeutet.

Unter einer Konzeption verstehen wir theoretische Vorstellungen T und Informationen I . Die theoretischen Vorstellungen T entsprechen dem ersten Teil X,K,D der joint proposition, die Information I dem zweiten Teil T,K',A ; die durch die Konzeption K festgelegten Maßfunktionen m_K entsprechen den Verteilungsfunktionen aus D .
Als likelihood einer einfachen joint proposition h:= (X,K,D;T,K',A) bezeichnen wir die Chance von A auf Grund der Verteilungsfunktion D : likelihood(h):= $p_D(A)$.

C) Likelihoodgesetz

Unter einer statistischen Hypothese H verstehen wir die Aussage, daß die wahre Verteilungsfunktion in einer Klasse D von Verteilungsfunktionen liegt. Von einer einfachen Hypothese sprechen wir, wenn D aus einem Element besteht, ansonsten sprechen wir von einer komplexen Hypothese.
Angenommen wir haben zwei einfache Hypothesen H und H' .
Auf Grund der Hypothese H habe das Ereignis A die Chance p(A) , auf Grund von H' die Chance p'(A) .
Dann ist die Hypothese H auf Grund von A besser bestätigt als H' , genau dann wenn p(A) > p'(A) .
Dieses Prinzip der Bestätigung ist ein Spezialfall dessen, was HACKING Gesetz der Likelihood nennt; das allgemeine Gesetz der Likelihood von HACKING wollen wir hier nicht wiedergeben.
Das Gesetz der Likelihood ist der Kern der Bestätigungsaxiome, um die HACKING das KOLMOGOROFF-System erweitert, um zu einer vollständigen inhaltlichen Klärung des Begriffs Chance zu kommen.

D) Statistische Tests. Der Likelihood-Test.

Unter einem statistischen Test versteht HACKING die Annahme oder Verwerfung einer statistischen Hypothese.
Sein zentraler Begriff ist der des Likelihood-Tests.
Darunter versteht er das Folgende:
Besteht die Wahl zwischen den einfachen Hypothesen H und H' , dann wird auf Grund des Ereingisses A die Hypothese H akzeptiert und H' verworfen, wenn gilt $p'(A)/p(A) < \alpha$; der Wert α muß dabei ein "plausible" vorgegebene Größe sein.

Sind H und H' komplexe Hypothesen, so gilt das Folgende:
Die Hypothese H' wird beim Eintritt des Ereignisses A verworfen zu Gunsten von H , wenn für jede einfache Hypothese h' aus H' eine einfache Hypothese h aus H existiert, so daß gilt: $p_{h'}(A)/p_h(A) < \alpha$.
HACKING formuliert diese Regeln noch zusätzlich für joint propositions; darauf wollen wir hier verzichten.
HACKING betont, daß es eine Verwerfung von statistischen Hypothesen immer nur im Hinblick auf Alternativhypothesen gibt. Dabei kann es sein, daß auch die A-priori-Hypothese, dh. die Disjunktion von H und H' , in Zweifel gezogen wird und nun ihrerseits einem Test gegen eine neue Alternativhypothese unterworfen wird.

E) Das statistische Datum D_B der Binomialverteilung

Das wichtigste statistische Datum ist das der Binomialverteilung, bzw. der Polynomialverteilung; der Einfachheit halber beschränken wir uns hier auf die Binomialverteilung.
Die Binomialverteilung entsteht nämlich dann, wenn wir ein Experiment mit zwei möglichen Ereignissen unabhängig wiederholen.
Außerdem können wir, wie wir noch sehen werden, das statistische Datum der Binomialverteilung , das wir hier mit D_B bezeichnen wollen, in Beziehung zur Konzeption der Gleichartigkeit setzen.
Dem Begriff der experimentellen Anordnung X entspricht in unserer Terminologie das Versuchsschema B , dem Begriff Ereignisse von der Art K entsprechen die Ereignisse E(λ) ; dem Tripel T,K',A entspricht ein aus den E(λ) aufgebautes Ereignis A . Die Klasse D_B der Binomialverteilungen stellen wir mit den Bezeichnungen von (1.5) wie folgt dar:
$D_B = \{p_x | K : x \in [0,1]\}$.

F) Die theoretische Beziehung zwischen D_B und der Konzeption K

a) Zunächst können wir feststellen, daß die Konzeption K der Gleichartigkeit formal dem statistischen Datum D_B insofern entspricht, als die Konzeption K gemäß (1.5) aus den Mischungen der Binomialverteilungen bezüglich der subjektiven Maße ϕ besteht.

b) Weiterhin können wir das statistische Datum D_B wie folgt mit der Konzeption K der Gleichartigkeit in Beziehung setzen: Fügen wir nämlich zu der Konzeption K die Information $P(E) = x_o$ hinzu, so entsteht nach (3.7) die Konzeption K_{x_o}.
Fügen wir zu dem statistischen Datum D_B die Information $\lim_{n\to\infty} h_n(E) = x_o$ hinzu, so ergibt sich nach dem Likelihood-Test, wie wir gleich sehen werden, die Binomialverteilung zum Parameter x_o, also ebenfalls die Konzeption K_{x_o}.

Beweis:
Wir gehen von der A-priori-Hypothese D_B aus. Die Hypothese H ist gegeben durch die Funktion $p_{x_o}|K$, also $H := \{p_{x_o}\}$, die Hypothese H' ist wie folgt definiert: $H' := \{p_x : x \neq x_o\}$.
Das Ereignis A ist definiert als: $A := \{\omega : \lim_{n\to\infty} h_n(E,\omega) = x_o\}$.

Es gilt nun: $p_x(A)/p_{x_o}(A) = 0/1 = 0 < \alpha$ für jedes positive α, sofern p_x aus H'.

Das heißt: Auf Grund der Information A wird gemäß dem Likelihood-Test die Hypothese H' verworfen zu Gunsten der Hypothese H und zwar für jedes α. Das war zu beweisen.

Es besteht also eine gewisse Gleichwertigkeit zwischen der Konzeption K und dem statistischen Datum D_B. Allerdings ist diese Beziehung nur von theoretischer Bedeutung, da es unmöglich ist, die Information $\lim_{n\to\infty} h_n(E) = x_o$ zu besitzen.

G) Die praktische Beziehung zwischen D_B und K. Der Konfidenz-Test.

a) Wir setzen das statistische Datum D_B voraus. Dann wird der praktische induktive Prozeß der objektivistischen Statistik am einleuchtendsten durch den Konfidenz-Test beschrieben.
Darunter verstehen wir die folgende induktive Regel:
Tritt bei n Beobachtungen die relative Häufigkeit h_n auf, so wird behauptet, daß der wahre Parameter $\hat{x}$ der Binomialverteilung in dem folgenden Intervall liegt: $\hat{x} \in [h_n - \varepsilon, h_n + \varepsilon]$.

Das statistische Datum D_B geht also nach dieser Regel auf Grund von n Beobachtungen in die folgende Hypothese über:
$D_{B,h_n,\varepsilon} := \{p_x : x \in [h_n - \varepsilon, h_n + \varepsilon]\}$.

Als Irrtumswahrscheinlichkeit zu diesem Test bezeichnen wir die folgende Größe: $\alpha(x) :=$ Chance dafür, daß die Hypothese $D_{B,h_n,\varepsilon}$ falsch ist, unter der Voraussetzung, daß x der wahre Parameter der Binomialverteilung ist $= p_x(\{\omega : |h_n(E,\omega) - x| > \varepsilon\})$.

Es gilt:
$\alpha(x) = 1 -$ Chance dafür, daß der Konfidenz-Test zu einem richtigen Ergebnis führt, unter der Voraussetzung, daß x der wahre Parameter ist.
Die Größe $\alpha(x)$ läßt sich nach der BERNOULLI'schen Ungleichung abschätzen: $\alpha(x) \leq \alpha := 1/4n\varepsilon^2$.

Der Konfidenz-Test ist also durch drei Parameter charakterisiert: Den Stichprobenumfang n, die Irrtumsschranke α und die Schärfe ε. Sind zwei von diesen Größen vorgegeben, so läßt sich die dritte abschätzen.
Je kleiner die Irrtumsschranke α ist, desto sicherer ist es, daß der Konfidenz-Test zu einem richtigen Ergebnis führt; dabei ist natürlich, wie anfangs festgelegt, immer vorausgesetzt, daß D_B überhaupt richtig ist.

Bei jeder, noch so kleinen Irrtumsschranke α läßt sich durch Wahl eines hinreichend großen n ein immer kleineres Intervall $[h_n - \varepsilon, h_n + \varepsilon]$ aussondern.

Der Konfidenz-Test liefert also ein Verfahren, nach dem das statistische Datum D_B durch sukzessiven Übergang zu $D_{B,h_n,\varepsilon}$ immer mehr gegen die unbekannte wahre Hypothese $\{p_{\hat{x}}\}$ konvergiert.

In dem wir das starke Gesetz der großen Zahlen heranziehen, können wir dies wie folgt symbolisieren:
$p_x(D_{B,h_n,\varepsilon} \to \{p_x\}) = 1$. Wesentlich ist dabei, daß über das "Ausmaß" dieser Konvergenz bereits im Endlichen mit Hilfe der Irrtumsschranke eine Aussage gemacht werden kann.

b) Vom Standpunkt unseres neuen Aufbaus der Statistik läßt sich das Verfahren des Konfidenz-Tests rekonstruieren:
Nach Beobachtung der relativen Häufigkeit h_n gehen wir von der Konzeption K zu der Konzeption $K_{h_n,\varepsilon} := K \wedge (P(E) \in [h_n-\varepsilon, h_n+\varepsilon])$ über.
Wie man sich leicht überlegt (siehe dazu (3.7a)) besteht die Konzeption $K_{h_n,\varepsilon}$ aus den Mischungen der Binomialverteilungen von $D_{B,h_n,\varepsilon}$ bezüglich solcher Maße ϕ, die ihr gesamtes Gewicht auf $[h_n-\varepsilon, h_n+\varepsilon]$ konzentrieren.

Der Aussage, $P(E) \in [h_n-\varepsilon, h_n+\varepsilon]$, läßt sich auf Grund des Analogons zur BERNOULLI'schen Ungleichung eine quantitative logische Wahrscheinlichkeitsschranke zuordnen, die als Irrtumsschranke fungiert: $m_K(|h_n - P(E)| \geq \varepsilon) \leq 1/4n\varepsilon^2 =: \alpha$.
Gestützt auf (3.1) können wir also analog zu a) feststellen:

$$l_K(K_{h_n,\varepsilon} \to K_{P(E)}) = 1 .$$

Wesentlich ist dabei wieder, daß sich das "Ausmaß" dieser Konvergenz bereits im Endlichen völlig analog wie zu a) mit Hilfe einer Irrtumsschranke angeben läßt.

Damit haben wir also das folgende Ergebnis:
Zwischen statistischem Datum D_B und der Konzeption K der Gleichartigkeit besteht eine gewisse praktische Gleichwertigkeit, die darin besteht, daß im Verlauf der Beobachtungen gemäß dem Konfidenz-Test eine immer stärkere im Endlichen charakterisierbare Konvergenz gegen die wahre Chance, bzw. die tatsächliche statistische Wahrscheinlichkeit stattfindet.

H) Die Beziehung zwischen D_B und K^*

a) Zunächst können wir feststellen, daß die Konzeption K^* dem statistischen Datum D_B formal insofern entspricht, als K^* aus den Mischungen der Binomialverteilungen bezüglich aller Maße ϕ^* mit positiver stetiger Dichte entspricht.

b) Gemäß (3.8) konvergiert die Konzeption K^* mit quantitativer logischer Wahrscheinlichkeit Eins im Verlauf der Beobachtungen gegen die Konzeption $K_{P(E)}$. Damit besteht nach dem vorangegangenen Abschnitt eine weitere Entsprechung zwischen D_B und K^*.
Allerdings erfolgt die Konvergenz von K^* gemäß (3.8) in einem anderen Sinne als die Konvergenz von D_B gemäß dem vorangehenden Abschnitt. Insbesondere läßt die Konvergenz von K^* gemäß (3.8) keine Aussage über die Geschwindigkeit bzw. das "Ausmaß" dieser Konvergenz zu.
Dem wollen wir in dem nächten Abschnitt auf den Grund gehen.

I) Der subjektivistische Konfidenz-Test

a) Dem objektivistischen Konfidenz-Test läßt sich im Rahmen unserer Begriffsbildungen die folgende Überlegung entgegen stellen:
Wir setzen die Konzeption K^* voraus und betrachten einen beliebige Maßfunktion m_{K^*} mit dem Maß ϕ^*. Auf Grund der Beobachtung der relativen Häufigkeit $h_n = r_n/n$ gehen wir zu dem Maß ϕ^*_{n,r_n} über gemäß der Formel, - siehe dazu (3.8) :

$$d\phi^*_{n,r_n} = p_X(\{\omega: h_n(E,\omega) = h_n\})\, d\phi^* / \int p_X(\{\omega: h_n(E,\omega) = h_n\})\, d\phi^* .$$

Der Konfidenzschluß erscheint vom Standpunkt der Bewertung ϕ^* berechtigt, wenn gilt:

$$\phi^*_{n,r_n}([h_n-\varepsilon, h_n+\varepsilon]) = m_{K^*_n}(P(E) \in [h_n-\varepsilon, h_n+\varepsilon]) \geq 1 - \alpha_s ;$$

K^*_n ist dabei definiert wie in (3.8) als die um die Information h_n vermehrte Konzeption K^* ; α_s muß eine kleine Größe sein, - die subjektivistische Irrtumswahrscheinlichkeit für den Konfidenz-Test. *)

*) Warnung: $m_K(P(E) \in [h_n-\varepsilon, h_n+\varepsilon]) = m_K(\{\omega: |h_n(E,\omega) - p(E,\omega)| \leq \varepsilon\})$ ist im allgemeinen nicht gleich $\phi([h_n-\varepsilon, h_n+\varepsilon])$; der letztere Wert ist nämlich abhängig von der Beschaffenheit der Wirklichkeit, während der erstere Wert rein logisch durch m_K festgelegt ist.

Aus den Sätzen von (3.8) folgt, daß das Maß ϕ^*_{n,r_n} bei hinreichend großem n sein ganzes Gewicht beliebig gut um den Wert $h_n = r_n/n$ konzentriert.

Bei vorgegebenem α_s, ε und vorgegebenem speziellen ϕ^* läßt sich also stets erreichen durch Wahl eines hinreichend großen n, daß gilt: $\phi^*_{n,r_n}([h_n-\varepsilon, h_n+\varepsilon]) \geq 1 - \alpha_s$.

Allerdings läßt sich nicht erreichen, daß dies gilt für alle Maße ϕ^* der Konzeption K^*, so daß $1 - \alpha_s$ also nicht die Bedeutung einer quantitativen logischen Wahrscheinlichkeitsschranke zur Konzeption K^* besitzt. Auch wenn die Funktion $f_{n,r_n}(x) := p_x(\{\omega: h_n(E,\omega) = h_n\})$ eine beliebig gute Konzentration um den Wert $h_n = r_n/n$ aufweist, kann doch das ϕ^* für $[h_n-\varepsilon, h_n+\varepsilon]$ einen beliebig kleinen Wert haben, so daß die obige Konfidenzbedingung für diese ϕ^* nicht besteht.

Wollen wir den Konfidenz-Test vom subjektivistischen Standpunkt rechtfertigen, so müssen wir die Klasse der Maßfunktionen weiter einschränken. Dabei genügt es, nur solche ϕ^* zuzulassen, deren Dichte im Intervall $[h_n-\varepsilon, h_n+\varepsilon]$ über einer Schranke γ liegt. Wir bezeichnen diese ϕ^* mit ϕ^*_γ un die Konzeption dieser ϕ^*_γ mit K^*_γ.

Bei dieser Einschränkung läßt sich dann durch Wahl eines hinreichend großen n für alle ϕ^*_γ der Konzeption K^*_γ erreichen:

$\phi^*_{\gamma,n,r_n}([h_n-\varepsilon, h_n+\varepsilon]) = m_{K^*_{\gamma,n}}(P(E) \in [h_n-\varepsilon, h_n+\varepsilon]) \geq 1 - \alpha_s$;

$K^*_{\gamma,n}$ ist dabei wieder die um die Information h_n vermehrte Konzeption K^*_γ ; $1 - \alpha_s$ hat die Stellung einer quantitativen logischen Wahrscheinlichkeitsschranke für die Konzeption $K^*_{\gamma,n}$.

Damit haben wir das folgende Ergebnis: Die Auswahl einer objektivistischen Irrtumswahrscheinlichkeit α für den Konfidenz-Test, die durch gewisse formal nicht voll erfaßbare Plausibilitätsüberlegungen gegeben ist, entspricht der subjektivistische Konfidenz-Test, wobei zusätzlich zur Wahl der subjektivistischen Irrtumswahrscheinlichkeit α_s noch eine Schranke

γ bestimmt werden muß. Genauso wie man bei der Wahl der Irrtumswahrscheinlichkeit auf Plausibilitätsüberlegungen angewiesen ist, muß man bei der Wahl der Schranke γ Plausibilitätsüberlegungen anstellen.
Es gilt dabei: Wird der Konfidez-Test angenommen auf Grund einer objektivistischen Irrtumswahrscheinlichkeit α, so gibt es eine geeignete subjektivistische Irrtumswahrscheinlichkeit α_s und eine Schranke γ, so daß der subjektivistische Konfidenz-Test auch angenommen wird; und umgekehrt.
Objektivistische und subjektivistische Überlegungen sind also äquivalent, wobei dem α vom objektivistischen Standpunkt die Größen α_s und γ vom subjektivistischen Standpunkt entsprechen.

b) Aus der Relation $m_K(|h_n(E) - P(E)| \geq \varepsilon) \leq 1/4n\varepsilon^2$, die die Situation beim objektivistischen Konfidenz-Test beschreibt, läßt sich durch Gleichsetzung von objektivistischer und subjektivistischer Irrtumswahrscheinlichkeit die Forderung motivieren, daß nur solche subjektivistischen Maße ϕ^O sinnvoll sind, für die gilt:

(*) $\phi^O_{n,r_n}([\frac{r_n}{n} - \varepsilon, \frac{r_n}{n} + \varepsilon]) \geq 1 - 1/4n\varepsilon^2$ und zwar für alle ε, n und alle möglichen Werte von r_n.

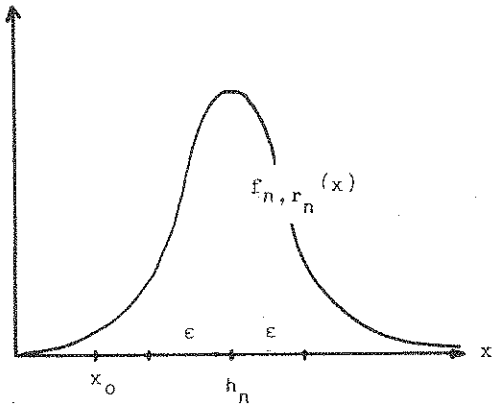
Die Gesamtheit all solcher Maße ϕ^O fassen wir zu der Konzeption K^O zusammen.
Die Frage nach der Beschaffenheit von K^O führt über den Rahmen dieser Arbeit hinaus.
Es sei jedoch darauf hingewiesen, daß der Zusammenhang zwischen K^O und den CARNAP'schen Bestätigungsfunktionen c_λ interessant ist. Möglicherweise führt nämlich die obige Bedingung (*) ausgesprochen für die CARNAP'schen ψ_λ (siehe (9.2b)) zu einer Einschränkung des zulässigen Parameterbereichs für das CARNAP'sche λ. Es sei allerdings bemerkt, daß sich diese Eingrenzung für den CARNAP'schen Parameter λ kaum analytisch wird ermitteln lassen, sondern mit Hilfe eines Computers für spezielle Werte von ε, n, r_n durchgetestet werden müßte.
Auch diese Fragestellung führt über den Rahmen dieser Arbeit hinaus.

J) Der Konfidenz-Test als Likelihood-Test

Schließlich wollen wir noch zeigen, wie sich der Konfidenz-Test als Likelihood-Test im Sinne von HACKING darstellen läßt.
Die A-priori-Hypothese ist das statistische Datum D_B.
Die zu verwerfende Hypothese lautet $H' := [0, h_n - \varepsilon] \cup [h_n + \varepsilon, 1]$.
Die Alternativhypothese ist $H := [h_n - \varepsilon , h_n + \varepsilon]$.

Wir wissen, daß die Likelihoodfunktion
$f_{n,r_n}(x) = p_x(\{\omega : h_n(E,\omega) = h_n\})$
ihr Gewicht bei hinreichend großem n beliebig gut um den Wert $h_n = r_n/n$ konzentriert.

Bei vorgegebenem α und hinreichend großem n existiert also zu jedem x_o aus H' ein $\overline{x}$ aus H, nämlich $\overline{x} = h_n$, so daß gilt: $f_{n,r_n}(x_o)/f_{n,r_n}(h_n) \leq \alpha$.



Zu dieser Überlegung ist allerdings zu bemerken, daß die Formulierung der Hypothesen H, H' von dem Beobachtungsergebnis h_n abhängt, auf Grund dessen zwischen den Hypothesen entschieden werden soll. Ob das der Intention von HACKING entspricht, ist zweifelhaft. Aber anders läßt sich der Konfidenz-Test nicht als Likelihood-Test darstellen.

K) Die Überprüfung von D_B und K

a) HACKING betont, daß das statistische Datum nie etwas Letztes, Unanfechtbares darstellt, sondern ein wissenschaftlicher Ansatz ist, der der Überprüfung fähig ist.
Dies steht im Einklang mit unserer Konzeption K, die ebenfalls aufgegeben werden kann.
Das statistische Datum D_B wird verworfen, wenn auf Grund der Beobachtungen ein anderes Datum sichtbar wird, das "besser" mit den Beobachtungen im Einklang steht als D_B .
Der Likelihood-Test gibt an, wie diese Verwerfung eines statistischen Datums formal durchgeführt wird.
Dasselbe gilt für die Konzeption K ; als Beispiel führten wir in (5.2) die deterministische Konzeption K_{neu} an.

b) Da wir hier keine Theorie von Wahrscheinlichkeitskonzeptionen, unverträglich mit der Konzeption K der Gleichartigkeit betreiben wollen, geben wir nur die folgende Regel an, die eine Übertragung des für Chancen konzipierten Likelihood-Tests auf Maßfunktionen darstellt:
Die Konzeption K wird auf Grund eines Ereignisses A verworfen zu Gunsten einer Konzeption $\overline{K}$, wenn zu jeder Maßfunktion m_K von K eine Maßfunktion $m_{\overline{K}}$ von $\overline{K}$ existiert, so daß gilt: $m_K(A)/m_{\overline{K}}(A) \leq \alpha$; α muß dabei eine vorgegebene "plausible" Größe sein.

c) Um die in §5 durchgeführten Überlegungen zu erläutern, betrachten wir den Fall, daß das obige Ereignis A von der folgenden Gestalt ist: $A = \{\omega : |h_n(E,\omega) - h_r(E,\omega)| \geq \varepsilon\}$.
A sei nun eingetreten und es gelte die obige Relation: $m_K(A)/m_{\overline{K}}(A) \leq \alpha$. Dann wird die Konzeption K zu Gunsten der Konzeption $\overline{K}$ verworfen.

Es gilt nun weiterhin: $m_K(A) \leq 1/4N\varepsilon^2$ für $n,r \geq N$.

Die Zahl $1/4N\varepsilon^2$ stellt die quantitative logische Wahrscheinlichkeitsschranke dafür dar, daß K verworfen wird, dh. A eintritt, obwohl K richtig ist. Diese Zahl läßt sich also als logische Irrtumswahrscheinlichkeit bei der Verwerfung von K ansehen.

Im Allgemeinen läßt sich jedoch eine solche logische Irrtumswahrscheinlichkeit nicht angeben, da das die Entscheidung zwischen K und $\overline{K}$ erzeugende Ereignis A nicht von der obigen Gestalt ist. Insbesondere ist das Ereignis A nicht von der obigen Gestalt bei der Entscheidung zwischen K und K_{neu} von (5.2); bei K_{neu} konvergieren ebenfalls die relativen Häufigkeiten, so daß das obige Ereignis A in diesem Sinne keine Entscheidung zwischen K und K_{neu} herbeiführen würde.

(10.2) Vergleich mit dem System von KYBURG [Ky]

A) Das formale Sprachsystem L

KYBURG legt seinen Überlegungen ein formales Sprachsystem L zu Grunde, das wir hier nicht näher angeben wollen. Es soll nur bemerkt werden, daß dieses Sprachsystem so reich ist, daß sich der gesamte mathematische Apparat der Wahrscheinlichkeitsrechnung darin formulieren läßt; außerdem lassen sich darin noch gewisse empirische Aussagen ausdrücken, von denen wir einige unter C) angeben werden.

B) Der Rationale Korpus

Als Rationalen Korpus bezeichnet KYBURG eine Menge von Sätzen, die als gültig, - etwa von einem Wissenschaftler - , akzeptiert werden. Der Rationale Korpus soll dabei nicht nur Sätze umfassen, die "definitiv" verifiziert sind, sondern auch Sätze, die von hoher Wahrscheinlichkeit sind. Die Tatsache, daß der Rationale Korpus auch Sätze enthält, die nur sehr wahrscheinlich sind, macht es schwierig, die logische Konsistenz des Rationalen Korpus zu definieren. Wir geben dazu das folgende Beispiel:
Wir betrachten die Lose zu einer Lotterie. Es ist sehr unwahrscheinlich, daß das Los Nummero k gewinnt, dh. wir können den Satz " das Los Nummero k gewinnt nicht " in unseren Rationalen Korpus aufnehmen und zwar für jede Nummer k. Die Konjunktion all dieser Sätze liegt jedoch nicht in dem Rationalen

Korpus, da ein Los mit Sicherheit gewinnt. Um konsistent zu sein, darf der Rationale Korpus also nicht alle Folgerungen enthalten, die man rein logisch aus ihm ableiten kann.
Die Komplikationen der KYBURG'schen Theorie, die sich aus dieser Eigenschaft des Rationalen Korpus ergeben, wollen wir hier nicht referieren; wir beschränken uns in dieser Darstellung auf den Hinweis, daß der Begriff der Konsistenz des Rationalen Korpus mit Schwierigkeiten verbunden ist und zu gesonderten Überlegungen Anlaß gibt, über die wir aber hier hinweggehen.

C) Die statistische Aussage

a) Ein fundamentaler Begriff in KYBURGs System ist der Begriff der statistischen Aussage, abgekürzt: $S(x,y,b,p,q)$; dabei soll x eine unendliche Folge von Mengen x_n sein; die Menge x_n ist dabei die Menge der Beobachtungsergebnisse nach n "Versuchen" . y ist eine Zufallsgröße, definiert auf der Menge $\bigcup_n x_n$. b ist eine Borel-menge. p , q sind zwei Zahlen zwischen 0 und 1 . $S(x,y,b,p,q)$ soll heißen:
Der Limes der relativen Häufigkeiten der Elemente aus x_n mit y-Werten in b für n gegen Unendlich liegt zwischen p und q ; in Formeln:

$$\lim_{n\to\infty} \frac{\text{Mächtigkeit der Menge } \{\, a:\ a \in x_n \ \wedge\ y(a) \in b \,\}}{\text{Mächtigkeit der Menge } x_n} \in [p,q] \ .$$

b) Die statistische Aussage $s := S(x,y,b,p,q)$ heißt <u>stärker</u> als die statistische Aussage $s' := S(x',y',b',p',q')$ bezüglich des Rationalen Korpus w in der Sprache L , in Formeln $sStr_{w,L}s'$, wenn gilt: $[p,q] \subset [p',q']$.

c) Die statistischen Aussagen s , s' <u>differieren</u>, in Formeln $sDif_{w,L}s'$, wenn gilt: $[p,q] \not\subseteq [p',q']$ und $[p',q'] \not\subseteq [p,q]$.

d) Die Aussage $PS_{w,L}(s,x,y,b)$ besagt, daß die statistische Aussage s die stärkste statistische Aussage des Rationalen

Korpus w in der Sprache L ist bezüglich der Terme x ,y , b ; genauer: Jede Aussage S(x,y,b,p',q') , die in w enthalten ist, ist nicht stärker als s und s ist nicht stärker als eine außerhalb von w liegende statistische Aussage S(x,y,b,p',q') .

D) Randomness

Ein wichtiger Bestandteil in KYBURGs Definition der Wahrscheinlichkeit ist der Begriff der Randomness. Dieser Begriff bezieht sich dabei auf die folgende Problematik:
Angenommen wir wissen, daß gilt $a \in \bigcup_n x_n$; außerdem gelte $s := S(x,y,b,p,q)$. Dann besteht für die Aussage $y(a) \in b$ auf Grund der Relation s eine bestimmte statistische Information. Jetzt kann es aber sein, daß der Rationale Korpus w weiterhin die Aussage $s' := S(x',y,b,p',q')$ enthält und außerdem gilt $a \in \bigcup_n x'_n$; dann ist es möglich, daß auf Grund von s' eine andere statistische Information für die Aussage $y(a) \in b$ besteht; vergleiche dazu (3.7d) . Diese Situation wird ausgeschlossen, wenn die Aussage $RAN_L(a,x,y,b,w)$ gilt, die wie folgt definiert ist:

$RAN_L(a,x,y,b,w)$: $\Leftrightarrow$ w ist ein konsistenter Rationaler Korpus
$\wedge$ " $a \in \bigcup_n x_n$ " $\in$ w
$\wedge$ $\exists$ s($PS_{w,L}(s,x,y,b)$
$\wedge$ $\neg \exists$ x'($x'D_L(a,x,y,b,w) \vee x'S_L(a,x,y,b,w)$) .

Die Relation $x'D_L(a,x,y,b,w) \vee x'S_L(a,x,y,b,w)$ besagt dabei im Wesentlichen:
Wenn auf Grund von w die Menge $\bigcup_n x'_n$ nicht Obermenge von $\bigcup_n x_n$ ist und weiterhin "$a \in \bigcup_n x'_n$" $\in$ w , dann gibt es ein $s' := S(x',y,b,p',q')$ in w , so daß:
$PS_{w,L}(s',x',y,b) \wedge (s'Str_{w,L}s \vee s'Dif_{w,L}s)$.

Wir bemerken zu dieser Erläuterung der Relation RAN , daß wir hier nur die wesentlichsten Grundgedanken wiedergegeben haben; auf die genaue Definition können wir hier nicht eingehen und verweisen dazu auf KYBURG .

E) Wahrscheinlichkeit

Die Wahrscheinlichkeit einer Aussage $y(a) \in b$ im Hinblick auf einen Rationalen Korpus w ist bei KYBURG das Intervall [p,q] der bezüglich dieser Aussage schärfsten statistischen Aussage, vorausgesetzt, daß die Relation RAN_L erfüllt ist; genauer:

$prob_L(y(a) \in b , w) = [p,q]$ dann und nur dann:
Es gibt eine Beobachtungsfolge
x : $RAN_L(a,x,y,b,w)$
und
$PS_{w,L}(S(x,y,b,p,q),x,y,b)$.

Dabei wird zunächst vorausgesetzt, daß der Rationale Korpus w "strikt konsistent" ist und dann in einem weiteren Schritt die Definition von $prob_L$ auf inkonsistente Rationale Korpus' erweitert. Darauf wollen wir jedoch hier nicht eingehen und verweisen auf KYBURG.

F) Kritik des Systems von KYBURG

a) In der für das System von KYBURG grundlegenden Definition der statistischen Aussage wird die Konvergenz der relativen Häufigkeiten vorausgesetzt. Diese Voraussetzung wird im System von KYBURG nicht fundiert. KYBURG muß eine Theorie voraussetzen, die diese Bedingung liefert; dafür käme entweder die Theorie von MISES in Frage oder eine Theorie, die die KOLMOGOROFF-Axiome liefert und diese Voraussetzung mit Wahrscheinlichkeit Eins sichert oder etwa unsere Theorie.
Das System von KYBURG liefert keinen Beitrag zu der in unserer Arbeit untersuchten Grundlagenproblematik.

b) Der Umstand, daß KYBURG nicht den Limes der relativen Häufigkeiten Wahrscheinlichkeit nennt, sondern das bezüglich dieses Limes im Rationalen Korpus w akzeptierte Konfidenzintervall, hat soviel wie gar keinen Erkenntniswert, sondern stellt nur eine neue Namensgebung dar.
Für die Zweckmäßigkeit dieser Namensgebung läßt sich anführen,

daß man den Limes der relativen Häufigkeiten niemals kennt und in der Praxis immer nur Konfidenzintervalle bezüglich dieses Limes vorkommen.
Gegen diese Namensgebung läßt sich anführen, daß es adäquater erscheint, den Limes der relativen Häufigkeiten als den objektiven Grenzwert aller Konfidenzintervalle Wahrscheinlichkeit zu nennen, anstatt den Begriff Wahrscheinlichkeit auf die vom jeweiligen Wissenstande abhängigen Konfidenzintervalle zu relativieren.

c) Die von KYBURG eingeführte Wahrscheinlichkeitsrechnung mit Intervallen ist mathematisch komplizierter als die Rechnung mit einer Zahl; außerdem beruht die Wahrscheinlichkeitsrechnung mit Intervallen begrifflich auf der mit einer Zahl; sie ist damit nicht nur mathematisch umständlich, sondern auch überflüssig.

d) Zwischen dem Begriff des Rationalen Korpus und der Konzeption K besteht eine gewisse Analogie. Die Konzeption K stellt allerdings einen wissenschaftlichen Ansatz dar, der unter gewissen Bedingungen gegeben ist, jedoch der Überprüfung durch Beobachtungen fähig ist. Dagegen ist der Rationale Korpus die Klasse aller akzeptierten Sätze; darin geht bereits der Wahrscheinlichkeitsbegriff ein, während in unserer Theorie der Wahrscheinlichkeitsbegriff erst aus der Konzeption K entwickelt wird.

e) Abschließend sei noch darauf hingewiesen, daß es mir trotz intensiver Bemühung nicht gelang, zu einer vollen Klarheit zu kommen, was die von KYBURG eingeführte, in D) skizzierte Definition der Relation RAN und deren Funktion im weiteren Aufbau des KYBURG'schen Systems betrifft. Ich fühle mich außerstande den wissenschaftlichen Wert zu beurteilen, den die Einführung der Relation RAN für das unter D) skizzierte, an sich wissenschaftlich interessante Problem besitzt.

(10.3) Vergleich mit dem System von SCHNORR [SCH]

A) Ziel des Systems von SCHNORR

Das System von SCHNORR beruht auf dem Ansatz von MISES.
SCHNORR hat sich dabei die Aufgabe gestellt, das Regellosigkeitprinzip von MISES zu präzisieren. Bekanntlich ist die Formulierung des Regellosigkeitsprinzips von MISES widersprüchlich.
Die Verbesserungen des Regellosigkeitsprinzips von WALD , CHURCH und anderen lassen diese Widersprüchlichkeit zwar veschwinden, führen aber zu dem inadäquaten Ergebnis, daß die so definierten Zufallsfolgen das Gesetz vom iterierten Logarithmus nicht erfüllen. Das ist der Ausgangspunkt der Überlegungen von SCHNORR.
Wir wollen das genauer ausführen:

B) Die bisherige Theorie

a) Wir beschränken uns der Einfachheit halber auf ein BERNOULLI-Experiment mit den beiden möglichen Ergebnissen 0 , 1 ;
$X := \{0 , 1\}$ ist die Ergebnismenge zu diesem Experiment.
Wir betrachten unendliche Folgen (Wiederholungen) des BERNOULLI-Experiments; die Ergebnismenge dazu bezeichnen wir mit X^∞ ;
$z \in X^\infty : \Leftrightarrow z = z_1 z_2 z_3 \ldots$ mit $z_i \in X$. X^* bezeichne die Menge aller endlichen Folgen von Elementen aus X .
$\Lambda \in X$ bezeichne die leere Folge.
Zu $z \in X^\infty$ sei $z(n)$ die Anfangsfolge der Länge n .
Die Aneinanderreihung einer Folge $x \in X^*$ und $y \in X^* \cup X^\infty$ wird als Produkt xy geschrieben. Dies impliziert in natürlicher Weise ein Produkt $AB \subseteq X^\infty$ von Mengen $A \subseteq X^*$ und $B \subseteq X^\infty$.

b) Unter einer Auswahlregel verstehen wir eine Abbildung
$\phi : X^* \to \{0 , 1\}$.

Einer Auswahlregel ϕ ordnen wir in eineindeutiger Weise eine Abbildung $\Phi : X^* \to X^*$ wie folgt zu:

$$\Phi(\Lambda) = \Lambda$$

$$\Phi(x\alpha) = \begin{cases} \Phi(x)\alpha & \text{sofern } \phi(x) = 1 \\ \Phi(x) & \text{sonst} \end{cases} \qquad \text{für alle } x \in X^* , \alpha \in X .$$

Weiter ordnen wir einer Auswahlregel ϕ eine partielle Funktion $\overline{\Phi} : X^\infty \to X^\infty$ wie folgt zu:
Der Definitionsbereich $D(\overline{\Phi})$ sei gegeben durch
$D(\overline{\Phi}) := \{z \in X^\infty:$ Die Folge $\Phi(z(n))$ bricht nicht ab für $n \to \infty\}$.

Für $z \in D(\overline{\Phi})$ definieren wir dann $\overline{\Phi}(z)$ wie folgt:
$\overline{\Phi}(z) \in \Phi(z(n))X^\infty$ für alle $n \in N$.

c) Ein Kollektiv über X ist eine Folge $z \in X^\infty$, für die es ein normiertes Maß μ auf der Potenzmenge von X gibt, so daß für alle Auswahlregeln ϕ mit $z \in D(\overline{\Phi})$ gilt:

$$\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} \overline{\Phi}(z)_i = \mu(1) .$$

$\mu(1)$ heißt dann die Wahrscheinlichkeit von 1 bezüglich des Kollektivs z .

Es gilt nun der bekannte Satz:
$0 < \mu(1) < 1 \Rightarrow$ Es gibt kein Kollektiv.

d) Zur Existenz von Kollektiven gelangen wir, wenn wir die Klasse der möglichen Auswahlregeln einschränken
Sei A eine Menge von Auswahlregeln ϕ und μ ein normiertes Maß auf der Potenzmenge von X . Dann heißt $z \in X^\infty$ ein Kollektiv zu A und μ , wenn für alle $\phi \in A$ mit $z \in D(\overline{\Phi})$ gilt:
$\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} \overline{\Phi}(z)_i = \mu(1)$. $\mathfrak{K}(A,\mu)$ sei die Menge aller Kollektive zu A und μ .

Es gilt nun der folgende von WALD [W] stammende Satz:
Für jedes abzählbare System A von Auswahlregeln gilt:
$\mathfrak{K}(A,\mu)$ hat die Mächtigkeit des Kontinuums; genauer gilt:
$\overline{\mu}(\mathfrak{K}(A,\mu)) = 1$; dabei ist $\overline{\mu}$ das unendliche Produktmaß zu μ über X^∞ .

e) Das Gesetz vom iterierten Logarithmus lautet für unseren Fall (BERNOULLI-Experiment) :

$$\overline{\mu}\{z \in X^\infty: (\overline{\lim_{n\to\infty}}) \frac{\sum_{i=1}^{n} z_i - n\mu(1)}{D\sqrt{2n\log\log n}} = (\pm) 1\} = 1 ;$$

dabei ist D die Streung der Zufallsvariablen, in unserem Fall $D = \sqrt{\mu(1) - \mu^2(1)}$.

Das Gesetz vom iterierten Logarithmus läßt erwarten, daß für jedes Kollektiv $z \in X^\infty$ die folgende Relation erfüllt ist:

$$(\overline{\lim_{n\to\infty}}) \frac{\sum_{i=1}^{n} z_i - n\mu(1)}{D\sqrt{2n\log\log n}} = (\pm) 1 .$$

Aus einem Satz von VILLE [V] folgt jedoch, daß es bei jeder abzählbaren Klasse A von Auswahlregeln ein Kollektiv $z \in \mathfrak{K}(A,\mu)$ gibt, so daß gilt:

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} z_i - n\mu(1)}{D\sqrt{2n\log\log n}} = 0 .$$

Aus diesem Grund entwickelt SCHNORR eine schärfere Theorie der Zufallsfolgen, so daß die obige Relation erfüllt ist.

C) Hyperzufällige Folgen

a) Wir setzen in diesem Abschnitt voraus, daß $\mu | P(X)$ die Gleichverteilung ist, dh. $\mu(1) = 1/2$. $\overline{\mu}$ sei dabei wieder das Produktmaß über X^∞ . Sei $A \subseteq X^*$; dann bezeichnet $[A] \subseteq X^\infty$ die Zylindermenge in X^∞ zur Basis A .

b) $\mathcal{N} \subseteq X^\infty$ ist eine rekursive Nullmenge, wenn es eine rekursiv aufzählbare Menge $Y \subseteq N \times X^*$ gibt mit

(1) $\overline{\mu}[Y_i] \leq 2^{-i}$ für alle $i \in N$,

(2) $\mathcal{N} \subseteq \bigcap_{i \in N} [Y_i]$, $Y_i := \{x \in X^* : (i,x) \in Y\}$.

Y heißt ein rekursiver Sequentialtest zu $\mathcal{N}$.

c) Satz: Es gibt einen rekursiven Sequentialtest $U \subseteq N \times X^*$, so daß es zu jedem rekursiven Sequentialtest $Y \subseteq N \times X^*$ ein $k \in N$ gibt mit $[Y_{i+k}] \subseteq [U_i]$ für alle $i \in N$.
Ein solcher rekursiver Sequentialtest heißt universell.
Ist U ein universeller rekursiver Sequentialtest, dann enthält $\mathcal{N}_U := \bigcap_{i \in N} [U_i]$ jede rekursive Nullmenge.

$\mathcal{N}_U$ heißt die universelle rekursive Nullmenge.

d) Definition: Die Folgen in $\mathcal{K}_h := X^\infty - \mathcal{N}_U$ heißen hyperzufällig.
Es gilt: $\overline{\mu}(\mathcal{K}_h) = 1$.

D) Zufällige Folgen

a) Eine Funktion $f: N \to R$ heißt berechenbar, wenn es eine rekursive Funktion $g: N \times N \to Q$ gibt mit $|f(n) - g(n,i)| \leq 2^{-i}$, für alle $i , n \in N$.

Diese Definition gilt wie üblich in der entsprechenden Weise für Funktionen $f: Y \to R$, wobei Y eine ohne Wiederholungen effektiv aufzählbare Menge ist.
Eine reelle Zahl r heißt berechenbar, wenn es eine rekursive Funktion $h: N \to Q$ gibt mit $|h(i) - r| \leq 2^{-i}$ für alle $i \in N$.

b) Definition: Ein rekursiver Sequentialtest $Y \subseteq N \times X^*$ heißt total rekursiv, wenn durch $f(n) := \overline{\mu}[Y_n]$ eine berechenbare Funktion $f: N \to R$ definiert wird. Eine Menge $\mathcal{N} \subseteq \mathcal{N}_Y := \bigcap_{i \in N} [Y_i]$ heißt dann eine total rekursive Nullmenge.

c) Definition: Eine Folge $z \in X^\infty$ heißt zufällig, wenn sie in keiner total rekursiven Nullmenge enthalten ist.
$\mathcal{K} \subseteq X^\infty$ sei die Menge aller zufälligen Folgen.
Es gilt: $\mathcal{K}_h \subseteq \mathcal{K}$ und somit $\overline{\mu}(\mathcal{K}) = 1$.

E) Das Prinzip vom ausgeschlossenen Spielsystem

a) Ein Spiel soll gegeben sein durch zwei Funktionen $B_i: X^* \to R^+$, $i = 0 , 1$; R^+ sei die Menge der nicht negativen reellen Zahlen. $B_i(z_1 \ldots z_n)$ bedeute den Einsatz nach Kenntnis der Glieder $z_1 \ldots z_n$ auf das Ereignis $z_{n+1} = i$.

Die Einsatzfunktionen B_i erzeugen aus einem Anfangsvermögen $V(\Lambda)$ eine Vermögensfunktion $V: X^* \to R$. Dabei bezeichnet $V(z_1 \ldots z_n)$ das Vermögen nach dem Ergebnis $z_1 \ldots z_n$. Das Vermögen ändert sich von der n-ten zur (n+1)-ten Spielrunde wie folgt:

(A) $V(z_1 \ldots z_{n+1}) = V(z_1 \ldots z_n) + \sum_{i=0,1} (2\delta^i_{z_{n+1}} - 1) B_i(z_1 \ldots z_n)$;

δ^i_k ist das Kroneckersymbol.

Die Auszahlungsbedingung (A) bedeutet, daß der Spieler den doppelten Einsatz erhält, den er auf das eintretende Ereignis gesetzt hat, während der auf das nicht eintretende Ereignis gesetzte Betrag verloren geht.
Hieraus folgt für das Vermögen V die Funktionalgleichung:

(M) $V(x) = 2^{-1} V(x0) + 2^{-1} V(x1)$, $x \in X^*$.

Funktionen $V: X^* \to R$ mit der Eigenschaft (M) heißen Martingale.

Satz: Jedes Martingal wird von geeigneten Einsatzfunktionen aus einem Anfangsvermögen $V(\Lambda)$ erzeugt.

Definition: Eine Funktion $F: X^* \to R$ heißt subberechenbar, wenn es eine rekursive Funktion $g: N \times X^* \to Q$ gibt mit
(1) $g(i,x) \leq g(i+1,x)$ für alle $i \in N$, $x \in X^*$;
(2) $\lim_{i \to \infty} g(i,x) = F(x)$.

Es gilt jetzt der folgende

Satz: Eine Folge $z \in X^\infty$ ist genau dann hyperzufällig, wenn es kein subberechenbares Martingal gibt mit $\overline{\lim_{n \to \infty}} V(z(n)) = \infty$.

b) Unter einer Ordnungsfunktion verstehen wir eine rekursive, monotone und unbeschränkte Funktion $h: N \to N$.
Sei $f: N \to R$ eine Funktion, dann sagen wir, die Folge $(f(n) : n \in N)$ wachse mit der Ordnung h , wenn $\overline{\lim_{n\to\infty}} (f(n) - h(n)) > 0$.
Gibt es zu f eine solche Ordnungsfunktion h , so schreiben wir abkürzend: $k\overline{\lim_{n\to\infty}} f(n) = \infty$.

Satz: Eine Folge $z \in X^\infty$ ist genau dann zufällig, wenn es kein berechenbares Martingal V gibt mit $k\overline{\lim_{n\to\infty}} V(z(n)) = \infty$.

F) Allgemeines Maß μ

a) Die obigen Überlegungen lassen sich in natürlicher Weise auf beliebige berechenbare Maße $\mu | \mathcal{P}(X)$ übertragen.
Etwas schwieriger ist die Definition der Zufallsfolgen bei nicht berechenbarem μ . Wir wollen die formale Struktur der SCHNORR'schen Überlegungen zu nicht berechenbarem μ darstellen; auf die genaue Definition wollen wir jedoch hier nicht eingehen.

b) SCHNORR definiert zunächst einen verteilungsunabhängigen (kurz: v.u.) rekursiven Sequentialtest $Y \subseteq N \times X^*$, der ohne Bezug auf ein bestimmtes Maß μ eingeführt wird.
Eine v.u. rekursive Nullmenge wird dann analog wie oben definiert. Für jedes Produktmaß gilt nun: $\overline{\mu}(\mathcal{N}_Y) = 0$, wenn $\mathcal{N}_Y$ eine v.u. rekursive Nullmenge darstellt.
Es existiert nun wieder eine universelle v.u. rekursive Nullmenge.

c) Weiterhin definiert SCHNORR einen v.u. total rekursiven Sequentialtest.
Sei $\mathcal{K}_\mu$ die Menge aller Zufallsfolgen zu dem berechenbaren Maß μ . Dann gilt: $z \in \mathcal{K}_\mu$ dann und nur dann, wenn

(1) z liegt in keiner v.u. total rekursiven Nullmenge;

(2) $\lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} z_i = \mu(1)$.

d) Als Menge $\mathcal{K}$ der binären Zufallsfolgen definiert SCHNORR die Menge derjenigen Folgen, die in keiner v.u. total rekursiven Nullmenge liegen. Zu beliebigem Maß μ definiert SCHNORR wie folgt die Menge $\mathcal{K}_\mu$:

$$\mathcal{K}_\mu := \{z \in \mathcal{K} : \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} z_i = \mu(1)\} .$$

Es gilt: $1 = \overline{\mu}(\mathcal{K}) = \overline{\mu}(\mathcal{K}_\mu)$.

Damit ist die Definition der Zufallsfolgen von SCHNORR abgeschlossen. Laut SCHNORR erfüllen die Zufallsfolgen $\mathcal{K}$ das Gesetz vom iterierten Logarithmus.

G) Kritik des Systems von SCHNORR

a) Wie schon zu Anfang bemerkt, stellt das System von SCHNORR die Ausarbeitung des Systems von MISES dar.
Die von uns an der Theorie von MISES geübte Kritik überträgt sich wörtlich suf das System von SCHNORR: siehe dazu auch den folgenden Abschnitt H) .

b) Man könnte zunächst meinen, daß SCHNORR in seiner Definition der zufälligen und hyperzufälligen Folgen den Maßbegriff voraussetzt. Wie jedoch die Charakterisierung der SCHNORR'schen Folgen durch das Prinzip vom ausgeschlossenen Spielsystem zeigt, gibt es eine Definition der Zufallsfolgen, die vom Maßbegriff unabhängig ist; der Maßbegriff spielt nur eine heuristische Rolle.
Die von SCHNORR durchgeführte Begriffsbildung der zufälligen und hyperzufälligen Folgen kann im Hinblick auf sein Prinzip vom ausgeschlossenen Spielsystem als eine wertvolle Erkenntnis gelten.

c) Wir wollen hier noch bemerken, daß der Vorschlag von CHURCH (siehe [CH]) , als Auswahlregeln die rekursiven Funktionen zu wählen, nicht ausreicht. Man muß weiterhin fordern, daß die durch Experimente definierten Auswahlregeln den Limes der relativen Häufigkeiten invariant lassen: Beispiel: Wir machen die Auswahl des n-ten Glieds einer Folge von Experimenten davon abhängig,

ob eine von uns geworfene Münze auf den Kopf fällt oder nicht. Da die in dieser Weise von der Menschheit festlegbaren Auswahlfunktionen ebenfalls als abzählbar anzusehen sind, läßt sich gemäß dem WALD'schen Satz der Ansatz von MISES retten. Allerdings müßte SCHNORR in seiner Definition der Zufallsfolgen diese Problematik auch noch berücksichtigen, was jedoch nicht der Fall ist.

H) SCHNORRs Kritik an dem System von KOLMOGOROFF

a) SCHNORR übt am System von KOLMOGOROFF die folgende Kritik: " Das Axiomensystem von KOLMOGOROFF hat sich nicht zuletzt wegen seiner syntaktischen Einfachheit gegenüber anderen Vorstellungen durchgesetzt. Inhaltlich jedoch sind diese Axiome wohl weniger einfach. Die Wahrscheinlichkeit $p(A)$ bedeutet soviel wie den Grad der Gewißheit des Ereignisses A. Dieser Grad der Gewißheit eines Ereignisses wird als eine jedem Ereignis zukommende Grundeigenschaft aufgefaßt, ähnlich dem Gewicht, der Ausdehnung und der Farbe von Körpern. Man wird sicherlich zugeben, daß der Grad der Gewißheit, sofern er sich nicht durch einfachere Begriffe erläutern läßt, nicht sehr anschaulich ist. Z.B. ist es unklar, was der Grad der Gewißheit $1/2$ für ein einmaliges Ereignis bedeutet. Die Interpretation des Wahrscheinlichkeitsbegriffs, so argumentiert man, liegt eben außerhalb der mathematischen Aufgabenstellung. Dennoch legt man implizit meist die Vorstellung zu Grunde, daß Ereignisse mit einem hohen Grad der Gewißheit häufig auftreten, während solche mit einem kleinen Grad der Gewißheit recht selten sind. Diese mehr anschauliche Interpretation der Wahrscheinlichkeit kommt aber in den KOLMOGOROFF'schen Axiomen nicht zum Ausdruck. Sie kann daher auch nicht aus ihnen abgeleitet werden. "

b) Zu dieser Kritik bemerken wir zunächst, daß die KOLMOGOROFF'sche Erläuterung der Wahrscheinlichkeit als Grad der Gewißheit tatsächlich etwas vage ist. KOLMOGOROFF gibt nämlich nur die folgende Konsequenz an: Wenn $p(A)$ nahe bei Eins liegt, ist der Eintritt von A als praktisch sicher zu betrachten. Dies ist in der Tat nicht sehr präzis.

c) Wir müssen jedoch der Behauptung SCHNORRS widersprechen, daß die Häufigkeitsinterpretation in den KOLMOGOROFF'schen Axiomen nicht enthalten ist und folglich nicht daraus abgeleitet werden kann. Tatsächlich werden die KOLMOGOROFF'schen Axiome über die Häufigkeitsvorstellung motiviert.
Außerdem scheint uns das schwache und starke Gesetz der großen Zahlen die Ableitung der Häufigkeitsinterpretation zu gewähren. In diesen Gesetzen läßt sich nämlich die von KOLMOGOROFF vorgenommene, unter b) zitierte Interpretation einer hohen Wahrscheinlichkeit als praktische Sicherheit durchführen und ergibt die von SCHNORR beschriebene Häufigkeitsinterpretation, auch wenn SCHNORR diese Häufigkeitsinterpretation dem KOLMOGOROFF'schen System abspricht.
Die von uns formulierte,gewisse erkenntnistheoretische Zikularität der KOLMOGOROFF-Theorie scheint uns die Mängel dieser Theorie prägnanter zu erfassen als die etwas vagen Ausführungen von SCHNORR.

d) Wir wollen feststellen, daß die SCHNORR'sche Kritik an der klassischen Maßwahrscheinlichkeitstheorie gegenüber unserem System nicht verfängt. Die von uns eingeführte quantitative logische Wahrscheinlichkeit, bzw. die quantitativen logischen Wahrscheinlichkeitsschranken besitzen in ihrer Interpretation als faire Wettquotienten, bzw. als logisch determinierte Schranken für faire Wettquotienten eine präzise Bedeutung, die den Aspekt der Anwendung der Wahrscheinlichkeitstheorie, nämlich die Entscheidungstheorie, voll erfaßt.

e) Im Gegenteil müssen wir darauf hinweisen, daß die SCHNORR'sche Theorie erkenntnistheoretisch unvollständig ist und insbesondere die Reduktion des von SCHNORR attackierten Begriffs der Wahrscheinlichkeit als Maß der Gewißheit auf den Häufigkeitsbegriff nicht möglich ist. - Wie schon in unserer Kritik des Systems von MISES bemerkt, läßt nämlich das System von SCHNORR keine Aussage über die Geschwindigkeit der Konvergenz in einem Kollektiv zu und bleibt damit ohne jede praktische Relevanz.
Das Gegenteil der SCHNORR'schen These ist der Fall: Der Begriff der Wahrscheinlichkeit ist genuin zu definieren als Maß der Gewißheit, etwa präzisiert als fairer Wettquotient, und läßt sich

nicht auf "definitive" Gegebenheiten, wie etwa einen Limes im Unendlichen reduzieren.
Der Häufigkeitsbegriff, wie ihn MISES oder SCHNORR ausführen, bleibt unbrauchbar ohne ein zusätzliches Axiom, das den Limes der relativen Häufigkeiten als Maß der Gewißheit kennzeichnet, präzisiert etwa als fairer Wettquotient.
Wie schon in unserer Kritik am System von MISES bemerkt, ist ein solches Axiom nur auf der Grundlage des Begriffs der Gleichartigkeit möglich; wie gezeigt, genügt dann dieser Begriff zusammen mit dem irreduziblen Begriff der Wahrscheinlichkeit als Maß der Gewißheit, präzisiert als fairer Wettquotient, um alles abzuleiten, was man haben will; dies wollen wir im folgenden Abschnitt noch belegen.

I) Der Zusammenhang mit unserem System

Wir wollen abschließend noch darstellen, wie sich das Regellosigkeitsprinzip in unserem System ausnimmt, welche Stellung wir dem Gesetz vom iterierten Logarithmus zuweisen können und welche Bedeutung die SCHNORR'schen Folgen haben.

a) Wir wollen zunächst die Ergebnisse von (3.1) rekapitulieren:

Wir wählen eine beliebige, feste Realisierungsfolge $\{\lambda_i\}$ mit $\lambda_i \in \Lambda$. $h_n^{\{\lambda_i\}}(E,\omega)$ seien die relativen Häufigkeiten in dieser Folge.

Dann ist $p(E,\omega) := \begin{cases} \lim\limits_{n\to\infty} h_n^{\{\lambda_i\}}(E,\omega) & \text{, sofern dieser Limes existiert} \\ 0 & \text{sonst} \end{cases}$

definiert, sowie die Zahl $P(E) = p(E,\hat{\omega})$ begrifflich festgelegt.

Es gilt für eine beliebige Folge $\{\lambda_i'\}$:

$$l_K(\{\omega: h_n^{\{\lambda_i'\}}(E,\omega) \to p(E,\omega)\}) = 1 \,, \quad l_K(h_n^{\{\lambda_i'\}}(E) \to P(E)) = 1 \,.$$

Für eine abzählbare Klasse F von Folgen $\{\lambda_i'\}$ gilt:

$$l_K\Big(\bigwedge_{\{\lambda_i'\}\in F} \{\omega: h_n^{\{\lambda_i'\}}(E,\omega) \to p(E,\omega)\}\Big) = 1 \,,$$

$$l_K\Big(\bigwedge_{\{\lambda_i'\}\in F} (h_n^{\{\lambda_i'\}}(E) \to P(E))\Big) = 1 \,.$$

b) Wir betrachten eine Folge $\{\lambda_i\}$ und dazu eine Auswahlregel $\overline{\Phi}$. Die Auswahlregel $\overline{\Phi}$ erzeugt aus $\{\lambda_i\}$ abhängig von ω eine neue Folge $\{\lambda_i'\} = \overline{\Phi}(\{\lambda_i\},\omega)$. Für ein abzählbares System A von Auswahlregeln gilt dabei gemäß dem WALD'schen Satz:

$$p_x\Big(\bigwedge_{\overline{\Phi}\in A} \{\omega: h_n^{\overline{\Phi}(\{\lambda_i\},\,)}(E,\omega) \to x\}\Big) = 1 \,.$$

Daraus folgt, wegen $p_x(\{\omega: p(E,\omega) = x\}) = 1$:

$$p_x\Big(\bigwedge_{\overline{\Phi}\in A} \{\omega: h_n^{\overline{\Phi}(\{\lambda_i\},\omega)}(E,\omega) \to p(E,\omega)\}\Big) = 1 \,.$$

Durch Integration dieser Gleichung nach $d\phi$ ergibt sich:

$$l_K\Big(\bigwedge_{\overline{\Phi}\in A} \{\omega: h_n^{\overline{\Phi}(\{\lambda_i\},\omega)}(E,\omega) \to p(E,\omega)\}\Big) = 1 \,, \text{ bzw.}$$

$$l_K\Big(\bigwedge_{\overline{\Phi}\in A} (h_n^{\overline{\Phi}(\{\lambda_i\},\hat{\omega})}(E) \to P(E))\Big) = 1 \,.$$

Das auf eine abzählbare Klasse von Auswahlfunktionen reduzierte Regellosigkeitsprinzip gilt also in unserem System mit quantitativer logischer Wahrscheinlichkeit Eins.

c) Für die Klasse $\mathfrak{K}$ der binären Zufallsfolgen von SCHNORR gilt nach F) : $p_x(\mathfrak{K}) = 1$. Daraus folgt:
$l_K(\mathfrak{K}) = 1$, bzw. $l_K(\hat{\omega} \in \mathfrak{K}) = 1$.

Die Aussage, daß in der Realität eine SCHNORR'sche Folge vorliegt, hat also in unserem System die quantitative logische Wahrscheinlichkeit Eins.

d) Schließlich wollen wir noch die Stellung des Gesetzes vom iterierten Logarithmus in unserem System betrachten.
Das Gesetz lautet:

$$p_x\left(\{\omega: \left(\overline{\lim_{n\to\infty}}\right) \frac{nh_n(E,\omega) - nx}{\sqrt{x-x^2}\sqrt{2n\log\log n}} = \left(\overset{+}{-}\right) 1\}\right) = 1 .$$

Daraus folgt, wegen $p_x(\{\omega: p(E,\omega) = x\}) = 1$:

$$p_x\left(\{\omega: \left(\overline{\lim_{n\to\infty}}\right) \frac{nh_n(E,\omega) - np(E,\omega)}{\sqrt{p(E,\omega) - p(E,\omega)^2}\sqrt{2n\log\log n}} = \left(\overset{+}{-}\right) 1\}\right) = 1 .$$

Daraus folgt durch Integration nach $d\phi$ und Übergang von der statistischen Wahrscheinlichkeitsfunktion zur statistischen Wahrscheinlichkeit:

$$l_K\left(\left(\overline{\lim_{n\to\infty}}\right) \frac{nh_n(E) - nP(E)}{\sqrt{P(E) - P(E)^2}\sqrt{2n\log\log n}} = \left(\overset{+}{-}\right) 1\right) = 1 .$$

Das Gesetz vom iterierten Logarithmus gilt also in unserem System mit quantitativer logischer Wahrscheinlichkeit Eins.

§11 Schlußbemerkung

Der Grundbegriff der vorliegenden Arbeit war der Begriff der Konzeption K . Alle weiteren fundamentalen Begriffe ergaben sich durch sukzessive "Ableitung" aus dem Begriff K ; zunächst der Begriff der auf Grund von K bestehenden komparativen Struktur "$\underset{\sim}{\leq}_K$" , dann daraus der Begriff der möglichen Maßfunktion m_K zu K , daraus wieder der Begriff l_K der quantitativen logischen Wahrscheinlichkeit, so wie der Begriff der quantitativen logischen Wahrscheinlichkeitsschranken.
Das Ergebnis der vorliegenden Arbeit bestand im Wesentlichen in der folgenden

These: Aus der Konzeption K der Gleichartigkeit folgt das statistische Verhalten der Phänomene; insbesondere läßt sich auf Grund der Konzeption der Gleichartigkeit ein Begriff der statistischen Wahrscheinlichkeit P definieren. Es besteht zwar keine logische Gewißheit auf Grund von K , jedoch eine damit praktisch gleichwertige, als absolut zu bezeichnende Sicherheit, daß P alle von dem Begriff der statistischen Wahrscheinlichkeit üblicherweise gewünschten Eigenschaften hat, etwa, daß P additiv ist, oder daß die relativen Häufigkeiten gegen P konvergieren, etc.
Die Konzeption K ist dabei nichts Absolutes, sondern ein wissenschaftlicher Ansatz, eine Oberhypothese, die sich durch Beobachtungen überprüfen läßt.

Wie schon im Verlauf der Arbeit mehrfach bemerkt, sind die oben genannten Begriffe und die damit geführten Untersuchungen in hohem Maße erweiterungsfähig. Diese Erweiterungen sollen einem im Anschluß an diese Dissertation verfaßten Buch vorbehalten bleiben. - Dabei sind, zusätzlich zu den schon gegebenen Hinweisen, unter anderem die folgenden Punkte ins Auge zu fassen:

(11.1) Die σ-Additivität von P

Untersuchungen zu der Aussage

l_K(P ist ein Maß auf $\mathcal{E}$) = 1 ,

für den Fall, daß der Definitonsbereich $\mathcal{E}$ von P , der σ-Körper $\mathcal{E}$ der möglichen Ereignisschemata zu dem experimentellen Schema B , nicht endlich ist.

Die obige Aussage ist äquivalent zu der folgenden Aussage:

$l_K(\{\omega\colon p(\Sigma A_i,\omega) = \Sigma p(A_i,\omega)$ für alle $A_i \in \mathcal{E}\}) = 1$.

Die Untersuchungen zu dieser Aussage erscheinen schwierig. Sofort ableitbar auf Grund der Konzeption der Gleichartigkeit ist im Hinblick auf (3.1) und die σ-Additivität der relativen Häufigkeiten die folgende Aussage:

$l_K(\{\omega\colon p(\Sigma A_i,\omega) = \Sigma p(A_i,\omega)\}) = 1$.

(11.2) Genauer Bezug zu anderen Wahrscheinlichkeitssystemen

Ausarbeitung des genauen Bezugs zwischen den hier definierten Begriffen P , m_K , etc. und den Begriffen von RICHTER , wie objektive Wahrscheinlichkeit, Glaubwürdigkeit und Chance. Genaue Klärung der Unterscheidung Wahrscheinlichkeit$_1$ und Wahrscheinlichkeit$_2$ von CARNAP. Genauer Bezug zu dem System der objektiven Wahrscheinlichkeit von KOLMOGOROFF.

(11.3) Untersuchungen zu nicht gleichartigen Konzeptionen

Etwa Untersuchungen zu stationären Folgen, zu MARKOFF'schen Prozessen, zu Konzeptionen,für die mit quantitativer Wahrscheinlichkeit Eins die MISES-Axiome erfüllt sind, etc.

(11.4) Umkehrsatz zur statistischen Wahrscheinlichkeit

Umkehrung der oben gegebenen These in dem Sinn, daß sich nur unter der Voraussetzung der Konzeption der Gleichartigkeit eine statistische Wahrscheinlichkeit P definieren läßt; siehe dazu Seite 57 .

(11.5) Spezialuntersuchungen zur Theorie von CARNAP

Spezialuntersuchungen zu den Konzeptionen K_C , der von CARNAP erarbeiteten Bestätigungs- und Maßfunktionen; etwa Ermittlung von $\mathcal{L}_{K_C}$ und l_{K_C} .

a) Insbesondere Untersuchung von K_{C_λ} , der gleichartigen Konzeption der CARNAP'schen m_λ-Funktionen; siehe [C] .

b) Untersuchungen zu den verschiedenen nicht gleichartigen Konzeptionen K_C , der von CARNAP erarbeiteten,in [C'] und [C''] veröffentlichten Maß- und Bestätigungsfunktionen M , C .

(11.6) Spezialuntersuchungen zur Quantenmechanik

Spezielle Entwicklung der obigen Theorie für die Quantenmechanik, wobei das besondere der Quantenmechanik ist, daß der zu dem Experiment B(λ) gehörende Ereigniskörper $\mathcal{V}(\lambda)$ ein quantenmechanischer Verband, aber keine σ-Algebra ist.
Zu der in dieser Arbeit bereits behaupteten Gültigkeit unserer Theorie für die Quantenmechanik und insbesondere für das gebrachte Photonenbeispiel bemerken wir das Folgende:
Die in §3 gemachten Aussagen, etwa die Aussage,
$l_K(h_n(E) \to P(E)) = 1$,
ergeben sich auch für die Quantenmechanik aus der vorliegenden Arbeit. - Zur Ableitung dieser Aussagen,- vom Additionssatz abgesehen, den wir hier ausklammern wollen -, benötigen wir nämlich nur den folgenden Produkt-σ-Körper
$\mathcal{K}(E) := \prod_{\Lambda}^{\times}\{E(\lambda),\overline{E}(\lambda),\emptyset,\Omega(\lambda)\}$,
der auch quantenmechanisch immer sinnvoll ist, vorausgesetzt, daß $E(\lambda)$ und $\overline{E}(\lambda)$ quantenmechanisch sinnvolle Aussagen darstellen.
Speziell für das Photonenbeispiel ist $\Lambda = N$ die Folge der gleichartig polarisierten Photonen und $E(n)$, bzw. $\overline{E}(n)$ das Ereignis, daß das n-te Photon das Filter passiert, bzw. nicht passiert.

An dem Photonenbeispiel wurde illustriert, daß sich in der Quantenmechanik die Konzeption der Gleichartigkeit mit besonderer Stringenz ergibt.
Nur wenn wir analog zu (11.1) ganz allgemein für die Quantenmechanik die Frage untersuchen:

$l_K(\ P$ ist σ-additiv auf $\mathcal{V}\ ) = 1$,

wobei $\mathcal{V}$ der Verband der möglichen Ereignisschemata zu dem experimentellen Schema B der Quantenmechanik ist, ergibt sich die angedeutete Notwendigkeit von Spezialuntersuchungen für die Quantenmechanik.

(11.7) Untersuchungen zum HUME'schen Problem

a) Zur Charakterisierung des HUME'schen Problems, dh. des Problems des induktiven Schließens, legen wir die Arbeit von STEGMÜLLER [St] zu Grunde, außerdem verweise ich auf meine Diplomarbeit [H_0], in der die Problematik induktiver Schlüsse an den Systemen von CARNAP und RICHTER dargestellt ist.
Das HUME'sche Problem läßt sich entsprechend [St] dadurch charakterisieren, daß gemäß der HUME'schen Argumentation ein Schluß vom Beobachteten auf das Unbeobachtete nicht möglich ist.
Logisch ist ein solcher Schluß nicht möglich, die Heranziehung von Prinzipien, wie etwa des Uniformitätsprinzips, involvieren einen erkenntnistheoretischen Zirkel oder unendlichen Regreß und Wahrscheinlichkeitsschlüsse auf Unbeobachtetes sollen zumindest laut HUME auch nicht möglich sein.
Wir wollen zunächst an einem Beispiel illustrieren, wie sich im Rahmen unserer Begriffsbildungen eine gewisse Lösung des HUME'schen Problems ergibt, nämlich des Schlusses vom Beobachteten auf das Unbeobachtete, zwar nicht in Strenge, aber doch mit beliebig hoher Wahrscheinlichkeit.

b) Für eine Ereignisgesamtheit $\mathfrak{E} = \{E(\lambda) : \lambda \in \Lambda\}$ sei die Konzeption K der Gleichartigkeit vorausgesetzt.
Außerdem treten nur "deterministische" Beobachtungen auf, dh. es gelte stets: $h_n(E) = 1$.

Dann ergibt sich aus dem Analogon zur BERNOULLI'schen Ungleichung:

$m_K(|1 - P(E)| \leq \varepsilon) \geq 1 - 1/4n\varepsilon^2$.

Daraus läßt sich durch eine hinreichende Anzahl n von Beobachtungen eine beliebige Sicherheit erzielen, so daß gilt:

$P(E) \geq 1 - \varepsilon$.

Auf Grund solcher Beobachtungen können wir deshalb von der Konzeption K zu der Konzeption $K_{1-\varepsilon} := K \wedge (P(E) \geq 1-\varepsilon)$ gemäß (1.10) übergehen.
Eine kurze Überlegung zeigt: Auf Grund der Konzeption $K_{1-\varepsilon}$ gilt: $m_{K_{1-\varepsilon}}(E(\lambda_o)) \geq 1 - \varepsilon$ im Sinne einer quantitativen logischen Wahrscheinlichkeitsschranke auf Grund von $K_{1-\varepsilon}$.
$E(\lambda_o)$ ist dabei ein beliebiges unbeobachtetes Ereignis aus $\mathfrak{E}$.

Zusammenfassend läßt sich also sagen: Auf Grund der Konzeption der Gleichartigkeit läßt sich durch hinreichend viele Beobachtungen eine beliebige Sicherheit erreichen, um das unbeobachtete Ereignis $E(\lambda_o)$ zu prognostizieren, falls alle beobachteten $E(\lambda)$ eingetreten sind. *)

Damit ist das HUME'sche Problem einer gewissen Lösung zugeführt. Diese Lösung ist nicht absolut, da die Konzeption der Gleichartigkeit nichts Absolutes ist, sondern ein wissenschaftlicher Ansatz, der auch aufgegeben werden kann.
Die Konzeption K stellt eine Art von Uniformitätsprinzip dar, auf Grund dessen das HUME'sche Problem lösbar ist. Welche Stellung wir diesem "Uniformitätsprinzip" K zumessen wollen, soll unter g) diskutiert werden.

c) Es soll allerdings bemerkt werden, daß eine absolute Lösung des HUME'schen Problems unmöglich erscheint und nur Lösungen unter gewissen, nicht als absolut zu bezeichnenden, möglicherweise ganz verschiedenartigen Voraussetzungen angestrebt werden können.

---

*) Gemäß (10.1I) wird diese Überlegung als objektivistisch bezeichnet; die gemäß (10.1I) bestehende subjektivistische Entsprechung wird hier nicht dargestellt.

Die Konzeption der Gleichartigkeit scheint dabei eine besonders einfache, fundamentale und häufige Voraussetzung zu sein, unter der das HUME'sche Problem, wie oben ausgeführt, als lösbar erscheint.

d) Eine absolute Lösung des HUME'schen Problems, ohne irgendwelche Voraussetzungen erscheint uns unmöglich und zwar nicht nur in dem Sinne, wie in [ST] als theoretisches Problem definiert, was die Bewährung von Theorien im Sinne POPPERS (siehe [P]) betrifft, sondern auch des gemäß in [ST] als Problem des induktiven Räsonierens definierten Problems, dh. der Entscheidung unter Risiko oder explizit geasgt, der Angabe von fairen Wettquotienten für unbeobachtete Ereignisse.
Zur Illustrierung dieser These sei das folgende Ergebnis meiner Diplomarbeit referiert:

e) Das System von RICHTER [R] versucht in der folgenden Weise das HUME'sche Problem zu lösen: Bezüglich einer fingierten Situation des absoluten Nichtwissens ist auf den konzipierten Hypothesen in dieser Situation eine dieses Nichtwissen repräsentierende Glaubwürdigkeitsbewertung ϕ_o in Form einer "Gleichverteilung" über diesen Hypothesen gegeben. RICHTER entwickelt eine exakt formulierte Vorschrift F, nach der bei Eintritt eines Ereignisses E eine Glaubwürdigkeitsbewertung ϕ in eine Glaubwürdigkeitsbewertung $\phi^* = F(E,\phi)$ abzuändern ist; die Vorschrift F ist dabei eine Verallgemeinerung der BAYES'schen Regel.
Die Lösung des HUME'schen Problems stellte sich RICHTER so vor, daß unsere gegenwärtigen Vorstellungen ϕ über die Natur aus dem Nichtwissen ϕ_o gemäß der Formel $\phi = F(E,\phi_o)$ überschlägig abgeleitet werden können.
Fatalerweise zeigt meine Diplomarbeit, daß für eine Gleichverteilung ϕ_o, die dem Nichtwissen entspricht, gewissermaßen gilt: $F(E,\phi_o) = \phi_o$ für alle E.

Da das System von RICHTER sehr allgemein ist, scheint dieser Weg einer "absoluten" Lösung des HUME'schen Problems prinzipiell zum Scheitern verurteilt zu sein.

f) Auch die CARNAP'schen Theorien ergeben keine absolute Lösung des HUME'schen Problems, auch wenn man nur von der Entscheidung unter Risiko spricht und nicht die anspruchsvollere Frage der Annahme von Theorien stellt.
Die CARNAP'schen Wahrscheinlichkeitsfunktionen setzen nämlich, wie etwa in (9.2) am Beispiel der c_λ-Funktionen bemerkt, Prinzipien voraus, die als empirisch oder synthetisch zu bezeichnen sind, wie etwa die Konzeption der Gleichartigkeit. Auch die in [C''] veröffentlichten Wahrscheinlichkeitsfunktionen CARNAPS sind in diesem Sinne nicht absolut, sondern setzen etwas über die Wirklichkeit voraus.

g) Es erhebt sich nun die Frage, woher kommen diese benötigten Prinzipien, wie etwa die Konzeption der Gleichartigkeit.
Darauf geben wir die folgende Antwort:
Die Konzeption K ist eine wissenschaftliche Hypothese, die wir unter gewissen Bedingungen ansetzen und die sich im Sinne POPPERS an der Wirklichkeit bewähren kann. Das heißt, auch das induktive Räsonieren, wie STEGMÜLLER es nennt, setzt Hypothesen voraus, die wir nur im Sinne POPPERS einer Bewährungsprobe unterziehen können. Im Unterschied zu den POPPER'schen Ausführungen ist jedoch die Konzeption K nicht nur nicht verifizierbar, sondern auch nicht definitiv falsifizierbar, kann also nur einer mehr oder weniger positiven Bewährungsprobe unterzogen werden.
Es ergibt sich also eine gewisse Verschmelzung von CARNAP und POPPER.

h) Als weiteres Forschungsprogramm ist somit die Ausarbeitung eines Bewährungsgrades für Konzeptionen ins Auge zufassen.
Dabei sind wir der Meinung, daß einem solchen Bewährungsgrad mehr oder weniger direkt eine additive Glaubwürdigkeitsbewertung von Konzeptionen zu Grunde gelegt werden muß.

i) Die Formalisierung des Prozesses, wie eine Konzeption K durch eine andere $\overline{K}$ ersetzt wird, wäre ein weiteres in diesem Zusammenhang zu untersuchendes Problem. - Dabei scheint die Entscheidung zwischen K und $\overline{K}$ gemäß (10.1K) vollständig erfaßt zu sein. Das Problem liegt darin, zu beschreiben, wie bei einer einmal akzeptierten Konzeption K allmählich eine andere

Konzeption $\overline{K}$ sichtbar wird. Der Hinweis auf die Genialität des Forschers scheint dabei wissenschaftstheoretisch unbefriedigend zu sein.
Auch um diesem Problem beizukommen, scheint sich die Übertragung des RICHTER'schen Begriffs der Glaubwürdigkeitsbewertung auf eine Menge von vorgegebenen Konzeptionen anzubieten.
Ein wichtiger Begriff scheint dabei die Gesetzesartigkeit von Konzeptionen zu sein.

j) Der CARNAP'sche Ansatz scheint im Zusammenhang mit unseren Begriffsbildungen erweiterungsfähig zu sein.
Wir betrachten ein Sprache L solcher Allgemeinheit, daß sich etwa die Konzeption K in L ausdrücken läßt.
Dann läßt sich für gewisse Ereignisse E, die ebenfalls in L bezeichenbar sein sollen, wie folgt eine Funktion $c(E,K)$ definieren:

$$c(E,K) := l_K(E) \quad .$$

In diesem Sinn haben wir, was die Untersuchungen zum HUME'schen Problem betrifft, die folgende Zielvorstellung:
Alle bislang von der Menschheit gemachten Induktionsüberlegungen sollen in eine entsprechend reichhaltige Sprache L eingebettet werden. Diese Induktionsüberlegungen sollen durch eine Funktion $c(h,e)$ erfaßt werden, die für gewisse Sätze h, e aus L definiert ist und die sich als logische Wahrscheinlichkeit von h im Hinblick auf e deuten läßt, gemäß dem angegebenen Beispiel: $c(E,K) = l_K(E)$.
In dieser modifizierten Gestalt scheint der CARNAP'sche Ansatz nach wie vor das Ideal dafür zu sein, wie sich das HUME'sche Problem klären läßt.
Als illusorisch ist allerdings die Hoffnung CARNAPs anzusehen, daß sich eine derartige Funktion $c(h,e)$ auf eine in L definierte A-priori-Wahrscheinlichkeit m stützen läßt, gemäß der einfachen Formel: $c(h,e) = m(h \wedge e)/m(e)$.
Die Funktion $c(h,e)$ läßt sich sicher nur für spezielle e, h definieren und benötigt dazu spezielle, in jedem Fall verschiedene auf e, h bezogene Überlegungen, entsprechend dem gegebenen Beispiel: $c(E,K) = l_K(E)$.

Die angegebenen Ausblicke (11.1) - (11.7) sind natürlich in keiner Weise vollständig, sondern sollen nur die Tragfähigkeit der entwickelten Begriffsapparatur im Hinblick auf die angedeutete Erweiterungsfähigkeit sichtbar machen.

Liste der in der Arbeit eingeführten Symbole

Grundsymbole

E , A , B , C , ... beliebige Ereignisse ..... Seite 6

$(\Omega,\mathcal{K})$ zu Grunde gelegter Meßraum ..... Seite 6

$A \lesssim_K B$ A ist auf Grund von K höchstens so wahrscheinlich wie B ..... Seite 7

$A \sim_K B$ A ist auf Grund von K gleichwahrscheinlich zu B ..... Seite 8

$\mathcal{H}_K$ Menge der Paare (A,B) , so daß $A \lesssim_K B$ ..... Seite 10

m_K Maßfunktion zu K ; ϕ zugeordnete DE FINETTI-Bewertung ..... Seite 10,12

l_K quantitative logische Wahrscheinlichkeit ..... Seite 19

$\mathcal{L}_K$ Definitionsbereich von l_K ..... Seite 19

$p_x = \mu_x$ Produktmaß auf $\Pi^{\times}\{E(\lambda),\overline{E}(\lambda),\emptyset,\Omega(\lambda)\}$ mit $p_x(E(\lambda)) = x$, $p_x(\overline{E}(\lambda)) = 1-x$ ..... Seite 12,44

$p_{x_1,\ldots,x_k}$ Produktmaß auf $\Pi^{\times\, B}\{E_1(\lambda),\ldots,E_k(\lambda)\}$ mit $p_{x_1,\ldots,x_k}(E_i(\lambda)) = x_i$ und $\Sigma x_i = 1$ ..... Seite 31

Symbole zu Wetten

$\mathfrak{J}$ Bezeichnung für das vorausgesetzte Individuum ..... Seite 15

W Wette, $\overline{W}$ inverse Wette ..... Seite 15,16

$(W_1,\ldots,W_r) = \mathcal{W}$ Wettsystem ..... Seite 15

$\mathcal{W} \trianglelefteq \mathcal{W}'$ Wettsystem $\mathcal{W}'$ wird $\mathcal{W}$ vorgezogen ..... Seite 15

$\mathcal{W} \equiv \mathcal{W}'$ Wettsysteme sind gleich gut ..... Seite 15

$\mathcal{W} \trianglelefteq \mathcal{W}'$ Wettsystem $\mathcal{W}$ höchsten so gut wie $\mathcal{W}'$ ..... Seite 16

$q(A)$ fairer Wettquotient auf Ereignis A ..... Seite 17

$q_{\mathfrak{J}}(A)$ fairer Wettquotient des Individuums $\mathfrak{J}$ ..... Seite 17

Symbole zu einem allgemeinen experimentellen Schema

$B(\lambda)$ konkretes Experiment zum Parameter λ ..... Seite 27

$B = \{B(\lambda) : \lambda \in \Lambda\}$ experimentelles Schema ..... Seite 27

$E_i(\lambda)$ atomare Ereignisse zu $B(\lambda)$ ..... Seite 27

$E(\lambda)$ beliebiges Ereignis zu $B(\lambda)$ ..... Seite 28

$\mathcal{E} = \{E(\lambda) : \lambda \in \Lambda\}$ Ereignisschema zu B ..... Seite 28

Algebra der Ereignisschemata

$\mathcal{E} + \mathcal{E}' = \{E(\lambda) + E'(\lambda) : \lambda \in \Lambda\}$ ..... Seite 39

$\mathcal{E}.\mathcal{E}' = \{E(\lambda).E'(\lambda) : \lambda \in \Lambda\}$ ..... Seite 52

Symbole zur statistischen Wahrscheinlichkeit

$h_n(\mathcal{E},\omega)$ relative Häufigkeit des Ereignisschemas $\mathcal{E}$ in der fest vorgegebenen Realisierungsfolge $B(\lambda_i)$ ..... Seite 33

$p(\mathcal{E},\omega)$ statistische Wahrscheinlichkeitsfunktion ..... Seite 33

$\hat{\omega}$ wahres, unbekanntes Element aus Ω ..... Seite 37

$\hat{A}$ Aussage, daß das Ereignis A eingetreten ist ..... Seite 37

$m_K(\hat{A}) := m_K(A)$ Maßfunktion, übertragen auf die dem Ereignis A zugeordneten Aussagen ..... Seite 37

$l_K(\hat{A}) := l_K(A)$ quantitative logische Wahrscheinlichkeit, übertragen auf Aussagen ..... Seite 37

$P(\mathcal{E}) := p(\mathcal{E},\hat{\omega})$ statistische Wahrscheinlichkeit ..... Seite 37

$h_n(\mathcal{E}) := h_n(\mathcal{E},\hat{\omega})$ tatsächlich auftretende relative Häufigkeit ..... Seite 37

Symbole zu Konzeptionen

$K_p := K \wedge (P(\mathcal{E}) = p)$ ..... Seite 44

analog definiert:

$K_{p(\mathcal{E},\omega)}$ ..... Seite 50

$K_{P(\mathcal{E})} := K_{p(\mathcal{E},\hat{\omega})}$ ..... Seite 50

K^* Konzeption mit Maßfunktionen, so daß die zugeordneten DE FINETTI-Maße ϕ^* eine stetige, positive Dichte haben. Seite 46

$K^*_{n,r_n} := K^* \wedge (h_n(E) = r_n/n)$ ..... Seite 47

ϕ^*_{n,r_n} das gemäß dem DE FINETTI-Theorem der Konzeption K^*_{n,r_n} zugeordnete Maß ..... Seite 47

Φ die dem Maß ϕ zugeordnete Verteilungsfunktion ..... Seite 48

$r_n(\omega)$ absolute Häufigkeit von E in einer vorgegebenen Folge $B(\lambda_i)$ nach n Beobachtungen, abhängig von dem auftretenden Ergebnis ω ..... Seite 50

$n = n = n$ bezeichnet gleichbedeutend die Zahl der Beobachtungen.

$\phi^*_n(\omega) := \phi^*_{n,r_n}(\omega)$ ..... Seite 50

$K^*_n(\omega) := K^*_{n,r_n}(\omega)$ ..... Seite 50

$K^*_n := K^*_n(\hat{\omega}) = K^*_{n,r_n}(\hat{\omega})$ ..... Seite 50

$m_n(\omega) := m_{K^*_{n,r_n}(\omega)}(E(\lambda_{n+1}))$ bedingte subjektive Wahrscheinlichkeit..... Seite 51

$m_n := m_n(\hat{\omega})$ tatsächlich auftretenden bedingte Wahrscheinlichkeit ..... Seite 51

$K_{h_n,\varepsilon}$ Konzeption der Gleichartigkeit mit Maßfunktionen, wobei das zugeordnete ϕ einen Träger hat, der im Intervall $[h_n-\varepsilon, h_n+\varepsilon]$ liegt ..... Seite 76

K^*_γ Konzeption K^* mit Schranke γ für Maßfunktionen .... Seite 78

$K^*_{\gamma,n}$ Konzeption K^*_γ vermehrt um die Information $h_n = r_n/n$ Seite 78

Weitere Symbole

$f_{n,r_n}(x) := p_x(\{\omega: h_n(E,\omega) = r_n/n\})$ Likelihoodfunktion nach Auftritt der Beobachtung $h_n = r_n/n$ ..... Seite 78

D_B statistisches Datum der Binomialverteilung ..... Seite 73

$D_{B,h_n,\varepsilon}$ statistisches Datum der Binomialverteilung, wobei der Parameter x der Binomialverteilung p_x im Intervall $[h_n-\varepsilon, h_n+\varepsilon]$ liegt ..... Seite 75

$Ea := \int a \, dm_K$ bezeichnet den Erwartungswert der Zufallsgröße a bezüglich einer Maßfunktion m_K ..... Seite 40

LITERATURVERZEICHNIS

[C] CARNAP,R. : STEGMÜLLER,W. : Induktive Logik und Wahrscheinlichkeit.
(Wien 1958)

[C'] CARNAP,R. : JEFFREY, C.R. : Studies in Inductive Logic and Probability. Volume I .
(University of California Press. Berkely . Los Angeles. London. 1971)

[C''] CARNAP,R. : JEFFREY, C.R. : Studies in Inductive Logic and Probability. Volume II.
(Berkely. Los Angeles . 1980)

[CH] CHURCH,A.: On the concept of a random sequence.
(Bull.Amer.math.Soc. 46 (1940) , 130-135)

[F] FINETTI,B.DE: La prévision: ses lois logiques, ses sources subjectives.
(Institut H. Poincaré , Vol. VII,1 ;Paris 1936)

[F'] FINETTI,B.DE: Sur la condition d'équivalence partielle.
(Colloque d'Octobre 1937 sur la Théorie des Probabilités; Tome 6 ,p. 5-18 , Paris 1938)

[HA] HACKING,I.: Logic of Statistical Inference.
(Cambridge 1965)

[H_0] HUMBURG,J.: Die Problematik des Induktiven Schließens bei Carnap und Richter.
(Diplomarbeit bei Prof.H. Richter an der Universität München, 1965)

[H] HUMBURG,J.: The Principle of Instantial Relevance.
(in [C'] Seite 225)

[H'] HUMBURG,J.: Wahrscheinlichkeits- und Wettbegriff.
(Metrika 1972, Vol.19 ; Fasc.1)

[K] KOLMOGOROV,A.N.: Foundations of the Theory of Probability.
(New York 1951)

[KY] KYBURG,H.E.: The Logical Foundations of Statistical Inference.
(Dordrecht 1974)

[L] LAPLACE,P.S.: Théorie analytique des probabilités.
(Paris 1812)

[M] MISES,R.v.: Wahrscheinlichkeit, Statistik und Wahrheit.
(Wien 1928 , 1951)

[N] NEUMANN,J.v. : MORGENSTERN,O.: Theory of Games and Economic Behaviour.
(Princeton 1947)

[P] POPPER,K.R.: Logik der Forschung.
(3.Aufl.Tübingen 1969)

[R] RICHTER,H.: Zur Grundlegung der Wahrscheinlichkeitstheorie.
(Mathematische Annalen: Bd. 125,126,128)

[R'] RICHTER,H.: Wahrscheinlichkeitstheorie.
(Berlin 1966)

[R''] RICHTER,H.: Wahrscheinlichkeitstheorie III .
(München Wintersemester 1971/72, Vorlesungsausarbeitung)

[S] SAVAGE,L.J.: The Foundations of Statistics.
(New York 1954)

[SCH] SCHNORR,C.P.: Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Begründung der Wahrscheinlichkeitstheorie.
(Springer-Verlag: Lecture Notes in Mathematics 1970)

[ST] STEGMÜLLER,W.: Das Problem der Induktion: Humes Herausforderung und moderne Antworten.
(1971 by Friedrich Vieweg, Braunschweig. Wissenschaftliche Buchgesellschaft, Darmstadt)

[V] VILLE,J.: Etude critique des la notion des collectifs.
(Paris, Gauthiers-Villars 1939)

[W] WALD,A.: Die Widerspruchsfreiheit des Kollektivbegriffs.
(Ergebnisse eines math.Koll. 8 (1937), 38-72)

Wir zitieren noch eine Originalarbeit von HUME:

HUME,D.: An Inquiry concerning Human Understanding,Sec. IV.

Außerdem verweisen wir auf:

STEGMÜLLER,W.: Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, Band IV. Personelle und Statistische Wahrscheinlichkeit.
(Springer-Verlag Berlin - Heidelberg - New York 1973)