Ökonomische Entscheidungen bei Ungewißheit

von Hans-Werner Sinn

J. C. B. Mohr (Paul Siebeck): Tübingen 1980

Kapitel 1: Das Wahlobjekt bei Unsicherheit

Erstes Kapitel

Das Wahlobjekt bei Unsicherheit

Abschnitt A
Der grundlegende entscheidungstheoretische Ansatz

1. *Die Ordnung der Alternativen*

Die Aufgabe einer Präferenztheorie ist es, allgemeine Kriterien anzugeben, nach denen die Menschen eine Auswahl aus einer Menge einander ausschließender Handlungsalternativen $(a_1, a_2, \ldots, a_m)$ treffen oder aus der Sicht ihrer Präferenzen treffen sollten.

Der hier verfolgte ökonomische Ansatz[1] zur Lösung dieser Aufgabe besteht in der Suche nach einer Bewertungsfunktion $R(.)$, die einem jeden der Handlungsergebnisse $(e_1, e_2, \ldots, e_m)$ eine reelle Zahl mit der Eigenschaft[2,3]

$$(1) \qquad R(e_i)\left\{\substack{>\\=\\<}\right\}R(e_j) \Leftrightarrow e_i\left\{\substack{\succ\\\sim\\\prec}\right\}e_j$$

zuordnet, so daß die zu wählende Handlungsalternative im Prinzip durch das Heraussuchen des größten Zahlenwertes, also

---

[1] Er wurde von PARETO (1906, S. 176) entwickelt. Vgl. dazu auch die "reconsideration" von HICKS und ALLEN (1930).

[2] In dieser Arbeit bedeuten die Symbole $\succ$ „ist besser als", $\prec$ „ist schlechter als", $\sim$ „ist genauso gut wie" und $\Leftrightarrow$ „genau dann, wenn". Die geschweiften Klammern sollen verdeutlichen, daß die eingeschlossenen Symbole, die in gleicher Höhe stehen, zusammengehören; sie über Kreuz zu lesen, ist also nicht gestattet.

[3] Wir sehen die Funktion $R(.)$ für alle in dieser Arbeit zu untersuchenden Entscheidungsprobleme als gegeben an. Da e_i im intertemporalen Teil IV B einen Zeitpfad von Zielgrößen beschreibt, ist allerdings der Fall sich im Zeitablauf ändernder für die aktuelle Entscheidung *abgeleiteter* Präferenzen erfaßt; in der Tat werden wir eine ganz charakteristische Zeitabhängigkeit der abgeleiteten Präferenzen nachweisen. Nicht alle entscheidungstheoretischen Ansätze implizieren eine feste Funktion $R(.)$. So hat beim sog. Minimax-Regret-Prinzip (vgl. Fn. 26, S. 23) von NIEHANS (1948) und SAVAGE (1951) die Größe des Alternativenvorrats einen Einfluß auf $R(.)$, was MILNOR (1954) zu recht kritisiert. Z.B. lautet meine Präferenz Apfel $\succ$ Birne $\succ$ Butterbrot. Kann ich zwischen einem Apfel und einer Birne wählen, nehme ich den Apfel. Kann ich entweder einen Apfel oder eine Birne oder ein Butterbrot bekommen, wähle ich immer noch den Apfel.

(2) $$\max_{a_i} R(e_i)$$

gefunden werden kann. Der Ansatz läßt offen, welche Handlung im Fall eines nicht eindeutigen Maximums, wenn also der höchste Wert $R(.)$ auf mehrerlei Wegen erreichbar ist, gewählt wird. Das ist der Fall, wo man an den Hemdknöpfen abzählt, wie man sich entscheiden soll. Es dürfte schwer fallen, hierfür eine Theorie zu finden. Die Funktion $R(.)$ kann man Nutzenfunktion nennen, sollte sich aber darüber im klaren sein, daß der hier angesprochene Nutzen nur bis auf eine streng monotone Transformation eines subjektiv empfundenen Befriedigungsniveaus bestimmt, also nur ordinaler Natur ist.

Eine unumgängliche Voraussetzung dieses Ansatzes ist es natürlich, daß der Möglichkeitsbereich nur solche Alternativen umfaßt, für die vermöge einer Präferenzordnung eine eindeutige Bewertung möglich ist. Wir sichern diese Voraussetzung durch das fundamentale

Ordnungsaxiom: *Über der Menge der erreichbaren Handlungsergebnisse existiert eine vollständige, schwache Ordnung.*

Es besagt vor allem[4],
- daß der Entscheidungsträger beim Vergleich zweier beliebiger erreichbarer Ergebnisse das Urteil „nicht schlechter als" ($\succsim$) abgeben kann und
- daß aus $e_i \succsim e_j$ und $e_j \succsim e_k$ folgt $e_i \succsim e_k$ (Transitivität).

Man kann sich leicht überlegen, daß das Ordnungsaxiom die Existenz der Präferenzfunktion $R(e_i)$ sichert, obwohl nur eine schwache Ordnung verlangt wird. Es gilt nämlich

(3) $$\begin{aligned} R(e_i) > R(e_j) &\Leftrightarrow e_i \succsim e_j \quad \text{und nicht} \quad e_j \succsim e_i \ , \\ R(e_i) = R(e_j) &\Leftrightarrow e_i \succsim e_j \quad \text{und ebenso} \quad e_j \succsim e_i \ . \end{aligned}$$

---

[4] Eine weitergehende Auflistung der Implikationen des Axioms mit X als dem kartesischen Produkt der Menge aller möglichen Ergebnisse mit sich selbst, Y als der Menge der Ergebnispaare, die der Entscheidungsträger überhaupt mit Hilfe der Relation $\succsim$ ordnen kann, und Y' als dem Konversen von Y lautet:

$$\begin{array}{ll} X \subseteq Y \cup Y' & \text{(Vollständigkeit),} \\ \neg \forall e_i, e_j : e_i \succsim e_j \Rightarrow e_j \succsim e_i & \text{(Nichtsymmetrie),} \\ e_i \succsim e_j \ \text{ und } \ e_j \succsim e_k \Rightarrow e_i \succsim e_k & \text{(Transitivität).} \end{array}$$

Eine Implikation der Vollständigkeit ist die Reflexivität der Relation $\succsim$, d.h. $e_i \succsim e_i$. Die Menge der geordneten Ergebnispaare, für die die (reflexive, symmetrische, transitive) Äquivalenzrelation $\sim$ gilt, ist $Y \cap Y'$. Nennen wir $\bar{Y}'$ die Komplementmenge von Y', dann ist die Menge der geordneten Ergebnispaare, für die die starke Präferenzrelation $\succ$ gilt, $Y \cap \bar{Y}'$. $Y \cap \bar{Y}'$ ist irreflexiv und transitiv. Letzteres impliziert wiederum Nichtsymmetrie von $\succ$. Vgl. dazu z.B. NACHTKAMP (1969, S. 66–81) und FISHBURN (1970, S. 9–15).

Anstatt auf eine schwache Präferenzordnung abzustellen, hätte man natürlich auch gleich eine starke Präferenzordnung fordern können. Doch vermutlich fällt den Menschen das Urteil „nicht schlechter als" leichter als die Urteile „genausogut wie" und „besser als". Im übrigen bringt die obige Formulierung vom behavioristischen Standpunkt den Vorteil, daß sie nur auf die Schlußfolgerungen, die man aus beobachtbaren Wahlhandlungen ziehen kann, zurückgreift: Wohl kann man nämlich aus einer beobachteten Entscheidung schließen, die gewählte Alternative sei nicht schlechter als alle anderen, doch kann man nicht unterscheiden, ob sie besser als oder genausogut wie andere Alternativen eingeschätzt wurde, da schließlich auch im Fall der Indifferenz irgendeine Alternative ergriffen werden muß[5].

So harmlos und selbstverständlich das Ordnungsaxiom vielleicht auch erscheinen mag, sowohl aus positiver wie aus normativer Sicht stellt es eine Idealisierung dar. Sicherlich ist kein Mensch in der Lage, eine völlig konsistente Ordnung über die in der Regel außerordentlich große Zahl der verfügbaren Alternativen zu erstellen[6]. Und selbst wenn er es wäre, würde er sich wohl auch gern einmal mit anderen Dingen beschäftigen, als fortwährend seine Präferenzen zu ordnen. Wir sehen daran, daß ein Mangel des Ordnungsaxioms darin besteht, daß es die *Mühe des Ordnens* vernachlässigt.

In der Praxis führt diese Mühe des Ordnens dazu, daß die Präferenzfunktion $R(.)$ eine stochastische Komponente erhält[7], was bei wenig voneinander abweichenden Handlungsergebnissen zu Intransitivitäten führen kann. Das läßt sich folgendermaßen erklären. Nehmen wir an, der Entscheidungsträger verfüge, ohne eine genaue Berechnung der Rangordnungsfunktion $R(.)$ vorgenommen zu haben, über Vorinformationen über den ungefähren Wert, den diese Funktion für verschiedene Ergebnisse annimmt. Dann ist es sehr gut möglich, daß er beim paarweisen Vergleich der Alternativen e_i, e_j und e_j, e_k zum Urteil kommt, die eine Alternative sei nicht schlechter als die andere, bloß weil der Vorteil, den er sich vom Auffinden der besseren Alternative verspricht, die Mühe des Ordnens nicht zu lohnen scheint. Obwohl in diesem Fall offenkundig $e_i \sim e_j$ und $e_j \sim e_k$ vorliegt, darf man nicht schließen, daß $e_i \sim e_k$, wie es bei einer transitiven Präferenzordnung zulässig wäre. Wird dem Entscheidungsträger nämlich die Aufgabe gestellt,

---

[5] Daher vertritt LITTLE (1950, S. 14–52) den Standpunkt, die Präferenztheorie solle sich von vornherein nur auf Handlungen beziehen. Warum sollte aber nicht mittels direkter Befragung Aufschluß über den Innenzustand des Entscheidungsträgers gefunden werden können?

[6] AUMANN (1962) hat sich aus diesem Grund bemüht, einen präferenztheoretischen Ansatz unter Verzicht auf die Vollständigkeit zu formulieren.

[7] Mit dem Problem der stochastischen Präferenzordnung hat sich auf ökonomischer Seite zunächst GEORGESCU-ROEGEN (1936) auseinandergesetzt. In die Psychologie fanden stochastische Empfindungsfunktionen aber bereits mit einem berühmt gewordenen Aufsatz von THURSTONE (1927) Eingang.

sich zwischen e_i und e_k zu entscheiden, dann interessiert ihn der Vorteil, der aus einer richtigen Feststellung der Präferenzordnung zwischen *diesen* beiden Alternativen entstehen kann. Dieser Vorteil kann sowohl jenen aus der Festlegung der Rangfolge zwischen e_i und e_j als auch jenen aus der Festlegung der Rangfolge zwischen e_j und e_k übersteigen und somit eine genaue Berechnung der Rangfolge zwischen e_i und e_k anregen. Das Resultat dieser Berechnung kann dann durchaus $e_i \nsim e_k$ lauten[8].

Es wäre sicher wünschenswert, eine ökonomische Präferenztheorie zu entwickeln, die auch die Genauigkeit der Ordnung zum Gegenstand des Optimierungsprozesses macht. Doch leider ist eine solche Theorie nicht vorhanden und kann auch hier nicht angeboten werden[9].

2. *Die Handlungsergebnisse bei Unsicherheit*

In einer sicheren Welt ist die Regel $\max_{a_i} R(e_i)$ für das Auffinden einer optimalen Handlung leicht interpretierbar: Hier sind die e_i irgendwelche mit Sicherheit eintretende Ergebnisse, deren Bewertung mit Hilfe der Funktion $R(.)$ keine prinzipiellen Schwierigkeiten bereiten sollte. So können in der Haushaltstheorie die Ergebnisse Konsumgüterbündel sein; e_i beschreibt dann einen Vektor von Gütermengen. Speziell in der Unternehmenstheorie bezeichnet e_i häufig einen in Geldeinheiten gemessenen Gewinn; in diesem Fall reduziert sich unsere Regel auf die altbekannte Zielsetzung der Gewinnmaximierung.

---

[8] Vgl. dazu SCHNEEWEISS (1967a, S. 35f. u. 81–84) und KRELLE (1957, S. 637; 1961, S. 112–116; 1968, S. 21–24). Dort werden die Möglichkeiten einer durch Fühlbarkeitsschwellen verursachten Intransitivität und das Problem der Kalkulationskosten diskutiert. Unsere Argumentation vereint beide Gesichtspunkte, indem die Fühlbarkeitsschwellen mit den Kalkulationskosten erklärt werden. Für eine automatentheoretische Erklärung von scheinbaren Intransitivitäten siehe RÖDDING und NACHTKAMP (1978, 1980).

[9] Unsere Forderung ist nicht mit dem Ziel der durch SIMON (1957, S. 241–260), SIEGEL (1957), SAUERMANN und SELTEN (1962), STARBUCK (1963a u. b) u.a. vertretenen Anspruchsanpassungstheorie zu verwechseln, die den Informationsgewinnungsprozeß für das Auffinden des Möglichkeitsbereichs von Handlungsalternativen mit in die Optimierungsentscheidung einbeziehen will. Im Gegensatz zum ersten Anschein steht diese Theorie nicht in einem prinzipiellen Widerspruch zum Ordnungsaxiom. Man braucht ja nur die verschiedenen Möglichkeiten der Informationsgewinnung als zusätzliche Handlungsalternativen aufzufassen. Auch bei dem auf diese Weise entstehenden sequentiellen Entscheidungsprozeß gibt es dann nämlich zu jedem Zeitpunkt einen festen Möglichkeitsbereich an Handlungsalternativen, aus denen in völliger Übereinstimmung mit dem Ordnungsaxiom die beste auszuwählen ist. Daß die Wahl einer solchen besten Alternative so interpretiert wird, als strebe der Entscheidungsträger nur ein Anspruchsniveau unterhalb des „wahren" Optimums an, ist irreführend.

Was ist nun aber das „Ergebnis" einer Handlung unter Unsicherheit? Denken wir z.B. an einen Unternehmer, der trotz Unsicherheit über die zukünftigen Erträge eines aus einer Menge unterschiedlicher Investitionsprojekte auswählen muß. Könnte das Ergebnis, von dem wir hier sprechen, der im Nachhinein festgestellte Gewinn sein? Das machte nicht viel Sinn, denn die Entscheidung für ein Investitionsprojekt muß ja gefällt werden, bevor man weiß, wie ertragreich es ist. Entscheidungsgrundlage kann also nur ein Ergebnis sein, wie es sich *ex ante* darstellt. Ein solches Ergebnis ist aber kein fester Wert, sondern läßt sich nur als „Zufallsvektor" von möglichen Einzelergebnissen

(4) $$e_i = (e_{i1}, e_{i2}, \ldots, e_{in})$$

darstellen, was TINTNER (1941, S. 301) dazu veranlaßt hat, die Bewertungsfunktion $R(.)$ „Präferenzfunktional" zu nennen. Ein anschauliches Beispiel für einen solchen Zufallsvektor ist ein Lotterielos.

Um zu einem Ergebnisvektor e_i zu gelangen, wird der Entscheidungsträger berücksichtigen müssen, daß das Einzelergebnis nicht nur von seiner eigenen Handlung, sondern auch von den verschiedensten Umwelteinflüssen abhängt, die er selbst weder steuern noch mit Sicherheit voraussehen kann[10]. In übersichtlicher Form kann man die Entscheidungssituation daher in Form einer Fallstudie darstellen, wie sie leicht an Hand der folgenden auf VON NEUMANN und MORGENSTERN (1947) zurückgehenden Ergebnismatrix vorgenommen werden kann.

Tabelle 1

Handlung \ Zustandsklasse der Welt	Z_1	…	Z_j	…	Z_n
a_1	e_{11}	…	e_{1j}	…	e_{1n}
⋮	⋮		⋮		⋮
a_i	e_{i1}	…	e_{ij}	…	e_{in}
⋮	⋮		⋮		⋮
a_m	e_{m1}	…	e_{mj}	…	e_{mn}

Hier kennzeichnen die Größen $(Z_1, \ldots, Z_m)$ einander ausschließende Klassen von Zuständen der Umwelt, die der Entscheidungsträger unterscheiden

[10] Vgl. VON NEUMANN und MORGENSTERN (1947, S. 10f.).

möchte[11]. Der Entscheidungsträger weiß, daß, falls er die Handlung a_i wählt und die Umwelt die Zustandsklasse Z_j annimmt, das Ergebnis e_{ij} eintritt. Er weiß aber nicht, in welche Klasse der wirkliche Umweltzustand fallen wird; das ist das Spezifikum des Entscheidungsproblems unter Unsicherheit.

Die Sachlage wird komplizierter, wenn man das Zeitproblem berücksichtigt. In einer sicheren Welt ändert sich die Struktur des Entscheidungsproblems kaum. Die Handlung a_i beschreibt dort einen Zeitpfad eigener Aktivitäten, denen in eindeutiger Weise ein Ergebniszeitpfad zugeordnet werden kann. Der optimale Aktivitätspfad wird einmal festgelegt und dann ohne eine neue Entscheidung verfolgt. Anders in der stochastischen Welt[12]. Hier wäre es verfehlt, einen einmal gewählten Zeitpfad eigener Handlungen unabhängig von den Ausprägungen des Umweltzeitpfades aufrecht zu erhalten, selbst wenn das möglich wäre. Stellt man sich nämlich die oben beschriebene Ergebnismatrix als für eine Zeitperiode gültig vor, zu deren Anfang gehandelt und zu deren Ende über einen bestimmten Umweltzustand das Periodenergebnis festgelegt wird, dann ist leicht einzusehen, daß die Gestalt dieser Ergebnismatrix in der Regel von dem am Ende der Vorperiode realisierten Zustand der Welt abhängen wird. Daraus folgt, daß die Handlung, die man sich gestern bereits für heute vorgenommen hat, nur per Zufall mit jener übereinstimmt, die man heute bei Kenntnis der Ergebnismatrix als optimal ansieht. Vernünftigerweise schiebt man daher Entscheidungen soweit wie möglich vor sich her. Das heißt freilich nicht, daß man von dem ganzen Zeitproblem einfach abstrahieren dürfte, denn sicherlich muß bei der heutigen Handlung mitbedacht werden, daß ihr Ergebnis den morgen verfügbaren Alternativenvorrat und damit die gesamte der morgigen Entscheidung zugrundeliegende Ergebnismatrix beeinflußt.

Bei diesen Bemerkungen zum Zeitproblem wollen wir es hier bewenden lassen. Der Problemkreis wird im Kapitel IV unter einer etwas stärker spezifizierten Fragestellung wiederaufgegriffen. Bis dahin begnügen wir uns mit der Vorstellung, daß ein einziges Mal eine Wahlhandlung vorzunehmen ist und nach Ablauf einer nicht näher spezifizierten Zeitperiode, innerhalb derer die Entscheidung nicht mehr revidiert werden kann, das Ergebnis bekannt wird. Es wird sich zeigen, daß diese Vorstellung, so unrealistisch sie für sich genommen sein mag, als Baustein eines Mehrperiodenansatzes verwendbar ist.

---

[11] Es kann sich immer nur um Zustandsklassen und nicht etwa um vollständig beschriebene Zustände handeln, denn der Entscheidungsträger wird die Zustände der Welt nur an Hand der ihn interessierenden Kriterien, aber keineswegs an Hand aller Kriterien kategorisieren. Die Unterscheidung ist für die weiter unten gegebene Diskussion des *Bayes-Theorems* von Bedeutung.

[12] Bisweilen wird das Zeitproblem durch die zwar recht trickreiche, doch ein wenig inhaltsleere Konstruktion überwunden, daß die Z_j Ausprägungen von Umweltzeitpfaden und die a_i „Strategien" so etwa im Sinne von Lebensweisheiten, denen man immer folgen sollte, bezeichnen.

Wir sind damit wieder bei der Entscheidungssituation, wie sie durch die obige Ergebnismatrix verkörpert wird. Die Frage ist nun, nach welchen Kriterien die Zeilen dieser Matrix, die das einer jeden Handlung a_i zugeordnete *ex-ante*-Ergebnis e_i beschreiben, zu bewerten sind. Die Antwort wird in zwei Schritten gegeben: In diesem Einführungskapitel soll zunächst die Vorfrage geklärt werden, welche Informationen der Entscheidungsträger über die Zustandsklassen der Welt benötigt, um überhaupt zu einer Entscheidung fähig zu sein. In den Kapiteln II und III geht es dann anschließend um eine möglichst weitgehende Spezifizierung der Bewertungsfunktion $R(e_i)$.

Abschnitt B
Wahrscheinlichkeiten

1. *Wahrscheinlichkeiten als Vertrauensgrade*

Wenn die oben dargestellte Ergebnismatrix richtig spezifiziert wurde, beschreibt das Element e_{ij} wirklich alles, was den Entscheidungsträger an der Situation, die durch das Zusammentreffen der Handlung a_i mit einem Umweltzustand aus der Klasse Z_j entsteht, interessiert. Insofern muß es ihm im nachhinein gleichgültig sein, unter welchem Umweltzustand ein bestimmtes Ergebnis erzielt wurde[1]. Anders *ex ante* in der Entscheidungssituation. Hier ist die Zustandsklasse, unter der ein Einzelergebnis zu erwarten ist, durchaus von Belang, denn sie gibt möglicherweise Aufschluß über den für eine optimale Entscheidung wichtigen Grad des Vertrauens für das Eintreten dieses Einzelergebnisses.

Was unter diesem Vertrauensgrad, den wir g nennen wollen, zu verstehen ist, soll noch geprüft werden. Festlegen wollen wir in diesem Sinne jedoch schon die folgende

Ergänzung zum Ordnungsaxiom: *Das Handlungsergebnis unter Unsicherheit besteht aus einem Zufallsvektor*

$$e_i = \begin{pmatrix} g_1 & g_2 & \dots & g_n \\ e_{i1} & e_{i2} & \dots & e_{in} \end{pmatrix},$$

[1] HIRSHLEIFER (1965, bes. S. 522) hält in einem anderen Zusammenhang eine Komplementarität zwischen Zuständen und Ergebnissen für realistisch. Die Komplementarität, die er im Auge hat, rührt indes nur von der Unzulänglichkeit der, Ergebnisbeschreibung her.

der neben einer Beschreibung von alternativ möglichen Einzelergebnissen auch Angaben über den Grad des Vertrauens, mit dem diese Ergebnisse erwartet werden können, enthält[2].

Bemerkenswert an dieser Ergänzung ist die Tatsache, daß der Weg, auf dem ein bestimmtes Ergebnis erreicht wird, nicht in die Bewertung einfließt. Damit wird implizit der in der Regel bei Glücksspielen zu beobachtende Umstand ausgeschlossen, daß nicht nur das Vertrauen, einen Gewinn zu erlangen, sondern auch die Spielprozedur von Bedeutung ist. Mit der Präferenz für bestimmte, meist langwierige und komplizierte Ausspielungsmodi zeigt sich etwas ähnliches wie bei der Präferenz für einen guten Kriminalroman, wo sich der anfänglich große Täterkreis nur allmählich zum wirklich Schuldigen verdichten darf. Wir vernachlässigen diesen Aspekt der Präferenzen, da er für nüchterne wirtschaftliche Entscheidungen, um deren Analyse es uns geht, keine Bedeutung haben dürfte[3]. Vor einer unbesehenen Anwendung der hier zu entwickelnden Präferenztheorie auf Glücksspiele muß aber gewarnt werden[4].

Gleichwohl werden wir uns in dieser Arbeit des häufigeren fiktiver Spielsituationen bedienen, um Entscheidungen unter Unsicherheit zu analysieren. So z.B. gleich anschließend. Wir wollen aber im Auge behalten, daß immer nur ernsthafte wirtschaftliche Entscheidungen modelliert werden sollen.

Um zu erfahren, was man unter den Vertrauens- oder Glaubwürdigkeitsgraden $(g_1, g_2, \ldots, g_n)$ verstehen kann, betrachten wir das folgende Entscheidungsproblem: Aus einer Urne mit schwarzen und weißen Kugeln wird wahllos eine Kugel gezogen, nachdem der Entscheidungsträger zuvor auf eine Farbe gesetzt hat. Erscheint die gewählte Farbe, dann werden 1000 DM ausgezahlt, andernfalls nichts. Der Entscheidungsträger hat keine Präferenz für eine bestimmte Farbe. Ihm ist bekannt, daß der Anteil der schwarzen Kugeln w_1 und der der weißen Kugeln w_2 beträgt. Die Ergebnismatrix für dieses Problem lautet:

---

[2] Man beachte, daß weder für die Ergebnisse noch für die Vertrauensgrade irgendeine Numerik gefordert wird.

[3] Die fehlende Berücksichtigung des Ausspielungsmodus hat ALLAIS (1952) am von Neumann-Morgenstern-Index kritisiert, für den das Ordnungsaxiom in einer strengeren Version benötigt wird; vgl. VON NEUMANN und MORGENSTERN (1947, S. 26).

[4] Vgl. unsere Kritik an den Ansätzen von TÖRNQVIST (1945), FRIEDMAN und SAVAGE (1948), FRIEDMAN (1952) und MARKOWITZ (1952), die allesamt aus dem Glücksspielverhalten der Menschen auf ihre Bewertung wirtschaftlicher Risiken schließen wollen, im Kapitel III B 1.3. Vgl. auch S. 31f.

Tabelle 2

Wahrscheinlichkeit	w_1	w_2
Zustandsklassen der Welt / Handlungen	Z_1 (schwarze Kugel wird gezogen)	Z_2 (weiße Kugel wird gezogen)
a_1 (es wird auf schwarz gesetzt)	$e_{11} = 1000$ DM	$e_{12} = 0$ DM
a_2 (es wird auf weiß gesetzt)	$e_{21} = 0$ DM	$e_{22} = 1000$ DM

Hier wird deutlich, was unter dem Vertrauen in das Auftreten bestimmter Zustände der Welt verstanden werden kann: Obwohl man bei beiden Handlungen gleichermaßen entweder 1000 DM oder nichts gewinnen kann, ist der Entscheidungsträger sicherlich im allgemeinen nicht indifferent, sondern wird auf die häufiger vertretene Farbe setzen. Damit wählt er die Handlung mit der größeren Gewinnwahrscheinlichkeit, denn die Anteilswerte der Kugeln, w_1 und w_2, geben ja ihre Wahrscheinlichkeit, gezogen zu werden, an.

Gegen die Behauptung, daß Wahrscheinlichkeiten als Glaubwürdigkeitsmaße fungieren können, hat SHACKLE (1952, S. 5f., bes. S. 109–111; 1955, S. 3–16) eingewendet, Wahrscheinlichkeiten seien nur dann von irgendwelcher Bedeutung, wenn der Entscheidungsträger die Möglichkeit erhalte, die gewählten Handlungen mehrfach vorzunehmen, denn in diesem Fall könnten die Wahrscheinlichkeiten wenigstens approximativ als relative Häufigkeiten interpretiert werden, so daß man zu Recht die Alternative, die in der größeren Zahl der Fälle zum Erfolg führe, vorziehe. Doch wenn es, wie hier unterstellt, um eine einmalige Entscheidung gehe, seien die Wahrscheinlichkeiten völlig belanglos, da ja doch immer nur der eine *oder* der andere Fall eintreten könne. Statt an Wahrscheinlichkeiten orientiere sich der Entscheidungsträger am *Grad der potentiellen Überraschung*, die er dem Eintritt eines bestimmten Ergebnisses beimesse. Doch wenn in der Urne z. B. 70 weiße und 30 schwarze Kugeln sind, bei welchem Ziehungsergebnis wäre man dann mehr überrascht, bei „weiß" oder „schwarz"? Sicherlich bei letzterem, denn es hat die kleinere Eintrittswahrscheinlichkeit. Obwohl Shackle selbst es nicht wahrhaben wollte, widerspricht das Grundkonzept der potentiellen Überraschung dem Wahrscheinlichkeitskonzept nicht, sondern liefert ganz im Gegenteil eine anschauliche Interpretation für das beobachtbare Faktum, daß man sich auch in einmaligen Risikosituationen nach Wahrscheinlichkeiten richtet. In dem obigen Beispiel setzt man auf die Farbe, die, sollte sie nicht gezogen werden, für die größere Überraschung sorgt; das ist die Farbe mit der größeren Wahrscheinlichkeit[5].

[5] Wir kommen damit zu der z.B. von KRELLE (1957, S. 648–651) und NACHTKAMP (1969, S. 199f.) vertretenen Interpretation. Vgl. aber die Beiträge von TURVEY (1949)

So bereitwillig man geneigt ist, diesem für eine idealisierte Entscheidungssituation gefundenen Ergebnis zuzustimmen, stellt sich doch sogleich die Frage nach der Übertragbarkeit auf realistischere Fälle. Finden wir in der Wirklichkeit Strukturen, die jenen des kontrollierten Zufallexperimentes entsprechen?

Ramsey (1931), de Finetti (1937, 1952) und Savage (1952, 1954) sehen so wenig Parallelen, daß sie subjektive Wahrscheinlichkeiten als Entscheidungsgrundlage vorschlagen[6]. Mit diesen subjektiven Wahrscheinlichkeiten wird den Vertrauensgraden (g_1, g_2, ..., g_n) anders als bei Shackle eine Numerik zugeordnet, die mehr oder weniger zufällig den Regeln des Wahrscheinlichkeitskalkulus gehorcht und inhaltlich ohne Rückgriff auf objektive Wahrscheinlichkeiten interpretiert werden kann. Mit der Einführung subjektiver Wahrscheinlichkeiten ist die sozialwissenschaftliche Entscheidungstheorie bei Unsicherheit sicherlich ein gutes Stück voran gekommen. Eine grundlegende Schwäche liegt aber in ihrer Natur: Es gibt keine allgemein als richtig angesehenen Wahrscheinlichkeiten mehr. Für eine erklärende Theorie ist das natürlich weniger angenehm, da ihr Aussagegehalt eingeschränkt wird. Besonders kraß zeigt sich das bei Savage (1954, S. 63–67), der sich anders als de Finetti (1937, S. 16–24) und vermutlich auch Ramsey (1931, S. 187f.) nicht einmal für kontrollierte Zufallsexperimente in der Art unseres Urnenbeispiels darauf festlegen lassen möchte, daß der Entscheidungsträger mit objektiven Wahrscheinlichkeiten kalkuliert.

So ist es nicht verwunderlich, daß Luce und Raiffa (1957, S. 299–302), Schlaifer (1959, S. 2–23; 1969, S. 106–127, 201–217), Pratt, Raiffa und Schlaifer (1965, Kap. 2 u. 3) und Raiffa (1968, S. 104–128) eine gewisse Rückbesinnung auf objektive Wahrscheinlichkeiten anstrengen, indem sie den subjektiven Wahrscheinlichkeiten eine engere Interpretation als Savage geben. Sie gehen davon aus, daß der Entscheidungsträger subjektive Vertrauensgrade in *äquivalente objektive Wahrscheinlichkeiten* überführen kann, indem er abzuschätzen versucht, mit welcher relativen Häufigkeit diese Zustände wohl bei einer vielfachen fiktiven Wiederholung der Entscheidungssituation aufträten. Diese Häufigkeitsinterpretation der subjektiven

---

und Graaf und Baumol (1949), wo besonders die Unterschiede zum Wahrscheinlichkeitskonzept herausgestellt werden (z.B. die Nichtaddierbarkeit des Shackleschen Maßes für einander ausschließende Ereignisse).

[6] Neuere mathematische Entwicklungen dieses Ansatzes werden bei Gottinger (1974) diskutiert.

[7] Ganz im Gegensatz zu Keynes (1921, S. 4): "The Theory of Probability is logical, therefore, because it is concerned with the degree of belief which it is *rational* to entertain in given conditions, and not merely with the actual beliefs of particular individuals, which may or may not be rational."

Wahrscheinlichkeiten ist bei Savage sicher unzulässig, doch mit de Finettis Ansatz ist sie vereinbar, und Ramsey empfand sie sogar als nützlich[8].

Die subjektiven Vertauensgrade in äquivalente objektive Wahrscheinlichkeiten umzuwandeln, bringt (mindestens) drei Vorteile: Erstens wird auf ganz natürliche Weise das oben beschriebene Problem beseitigt, daß die subjektive von der objektiven Wahrscheinlichkeit abweichen kann, selbst wenn letztere dem Entscheidungsträger bekannt ist. Zweitens können alle Entscheidungsprobleme unter Unsicherheit *einheitlich* auf die Aufgabe reduziert werden, ein Präferenzfunktional $R(.)$ für das Handlungsergebnis

$$e_i = \begin{pmatrix} w_1 & w_2 & \dots & w_n \\ e_{i1} & e_{i2} & \dots & e_{in} \end{pmatrix} \tag{1}$$

zu finden, wobei e_i der oben eingeführte Zufallsvektor von Einzelergebnissen ist (das „Lotterielos"), nur mit dem Unterschied, daß die Vertrauensgrade $(g_1, g_2, \dots, g_n)$ durch äquivalente objektive Wahrscheinlichkeiten $(w_1, w_2, \dots, w_n)$ ersetzt wurden. Drittens können wir bei Operationen mit diesen Wahrscheinlichkeiten den reichhaltigen Werkzeugkasten der mathematischen Wahrscheinlichkeitstheorie benutzen[9].

Ungeachtet dieser Vorteile einer Ungewißheitstheorie auf der Basis äquivalenter objektiver Wahrscheinlichkeiten leidet speziell der Luce-Raiffa-Schlaifer-Ansatz aber immer noch darunter, daß nicht untersucht wird, *wie* der rationale Entscheidungsträger aus seinen Glaubwürdigkeitsvorstellungen äquivalente objektive Wahrscheinlichkeiten bildet. Im übernächsten Abschnitt B 3 werden wir diese Frage zu beantworten suchen. Eine hierfür allerdings unerläßliche Vorfrage ist die nach einer inhaltlichen Präzisierung des bislang primitiv notierten Begriffs der objektiven Wahrscheinlichkeit. Ihr gehen wir im nun folgenden Abschnitt B 2 nach.

---

[8] Vgl. dazu SAVAGE (1954, S. 4), RAMSEY (1931, S. 158f. und 187f.) und DE FINETTI (1937, S. 18f.). Zur Vorstellung fiktiv wiederholter Entscheidungssituationen sagt RAMSEY (S. 188): "It is this connection between partial belief [im Sinne des Vertrauensgrades; der Verf.] and frequency which enables us to use the calculus of frequencies as a calculus of consistent partial belief. And in a sense we may say that the two interpretations are the objective and subjective aspects of the same inner meaning ..." Die Vereinbarkeit mit de Finettis Ansatz folgt daraus, daß de Finetti postuliert, der Entscheidungsträger müsse mit Hilfe eines objektiven Verfahrens aus seinen subjektiven Wahrscheinlichkeitsvorstellungen die relative Häufigkeit, mit der er für die Zukunft rechnet, herleiten können.

[9] Sie basiert auf nur drei Axiomen KOLMOGOROFFS (1933):

1. Die Wahrscheinlichkeit ist eine reelle Zahl im geschlossenen Einheitsintervall.
2. Die Wahrscheinlichkeit eines sicheren Ereignisses ist 1 und die eines unmöglichen Ereignisses 0.
3. Die Wahrscheinlichkeit für das Auftreten irgendeines aus einer Menge einander ausschließender Ereignisse ist gleich der Summe der Wahrscheinlichkeiten dieser Ereignisse.

2. *Die objektive Wahrscheinlichkeit und die wirkliche Indeterminiertheit*

Betrachten wir einmal die beiden folgenden, von FISHER (1906, S. 266–269) gestellten Aufgaben:

1. Man soll vor dem Wurf einer idealen Münze angeben, wie groß die objektive Wahrscheinlichkeit für das Ereignis „Kopf" ist.
2. Nachdem die Münze bereits geworfen ist, jedoch bevor das Wurfergebnis bekannt ist, soll die objektive Wahrscheinlichkeit für „Kopf" angegeben werden.

Im ersten Fall wird man ohne viel Zögern „1/2" sagen. Vielleicht sagt man das auch im zweiten Fall, doch ein weiteres Nachdenken stiftet einige Verwirrung. Muß nicht in Anbetracht der Tatsache, daß zum Zeitpunkt der Wahrscheinlichkeitsangabe das Wurfergebnis vollständig determiniert ist, die einzig vernünftige Antwort lauten, die objektive Wahrscheinlichkeit sei entweder 0 oder 1? Welchen Sinn macht es, nach dem Luce-Raiffa-Schlaifer-Konzept subjektive Wahrscheinlichkeiten als Schätzwerte von objektiven aufzufassen, wenn der Zufall überhaupt nicht mehr zum Zuge kommt? Muß man nicht folglich die Entscheidungssituationen der Wirklichkeit in zwei Kategorien unterteilen, nämlich eine Kategorie von Entscheidungen, wo Vermutungen von Tatsachen eine Rolle spielen (Gibt es Leben auf dem Mars? Finde ich Öl, wenn ich hier bohre?), und eine solche, bei der das Ergebnis wirklich indeterminiert ist (Wird es morgen regnen? Welche Nachfrage wird sich bei diesem Preis einstellen?)?

Diese Fragen lassen ahnen, daß unsere Vorstellungen, von dem, was unter einer objektiven Wahrscheinlichkeit zu verstehen ist, noch nicht ganz in Ordnung sind. In der Tat haben wir die objektive Wahrscheinlichkeit mit so etwas wie dem „Grad der wirklichen Indeterminiertheit" durcheinandergebracht.

Daß beides nicht dasselbe ist[10], tritt besonders klar hervor, wenn wir einmal einer Fiktion folgen, die von LAPLACE (1814, S. IIf.) entwickelt wurde. Nach dieser Fiktion ist der Ablauf der Welt deterministisch, da nach einem *Prinzip des hinreichenden Grundes* ein jedes Ereignis eine zeitlich vorher liegende Ursache hat. Das „Heute" folgt nach festen Gesetzen aus dem „Gestern" und das „Morgen" nach eben diesen Gesetzen aus dem „Heute". Es gibt weder einen freien Willen der Menschen[11] noch sonst eine Ursache

[10] Eine Unterscheidung zwischen der objektiven Wahrscheinlichkeit und dem Grad der wirklichen Indeterminiertheit wird nicht immer klar getroffen: Vgl. dazu REICHENBACH (1935, bes. S. 8–13), KNIGHT (1921, bes. S. 221f.) und DE FINETTI (1949, S. 91), die beide Größen einander gleichsetzen.

[11] Im freien Willen sieht KNIGHT (1921, S. 221) die eigentliche Ursache der Indeterminiertheit des Ablaufes der Welt. Das ist aber nicht sonderlich einsichtig, denn wie HEISENBERG (1955, S. 118) zu Recht bemerkt, kann der Mensch zwar tun, was er will, aber er kann nicht wollen, was er will.

für Indeterminiertheit. Für eine unbeschränkte Intelligenz ist daher der Ablauf der Ereignisse mit Sicherheit vorhersehbar[12].

In einer solchen Welt unterscheiden sich die beiden von Fisher gestellten Aufgaben gar nicht mehr prinzipiell. Auch wenn *vor* dem Münzwurf die Wahrscheinlichkeit für das Ereignis „Kopf" angegeben werden soll, ist das wirkliche Wurfergebnis bereits determiniert. Es hängt davon ab, wie der Werfende seine Hand bewegt und welche Luftströmungen während des Wurfes auf den Fall der Münze Einfluß nehmen, aber all dies kann nach festen Gesetzen aus zeitlich früher liegenden Ursachen erklärt werden. So gibt es keine wirkliche Indeterminiertheit, und die einzige Quelle der subjektiven Indeterminiertheit ist unsere eigene Dummheit.

Was ist nun aber die objektive Wahrscheinlichkeit, wenn alles vorherbestimmt ist? Nach dem Luce-Raiffa-Schlaifer-Konzept ist sie schlicht jener Wert, gegen den ganz im Sinne der durch von Mises (1936) gegebenen Wahrscheinlichkeitsdefinition die relative Häufigkeit einer Ergebnisausprägung konvergiert, wenn die Entscheidungssituation unter *nicht unterscheidbaren* Versuchsbedingungen ständig wiederholt wird.

So definiert gibt es für beide von Fisher gestellten Aufgaben, also vor allem auch für die zweite, eine objektive Wahrscheinlichkeit: Bei fortlaufender Wiederholung findet man etwa in der Hälfte der Fälle das Ergebnis „Kopf" vor. Spitzfindig könnte man zwar auch argumentieren, die Versuchswiederholung brauchte einfach nur darin zu bestehen, immer wieder von

[12] Laplace ist zu dieser Auffassung wohl auch unter dem Eindruck der Erfolge, die die makroskopische Physik im Bereich der Astronomie und Mechanik aufzuweisen hatte, gekommen. (Er führt sie jedenfalls zum Beleg seiner These an.) In Kenntnis der Heisenbergschen Unbestimmtheitsrelationen, die uns sagen, daß die für eine deterministische Beschreibung so unerläßlichen Begriffe wie Ort und Zeit im Bereich des Mikrokosmos gegenstandslos werden, sind wohl einige Zweifel am Laplaceschen Weltbild angebracht. Statt dessen scheint das schon bei Reichenbach (1925) begründete, doch von Hartwig (1956) so genannte *Ätialprinzip*, nach dem gleiche allgemeine Ursachen gleiche stochastische Verteilungsgesetze der Ergebnisse zur Folge haben, eine Stütze zu erhalten. Die Stochastik im Mikrobereich kommt zwar bei vielen Vorgängen auf Makroebene wegen der großen Zahl der beteiligten Moleküle häufig nicht zum Zuge. Doch die erhebliche Streuung, die man bei der Berechnung der Explosionskraft der Atombombe (vgl. Heisenberg (1954, S. 135)) in Kauf nehmen muß, und der Einfluß, der über stochastische Mutationen auf den Evolutionsprozeß ausgeübt wird (vgl. Monod (1971, S. 57 u. S. 141–150), wo genau die Gegenthese zum Prinzip des hinreichenden Grundes für die Biosphäre formuliert wird), liefern prägnante Beispiele für Einflüsse auf die Makrowelt. Dennoch weiß man nicht, ob die Mikrostochastik nicht letztlich auch bloß Ausdruck unserer Unkenntnis der wahren Vorgänge ist. So soll Einstein ja angesichts der philosophischen Interpretation der Unbestimmtheitsrelation einmal gesagt haben, er können sich nicht vorstellen, daß Gott würfele. Wie dem auch sei, wir bleiben in der deterministischen Welt, um klarzustellen, daß das Phänomen der Wahrscheinlichkeit nicht mit wahrer Indeterminiertheit erklärt werden *muß*.

neuem nachzuschauen, ob „Kopf" oder „Zahl" oben liegt, ohne die Münze jedes Mal von neuem zu werfen. Doch hätte man dann die Bedingung der nicht unterscheidbaren Versuchsbedingungen verletzt, da beim ersten Nachschauen die Lage der Münze unbekannt, in allen anderen Fällen jedoch bekannt wäre. Als realistisches Beispiel für diesen Fall könnte man an die Erdölexploration denken. Die objektive Wahrscheinlichkeit, in einem bestimmten Feld fündig zu werden, bestimmt sich durch die relative Erfolgsquote auf anderen, von ihren geologischen Daten her *a priori* als gleich zu beurteilenden Feldern und natürlich nicht durch den Anteil der Bohrungen, die in *diesem* Feld erfolgreich sind.

Dieses Beispiel zeigt auch, daß mit nicht unterscheidbaren Bedingungen nicht etwa identische Bedingungen gemeint sein können, denn identische Bedingungen liegen vor, wenn man immer an derselben Stelle bohrt. „Nicht unterscheidbar" heißt nur soviel wie „bei gleichen Vorinformationen". Daß diese Vorinformationen zwangsläufig beschränkt sind, ist der Grund dafür, daß eine Stochastik überhaupt erst ins Spiel kommt: Von Versuch zu Versuch variieren die unkontrollierten Einflüsse auf das Ergebnis in einer zwar deterministischen, jedoch nicht systematisch mit dem Ergebnis verbundenen Weise und produzieren so das, was wir Zufall nennen[13].

Überlegen wir einmal, was geschieht, wenn man einen Teil der bislang nicht kontrollierten Einflüsse bei der Versuchswiederholung mit berücksichtigt, so daß der Spielraum unsystematischer Einflüsse reduziert wird. In diesem Fall muß man wohl mit einem Einfluß auf die relative Häufigkeit eines interessierenden Ergebnisses rechnen. So könnte bei der Erdölsuche die Erweiterung der in Betracht gezogenen Einflußgrößen in der zusätzlichen Berücksichtigung eines seismologischen Tests bestehen. Bestimmt man die relative Erfolgsquote auf allen Feldern, die bezüglich der alten Informationen und des neu eingeführten Tests *a priori* als gleich zu bewerten sind, dann wird man in aller Regel einen anderen Wert als zuvor, ohne Berücksichtigung des Tests, finden. Als ein zweites Beispiel für diesen Effekt sei auf das Problem, die objektive Schadenswahrscheinlichkeit eines Versicherungsvertrages zu berechnen, angeführt. Damit dem Versicherungspraktiker die Haare zu Berge stehen, nehmen wir an, die Versicherungsunternehmen versicherten alle Kraftfahrzeuge zum Einheitstarif, weil sie die Mühe einer Kategorisierung der Kraftfahrzeuge scheuen. Die objektive Schadenswahr-

[13] Wer meint, daß Zufall wirkliche Indeterminiertheit voraussetzt, möge das Telefonbuch aufschlagen und nach einem beliebigen, aber festen System Namen herausschreiben und die Träger der Namen nach ihrer Körpergröße fragen. Obwohl von vornherein determiniert, müssen die von Befragung zu Befragung erfahrenen Zahlenwerte statistisch genauso als Zufallsvariable aufgefaßt werden, als hätte man die Namen mit Hilfe irgend eines mechanischen Zufallsgenerators, dessen Verhalten ja möglicherweise wirklich indeterminiert ist, ermittelt. Die systematischen Stichprobenauswahlverfahren der Statistik machen sich diesen Umstand zunutze.

scheinlichkeit eines einzelnen willkürlich herausgegriffenen Vertrages wäre in diesem Fall an der relativen Schadenshäufigkeit aller Verträge abzulesen, da der Gesamtbestand an Verträgen als die vielfache Wiederholung eines Zufallsexperiments bei gleichen Vorinformationen aufgefaßt werden kann[14]. Damit sich die Haare wieder legen, sei jetzt unterstellt, es werde eine Kategorisierung nach PS-Klassen eingeführt. In diesem Fall mißt die relative Schadenshäufigkeit in einer einzelnen Klasse (approximativ) die objektive Schadenswahrscheinlichkeit eines beliebigen Vertrages eben dieser Klasse. Nur im Spezialfall wird sich hier gegenüber der Situation ohne Kategorisierung keine Änderung der objektiven Schadenswahrscheinlichkeit ergeben.

Aus diesen Überlegungen folgt, daß es keine objektiven Wahrscheinlichkeiten „an sich" geben kann, sondern diese Wahrscheinlichkeiten nur in bezug auf irgendwelche Vorinformationen sinnvoll definiert werden können. Hier liegt die einzige Quelle individueller Einflüsse auf den Wert der objektiven Wahrscheinlichkeit. Insofern als zwei Personen unterschiedliche Informationen haben oder auch nur für relevant halten, gibt es für sie unterschiedliche objektive Wahrscheinlichkeiten. Gerade im Versicherungsfall ist keineswegs ausgeschlossen, daß vom Standpunkt des Versicherungsnehmers und des Versicherungsunternehmens die objektive Wahrscheinlichkeit unterschiedliche Werte hat.

Wie die Informationen Einfluß auf die Wahrscheinlichkeit gewinnen können, läßt sich leicht an Hand des Satzes von BAYES (1763, S. 381, Prop. 5) klarmachen[15], der sogleich auf das Versicherungsbeispiel bezogen werden soll. Wir nennen M die Menge der bei gegebenen Vorinformationen für möglich gehaltenen Zustände der Welt. Vom Standpunkt eines Versicherungsunternehmens, das sich für die Schadenswahrscheinlichkeit eines bestimmten Vertrages interessieren möge, unterscheiden wir die Zustände der Welt zweckmäßigerweise nach allen erdenklichen Kriterien, die für die Schadenswahrscheinlichkeit dieses Vertrages von Belang sein können[16], also z.B. danach, welcher PS-Klasse das versicherte Fahrzeug angehört, wieviel schadenfreie Jahre der Versicherungsnehmer nachweisen kann usw., aber vor allem danach, ob ein Schaden stattfinden wird oder nicht. $I \subset M$ sei die Teilmenge nur jener Zustände, die nach Erhalt einer bestimmten Information, also z.B. „der Vertrag bezieht sich auf ein Fahrzeug der Klasse

[14] Eine ausführlich Diskussion der Bedingungen, unter denen der Versicherungsfall als Zufallsexperiment interpretiert werden kann, findet man bei HELTEN (1973, S. 7–16).

[15] Für eine experimentelle Überprüfung des Problems, inwieweit die Menschen Wahrscheinlichkeiten in Übereinstimmung mit dem Bayesschen Theorem schätzen, inwieweit sie also in der Lage sind, objektive Wahrscheinlichkeiten korrekt zu ermitteln, siehe EDWARDS und PHILLIPS (1964).

[16] Für ein verwandtes Konzept zur konkreten Schätzung von Schadenswahrscheinlichkeiten siehe BAILEY und SIMON (1960) und HELTEN (1974).

90–110 PS mit 4 schadenfreien Jahren", noch möglich sind. $Z \subset M$ sei die Teilmenge der Zustände der Welt, die den Schadensfall kennzeichnen, und $\bar{Z}$ die Komplementmenge. Bekannt seien die zugehörigen *a priori*-Wahrscheinlichkeiten[17] $W(Z)$ (Anteil der Schäden am Gesamtbestand der Kraftverkehrsversicherung) und $W(\bar{Z}) = 1 - W(Z)$. Weiterhin sei die bedingte Wahrscheinlichkeit $W(I/Z)$ (Anteil der Schadensfälle mit den Vertragskennzeichen „90–110 PS, 4 schadenfreie Jahre" an der Gesamtzahl der Schadensfälle) und ebenso $W(I/\bar{Z})$ (Anteil der schadenfreien Verträge mit den Vertragskennzeichen „90–110 PS, 4 schadenfreie Jahre" an der Gesamtzahl der schadenfreien Verträge) bekannt. Dann folgt aus

$$(2) \qquad W(I \cap Z) = W(Z)\, W(I/Z) = W(I)\, W(Z/I)$$

die Bayes-Formel

$$(3) \qquad W(Z/I) = \frac{W(I \cap Z)}{W(I)} = \frac{W(Z)\, W(I/Z)}{W(I)} = \frac{W(Z)\, W(I/Z)}{W(Z)\, W(I/Z) + W(\bar{Z})\, W(I/\bar{Z})},$$

die uns sagt, wie die *a priori*-Wahrscheinlichkeit $W(Z)$ durch den Erhalt der zusätzlichen Informationen zu einer *a posteriori*-Wahrscheinlichkeit $W(Z/I)$ verändert wird. In der Abb. 1 sind diese Zusammenhänge dargestellt. Darüber hinaus wird dort die Wirkung einer weiteren Informationsvermehrung durch die zusätzliche Berücksichtigung anderer die Schadenswahrscheinlichkeit beeinflussender Kriterien wie „Höchstgeschwindigkeit", „Jahreskilometerleistung" u.ä. demonstriert. Sie besteht darin, daß die Menge der möglichen Umweltzustände über $I', I'', \ldots$ immer weiter eingeengt wird, bis schließlich irgendwann entweder

$$(4) \qquad I^{(n)} \subseteq Z, \quad \text{so daß} \quad W(Z/I) = 1\ ,$$

oder

$$(5) \qquad I^{(n)} \cap Z = \emptyset, \quad \text{so daß} \quad W(Z/I) = 0\ ,$$

gefunden wird, man sich also der Laplaceschen Intelligenz angenähert hat und weiß, ob ein bestimmter Vertrag einen Schaden bringen wird oder nicht.

[17] $W(.)$ heißt in dieser Arbeit durchweg „Wahrscheinlichkeit von (.)".

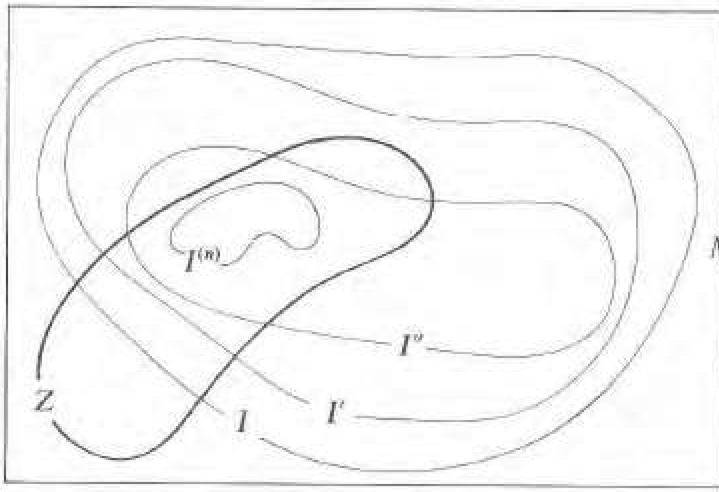


Abbildung 1

Natürlich stellen sich dem Versuch einer Informationsvermehrung praktisch sehr früh technische und ökonomische Schranken entgegen und verhindern die Beseitigung der Unsicherheit. Dennoch macht unser Gedankenspiel deutlich, daß sich eine objektive Schadenswahrscheinlichkeit immer nur für vorgegebene Klassifikationskriterien bestimmen läßt, also erst, nachdem der betrachtete Vertrag einer Kategorie nicht etwa identischer, sondern in bezug auf diese Kriterien *nicht unterscheidbarer* Verträge zugeordnet wird. *Die* einzig richtige objektive Wahrscheinlichkeit gibt es nicht. Immer kann man mit zusätzlichem Aufwand ein höheres Maß an Vorinformationen (eine feinere Klassifikation) anstreben, so daß sich die objektive Schadenswahrscheinlichkeit eines beliebig herausgegriffenen Vertrages ändert. Dies ist das schon von KNIGHT (1921, S. 217f. u. S. 224) bemerkte *Paradoxon der homogenen Gruppierung*. Es führt die Generationenfrage der Versicherungswissenschaft[18], ob auch „nicht-homogene" Verträge, nämlich solche mit unterschiedlicher „objektiver" Schadenswahrscheinlichkeit in der großen Zahl konsolidiert werden können, *ad absurdum*.

3. *Die Bildung äquivalenter objektiver Wahrscheinlichkeiten*

Wenn in der Kraftverkehrsversicherung ein Versicherungsnehmer seine Schadensklasse kennt, so weiß er noch lange nicht über seine Unfallwahr-

[18] Vgl. z.B. BRAESS (1960, bes. S. 40f.). Es sei nur nebenbei bemerkt, daß es bei unserem Verständnis objektiver Wahrscheinlichkeiten gar nicht verwunderlich ist, daß Lloyd's auch dann noch eine objektive Kalkulationsbasis findet, wenn es darum geht, die Beine von Marlene Dietrich zu versichern. Es werden zwar nicht mehrere Beinpaare versichert, doch gibt es wohl sehr viele andere Risiken, die in bezug auf die Beurteilungskategorien, die Lloyd's vermutlich entwickelt hat, als gleich anzusehen sind.

scheinlichkeit Bescheid. Der Besitz der zur Definition einer objektiven Wahrscheinlichkeit benötigten Information impliziert also keinesfalls automatisch die Kenntnis der objektiven Wahrscheinlichkeit. Bekannt kann eine objektive Wahrscheinlichkeit erst sein, wenn zusätzlich Informationen über die relative Häufigkeit des interessierenden Ereignisses bei unabhängigen oder korrekter: bei nicht systematisch verbundenen Risikosituationen vorliegen[19].

Nach dem Bekanntheitsgrad der Wahrscheinlichkeit lassen sich folgende Kennzeichen von Entscheidungssituationen unterscheiden:

– mit Sicherheit bekannte Wahrscheinlichkeiten – völlig bekannte Wahrscheinlichkeitshierarchien	} Risiko
– teilweise bekannte Wahrscheinlichkeitshierarchien – völlig unbekannte Wahrscheinlichkeiten	} Ungewißheit

Dabei ist mit dem Begriff der Wahrscheinlichkeitshierarchie gemeint, daß alternative Wahrscheinlichkeitsverteilungen über den Zuständen der Welt für möglich gehalten werden, hierfür unter Umständen selbst wieder alternative Wahrscheinlichkeitsangaben zu berücksichtigen sind usw. (Eine präzisere Beschreibung soll an geeigneter Stelle gegeben werden.) Unsere Aufgabe wird es sein, die drei letzten Fälle auf den ersten zurückzuführen[20]. Dabei wird sich zeigen, daß der zweite Fall mit dem ersten identisch ist. Beide belegen wir daher mit dem Begriff „Risiko". Um einen prinzipiellen Unterschied zu dokumentieren, kennzeichnen wir die beiden letzten Fälle mit dem Begriff „Ungewißheit"[21].

Es wäre angenehm, wenn für die Entscheidungsprobleme der Wirklichkeit mit Sicherheit bekannte objektive Wahrscheinlichkeiten vorlägen. Doch leider ist das äußerst selten der Fall. So findet man neben den Versicherungen und gewerblichen Lotterien kaum praktische Anwendungsfälle, und selbst die Erwähnung der Versicherungsgesellschaften in diesem Zusammenhang ist nicht unproblematisch. Wozu die nicht enden wollende Diskussion der Versicherungswissenschaft um die richtigen Modelle der Schadensverteilung, wenn die Verteilung mit Sicherheit bekannt ist[22, 23]?

---

[19] Weil „Unabhängigkeit" ein feststehender statistischer Begriff ist, behalten wir ihn bei, ohne allerdings die philosophische Frage, ob es Unabhängigkeit im eigentlichen Sinne gibt, präjudizieren zu wollen.

[20] Es sei in Erinnerung gerufen, daß wir dabei Rationalverhalten unterstellen. Die aus positiver Sicht zu konstatierenden Schwächen der Menschen beim Umgang mit Wahrscheinlichkeiten berücksichtigen wir hier allenfalls als Kontrast zum Rationalverhalten. Zu den nicht behandelten Schwächen siehe PHILLIPS (1970) und KAHNEMAN und TVERSKY (1973a u. b).

[21] In der Literatur wird bisweilen auch der Begriff „Unsicherheit" gebraucht. „Unsicher" möge hier aber schlicht „nicht sicher" heißen.

[22] Siehe HELTEN (1973).

[23] Vgl. aber den von SCHNEEWEISS (1967, S. 271 f.) unternommenen Versuch, die praktische Relevanz des Risikofalles herauszustellen. Neben den beiden genannten

Um Informationen über relative Häufigkeiten zu gewinnen, benötigt man allerdings nicht zwangsläufig empirische Erfahrungen. KRELLE (1957, S. 638) weist zu recht darauf hin, daß man bei vielen praktischen Problemen relative Häufigkeiten über Gedankenexperimente ermitteln kann[24]. Das ist ja auch das Grundkonzept des Luce-Raiffa-Schlaifer-Ansatzes[25], dem wir uns mit dem Bemühen um die Reduktion aller Entscheidungssituationen auf den Risikofall anschließen. Allein, in den seltensten Fällen kommen wir mit Hilfe von Gedankenexperimenten zu Wahrscheinlichkeiten, für die wir die Hand ins Feuer legen würden.

Daher ist die Analyse der drei letzten der oben unterschiedenen Fälle unerläßlich. Wir beginnen mit dem Extremfall völlig unbekannter Wahrscheinlichkeiten und kommen dann zur Stufentheorie der Wahrscheinlichkeit.

3.1. *Völlig unbekannte Wahrscheinlichkeiten*

Für diesen Fall sind in der unsicherheitstheoretischen Diskussion eine Reihe von Präferenzfunktionalen maßgeschneidert worden, die von vornherein selbst auf Surrogatwahrscheinlichkeiten verzichten. So das *Maximin-* (oder auch *Minimax-)Prinzip* von WALD (1945; 1950, S. 18) und VON NEUMANN und MORGENSTERN (1947, S. 101), der *Optimismus-Pessimismus-Index* von HURWICZ (zitiert nach MILNOR (1954, S. 50), der sich auf ein unveröffentlichtes Manuskript beruft) oder das *Minimax-Regret-Prinzip* von NIEHANS (1948) und SAVAGE (1951)[26]. Wir werden sehen, daß die bei der Konstruktion dieser Präferenzfunktionale freiwillig geübte Abstinenz von Wahrscheinlichkeiten unnötig ist. Es lassen sich nämlich äquivalente Surrogatwahrscheinlichkeiten finden.

Beispielen verweist Schneeweiß noch auf die im Bereich der Unternehmensforschung üblichen risikotheoretischen Modelle, scheint aber zu verkennen, daß diese Modelle von ihren Urhebern wohl vor allem der Not gehorchend und nicht dem eigenen Triebe für den Risikofall konstruiert wurden.

[24] Ähnliches scheint GEORGESCU-ROEGEN (1954) im Sinn zu haben, wenn er Risiko und Ungewißheit danach unterscheidet, ob der „Ausspielungsmodus" eines Ergebnisses bekannt ist oder nicht.

[25] Vgl. S. 14f.

[26] Ausgehend von der Ergebnismatrix wird beim *Maximin-Prinzip* als Präferenzfunktional das Zeilenminimum und beim *Optimismus-Pessimismus-Index* ein gewogenes Mittel aus Zeilenminimum und -maximum gewählt. Beim *Minimax-Regret-Prinzip* wird jedes Element der Ergebnismatrix zunächst durch die Differenz bis zum Spaltenmaximum ersetzt und mit -1 multipliziert, so dann wird das Zeilenmaximum der transformierten Matrix als Präferenzfunktional gewählt. Einen detaillierten Vergleich der Kriterien findet man bei MILNOR (1954) und LUCE und RAIFFA (1957, S. 275–297).

3.1.1. Das Ellsberg-Paradoxon

Betrachten wir dazu eine Lotterie von der Art wie sie von ELLSBERG (1961) konstruiert wurde, um gerade das Gegenteil unserer Behauptung zu belegen. Aus einer Urne mit weißen und schwarzen Kugeln wird eine Kugel blind gezogen. Will der Entscheidungsträger mitspielen, muß er einen Einsatz p zahlen und auf eine der beiden Farben setzen. Tippt er richtig, erhält er 100 DM, andernfalls nichts. Die Tab. 3 zeigt die zugehörige Ergebnismatrix.

Tabelle 3

Zustandsklassen der Welt / Handlungen	schwarz wird gezogen	weiß wird gezogen
es wird auf schwarz gesetzt	$100\,\text{DM} - p$	$-p$
es wird auf weiß gesetzt	$-p$	$100\,\text{DM} - p$
es wird auf die Teilnahme verzichtet	0 DM	0 DM

Es ist nun noch nichts darüber gesagt worden, in welchem Verhältnis die Urne schwarze und weiße Kugeln enthält. Hierzu unterscheiden wir zwei alternative Spiele:

(1) Der Entscheidungsträger erfährt, daß in der Urne $w_1 \cdot 100 = 50$ weiße und $w_2 \cdot 100 = 50$ schwarze Kugeln sind.
(2) Dem Entscheidungsträger werden die Anteile der beiden Farben nicht mitgeteilt.

Die Frage an den Entscheidungsträger lautet, welche Einsätze er äußerstenfalls für beide Spiele wagen würde.

Ein Charakteristikum der vom typischen Entscheidungsträger gegebenen Antwort ist, daß der für das erste Spiel genannte Maximaleinsatz jenen für das zweite Spiel übersteigt. Damit zeigt der Entscheidungsträger ein Verhalten, das nicht mit der Hypothese, er bilde sich beim zweiten, dem „Ungewißheitsspiel", subjektive Wahrscheinlichkeiten für das Auftreten der beiden möglichen Farben, vereinbar zu sein scheint. Würde er nämlich mit w_1 und w_2 solche subjektiven Wahrscheinlichkeiten benutzen, dann müßte er für das zweite Spiel den gleichen maximalen Einsatz zu zahlen bereit sein wie für das erste, wenn er vermutet, daß $w_1 = w_2$. Wenn er jedoch, weshalb auch immer, glaubt, daß $w_1 > w_2$ oder $w_2 > w_1$, dann müßte das Ungewißheitsspiel ihm sogar einen höheren Maximaleinsatz als das Risikospiel entlocken, denn durch die Wahl der richtigen Farbe könnte er sich aus seiner Sicht eine Gewinnchance von mehr als 50% verschaffen.

An der Beurteilung dieses mittlerweile unter dem Namen *Ellsberg-Paradoxon* bekannten Verhaltens scheiden sich die Geister. Krelle (1968, S. 178–184) akzeptiert es als Ausdruck einer besonderen Ungewißheitsaversion und trägt ihm durch die Einführung eines Informationsaxioms (S. 181) Rechnung. Doch Roberts (1963) glaubt[27], der Entscheidungsträger interpretiere die ihm angebotene Entscheidungssituation falsch, und Brewer (1963) und Schneeweiss (1968 b) halten es für möglich, daß das Verhalten des Entscheidungsträgers nur deshalb nicht mit der Existenz subjektiver Wahrscheinlichkeiten übereinzustimmen scheine, weil er in dem Experimentator einen rationalen Gegenspieler vermute. Es ist denkbar, daß diese Interpretationen richtig sind, doch vielleicht waren die von Ellsberg befragten Versuchspersonen, die um eine schnelle und intuitive Antwort gebeten wurden, einfach überfordert. Diese Vermutung drängt sich jedenfalls auf, wenn man einmal einer einfachen Modifikation des Ungewißheitsspiels folgt, wie sie von Raiffa (1961) vorgeschlagen wurde.

Man fragt den Entscheidungsträger, ob er für den Fall einer Teilnahme am zweiten Spiel eine bestimmte Farbe bevorzuge, und wenn er das, wie zu erwarten[28], verneint, empfiehlt man ihm ein drittes Spiel:

(3) Zunächst entscheidet ein Münzwurf darüber, welche Farbe gesetzt wird, und dann wird aus einer Urne mit unbekanntem Mischungsverhältnis eine Kugel gezogen. Alles andere bleibt gleich.

Wiederum um die Angabe seines Maximaleinsatzes gebeten, wird der typische Entscheidungsträger den gleichen Betrag wie beim Spiel (2) nennen. Das erscheint auch als sehr vernünftig. Denn wenn der Entscheidungsträger keine der beiden Farben bevorzugt, dann sollte es ihm einerlei sein, ob er die Farbe selbst wählt oder ob der Zufall entscheidet.

Dennoch hat sich der Entscheidungsträger aufs Glatteis führen lassen. Es läßt sich nämlich leicht zeigen, daß das dritte Spiel in gleicher Weise objektive Gewinn- und Verlustchancen von 50% aufweist wie das erste Spiel. Betrachten wir dazu die nachfolgende Ergebnismatrix, deren Kopf ein Baumdiagramm enthält, das zeigt, welche Zustandsklassen der Welt durch die Kombination von Münzwurf und Urnenzug entstehen.

[27] Siehe auch die Replik von Ellsberg (1963).
[28] Falls sich eine Farbenpräferenz zeigt, kann man die Kugeln ja anders anmalen.

Tabelle 4

Münzwurf entscheidet für	schwarz 1/2		weiß 1/2	
Urnenzug / Handlung	schwarz w_1^*	weiß w_2^*	schwarz w_1^*	weiß w_2^*
Spiel (3)	$100-p$	$-p$	$-p$	$100-p$
Status quo	0	0	0	0

Nennen wir die *unbekannten* Wahrscheinlichkeiten oder relativen Anteile der schwarzen bzw. weißen Kugeln w_1^* und w_2^*, dann kann mit

$$0{,}5\, w_1^* + 0{,}5\, w_2^* = 0{,}5\,(w_1^* + w_2^*) = 0{,}5$$

eine objektive Gewinnwahrscheinlichkeit und damit natürlich auch eine ebenso hohe Verlustwahrscheinlichkeit für das Spiel (3) errechnet werden, ohne daß dabei, und das ist erstaunlich, die objektiven Wahrscheinlichkeiten w_1^* und w_2^* bekannt zu sein brauchen.

Vergleicht man die für alle drei Spiele genannten Maximaleinsätze miteinander, dann zeigt sich, daß der Entscheidungsträger offenkundig einen Fehler gemacht hat[29]. Da Spiel (3) von den Gewinnchancen her gesehen mit Spiel (1) identisch ist, liegt dieser Fehler entweder darin, daß Spiel (1) und Spiel (2) für unterschiedlich, oder darin, daß Spiel (2) und Spiel (3) für gleich gehalten wurden. Welche dieser Alternativen zutrifft, läßt sich eindeutig ermitteln; jedenfalls dann, wenn man ein altbekanntes Axiom für rationale Entscheidungen bei Risiko akzeptiert.

3.1.2. Das Unabhängigkeitsaxiom

Raiffas Münzwurftrick soll deshalb näher untersucht werden. Die Analyse wird die Irrationalität einer Unterscheidung zwischen dem Risiko- und dem Ungewißheitsfall aufzeigen und eine im nächsten Abschnitt noch abzuleitende Regel vorbereiten, nach der man selbst bei völliger Unkenntnis irgendwelcher Wahrscheinlichkeiten zu äquivalenten objektiven Wahrscheinlichkeiten gelangt[30].

---

[29] Die von Raiffa befragten Personen haben dies nach Erläuterung eingesehen und ihre Bewertungen revidiert.

[30] Soweit bekannt, hat Raiffa eine präferenztheoretische Fundierung seines Tricks nicht versucht. Auch in seinem später erschienenen Lehrbuch (Raiffa (1968)) geht er über seinen Kommentar (1961) zu Ellsberg nicht hinaus.

Betrachten wir die Tab. 5. Sie entspricht der Tab. 4 nur mit dem Unterschied, daß zusätzlich das Spiel (2) mit den alternativen Möglichkeiten, auf „schwarz“ oder „weiß“ zu setzen, aufgeführt ist. Es mag zunächst irritieren, daß beim Spiel (2) anders als in der Tab. 3 Ergebnisvektoren mit vier Elementen auftauchen. Man kann sich aber leicht überzeugen, daß der Münzwurf irrelevant ist und es, auf welche Farbe man auch setzt, eine Gewinnmöglichkeit mit der (unbekannten) Wahrscheinlichkeit w_1^* und eine Verlustmöglichkeit mit der (ebenfalls unbekannten) Wahrscheinlichkeit w_2^* gibt.

Tabelle 5

		Münzwurf entscheidet für: schwarz 1/2		weiß 1/2	
Urnenzug / Handlung		schwarz w_1^*	weiß w_2^*	schwarz w_1^*	weiß w_2^*
Spiel (2)	schwarz wird gesetzt	$100-p$	$-p$	$100-p$	$-p$
	weiß wird gesetzt	$-p$	$100-p$	$-p$	$100-p$
Spiel (3)		$100-p$	$-p$	$-p$	$100-p$
Status quo		0	0	0	0

Man erkennt jetzt das Wesen des Raiffa-Tricks. Es besteht darin, die Gleichheit zwischen Spiel (2) und Spiel (3) durch eine Vertauschung der beiden letzten Elemente in Zeile 1 oder der beiden ersten Elemente in Zeile 2 zu suggerieren. Damit hat Raiffa implizit von dem in der Risikotheorie viel diskutierten[31] Unabhängigkeitsaxiom Gebrauch gemacht.

Dieses Axiom wurde bereits von Marschak (1950, S. 120–122, Postulat IV) vorgeschlagen, in einer von Samuelson (1952a, S. 147) etwas verschärften Version jedoch populär gemacht. Wir beziehen uns hier sofort auf die

[31] Im Zusammenhang mit dem unten (Kap. II C 2.) noch zu behandelnden *von Neumann-Morgenstern-Index*.

verschärfte Version[32]; sie wird an dieser Stelle zwar noch nicht benötigt, doch später werden wir sie gut gebrauchen können.

Axiom der starken Unabhängigkeit: *Gegeben seien drei Ergebnisvektoren, e_1, e_2, und e_3. (Im Spezialfall dürfen es auch Skalare sein.) Angenommen, es bestehe die Präferenz*

$$e_1 \{\gtrsim\} e_2,$$

dann gilt für mit beliebigem e_3 zusammengesetzte Vektoren

$$\begin{pmatrix} w & 1-w \\ e_1 & e_3 \end{pmatrix} \{\gtrsim\} \begin{pmatrix} w & 1-w \\ e_2 & e_3 \end{pmatrix},$$

solange $0<w\leq 1$, mit w als einer (bekannten) objektiven Wahrscheinlichkeit.

Bezieht man sich wegen der größeren Anschaulichkeit auf die Begriffswelt der Glücksspiele, dann kann man das Axiom auch so ausdrücken: Bei der Wahl zwischen zwei Lotterien, die mit der Wahrscheinlichkeit $(1-w)$ gleiche, jedoch mit der Wahrscheinlichkeit w unterschiedliche Gewinne versprechen, entscheidet man sich nur im Hinblick auf diese unterschiedlichen Gewinne, und zwar so wie man es auch bei unabhängiger Betrachtung täte.

Das Axiom entspricht Marschaks schwächerer Version, wenn nur auf das Indifferenzsymbol $\sim$ abgestellt wird, und erlaubt dann die hier benötigte Aussage, daß man in einer Lotterie jeden Gewinn durch einen in unabhängiger Betrachtung als äquivalent festgestellten Gewinn ersetzen darf, ohne an der Vorteilhaftigkeit der Lotterie im Vergleich zu dritten Lotterien etwas zu ändern. Man spricht daher auch vom Substitutionsaxiom[33].

Gerade diese Substitutionseigenschaft wurde bei der Elementvertauschung benutzt. Beachten wir dazu folgende Beziehungen:

$g_1 \equiv$ Glaubwürdigkeit für das Ziehen einer schwarzen Kugel
$g_2 \equiv$ Glaubwürdigkeit für das Ziehen einer weißen Kugel

---

[32] In der Fassung von SAMUELSON (1952a). Bei SAMUELSON (1952b) gibt es eine geringfügige Modifikation. In der Version

$$\begin{pmatrix} w & 1-w \\ e_1 & e_1 \end{pmatrix} \prec \begin{pmatrix} w & 1-w \\ e_2 & e_1 \end{pmatrix} \Leftrightarrow e_1 \prec e_2$$

entspricht das Axiom dem *Sure-Thing*-Axiom von SAVAGE (1954, S. 73).

[33] So ALLAIS (1953, S. 528).

$$\begin{pmatrix} 1/2 & 1/2 \\ e_1 & e_1 \end{pmatrix} = e_1 \equiv \begin{pmatrix} g_1 & g_2 \\ 100\,\mathrm{DM} - p & -p \end{pmatrix} \equiv \begin{cases} \text{Ergebnis des Spiels (2),} \\ \text{wenn auf „schwarz“} \\ \text{gesetzt wird.} \end{cases}$$

$$\begin{pmatrix} 1/2 & 1/2 \\ e_2 & e_2 \end{pmatrix} = e_2 \equiv \begin{pmatrix} g_1 & g_2 \\ -p & 100\,\mathrm{DM} - p \end{pmatrix} \equiv \begin{cases} \text{Ergebnis des Spiels (2),} \\ \text{wenn auf „weiß“} \\ \text{gesetzt wird.} \end{cases}$$

$$\begin{pmatrix} 1/2 & 1/2 \\ e_1 & e_2 \end{pmatrix} \equiv \text{Ergebnis des Spiels (3)}$$

Hieraus folgt unter Verwendung des Unabhängigkeitsaxioms:

$$(6) \qquad \begin{pmatrix} 1/2 & 1/2 \\ e_1 & e_1 \end{pmatrix} \sim \begin{pmatrix} 1/2 & 1/2 \\ e_1 & e_2 \end{pmatrix} \sim \begin{pmatrix} 1/2 & 1/2 \\ e_2 & e_2 \end{pmatrix} \Leftrightarrow e_1 \sim e_2.$$

Somit sind die Spiele (2) und (3) als gleich zu bewerten, wenn nur e_1 nicht schlechter als e_2 eingeschätzt wird und umgekehrt, also immer dann, wenn man beim Spiel (2) nicht weiß, auf welche Farbe man setzen soll. Und in Verbindung mit der oben schon gezeigten Identität der Spiele (1) und (3) folgt, daß es, wenn man das Unabhängigkeitsaxiom akzeptiert, falsch ist, das Ungewißheitsspiel (2) als unterschiedlich vom Risikospiel (1) anzusehen.

Daß man dann, wenn man unschlüssig über die zu setzende Farbe ist, nichts dagegen haben sollte, die Entscheidung per Münzwurf treffen zu lassen, leuchtet ein; insofern findet das Unabhängigkeitsaxiom sicher Zustimmung. Gleichwohl lassen sich andere Implikationen finden, die auf den ersten Blick nicht gerade plausibel erscheinen. Insbesondere ALLAIS (1952 u. 1953) hat dazu einiges zusammengetragen.

Man habe z.B. die Präferenz[34]

$$\text{Bauernschrank} \prec \text{Rokkokokommode}$$

Stellt man einen derben Tonkrug auf beide, ist dann automatisch

$$(\text{Bauernschrank, Tonkrug}) \prec (\text{Rokkokokommode, Tonkrug}),$$

wie man es nach dem Unabhängigkeitsaxiom vermuten könnte? Sicher nicht! Möglicherweise verkehrt sich die Präferenzrelation sogar in ihr Gegenteil. Ein erstes Befremden läßt sich aber zerstreuen: Mit dem Unabhängigkeitsaxiom werden nämlich keine Einrichtungsgegenstände, sondern, um im Bild zu bleiben, Chancen auf den Erhalt eines einzigen Gegenstandes

[34] Vgl. ALLAIS (1952, S. 316, Fn.). Eine ähnliche Kritik führt WOLD (1952) an; siehe auch die unmittelbar anschließend abgedruckte Diskussion mit Shackle und Savage und den Beitrag von SAMUELSON (1952b, S. 673f.).

nebeneinandergestellt. Der unangenehme Anblick des Tonkruges auf der Rokkokokommode bleibt in jedem Fall erspart.

Eine andere Kritik[35] bereitet aber schon etwas mehr Kopfzerbrechen. Fragen wir uns doch einmal, ob wir nicht der Wahlentscheidung

$$\text{(a)} \qquad \begin{pmatrix} 98\% & 2\% \\ 500 \text{ Mio DM} & 0 \text{ DM} \end{pmatrix} \prec \begin{pmatrix} 100\% \\ 100 \text{ Mio DM} \end{pmatrix}$$

und ebenso einer anderen Wahlentscheidung

$$\text{(b)} \qquad \begin{pmatrix} 0{,}98\% & 0{,}02\% & 99\% \\ 500 \text{ Mio DM} & 0 \text{ DM} & 1 \text{ DM} \end{pmatrix} \succ \begin{pmatrix} 1\% & 99\% \\ 100 \text{ Mio DM} & 1 \text{ DM} \end{pmatrix}$$

zustimmen könnten. Nicht jeder wird beide Wahlentscheidungen treffen, doch viele vernünftige Leute tun es[36]. Sie verstoßen damit gegen das Unabhängigkeitsaxiom, denn bei

$$e_1 \equiv \begin{pmatrix} 98\% & 2\% \\ 500 \text{ Mio DM} & 0 \text{ DM} \end{pmatrix}, \; e_2 \equiv \begin{pmatrix} 100\ \% \\ 100 \text{ Mio DM} \end{pmatrix} \;\text{ und }\; e_3 \equiv \begin{pmatrix} 100\% \\ 1 \text{ DM} \end{pmatrix}$$

wird trotz der im Fall (a) bekundeten Präferenz

$$e_1 \prec e_2$$

im Fall (b) so entschieden, daß

$$\begin{pmatrix} 1\% & 99\% \\ e_1 & e_3 \end{pmatrix} \succ \begin{pmatrix} 1\% & 99\% \\ e_2 & e_3 \end{pmatrix}.$$

Auch bei diesem Beispiel ist es allerdings ähnlich wie beim Ellsberg-Paradoxon. Durch eine etwas andere Präsentation[37] läßt sich die Inkonsistenz der Wahlentscheidung aufdecken. Man betrachte zunächst das folgende Problem:

---

[35] ALLAIS (1952, S. 316f.; 1953, S. 529f.).

[36] Der größere Teil einer Gruppe von ca. 20 vom Verfasser befragten Studenten, doch nicht alle. SAMUELSON (1952b, S. 678) braucht also nicht zu befürchten, daß er und Savage allein auf der ganzen Welt in der Lage seien, die von Allais gestellten Fragen konsistent zu beantworten.

[37] In der Art, wie sie MARKOWITZ (1970, S. 220–224) für ähnliche Beispiele vornimmt. Vgl. auch SAVAGE (1954, S. 103).

(c) $$\left(\begin{array}{cc}\overbrace{\begin{pmatrix}98\% & 2\% \\ 500\text{ Mio DM} & 0\text{ DM}\end{pmatrix} \prec \begin{pmatrix}100\% \\ 100\text{ Mio DM}\end{pmatrix}}^{1\%} & \begin{array}{c}99\% \\ 1\text{ DM}\end{array}\end{array}\right)$$

Die Schreibweise soll andeuten, daß der Entscheidungsträger hier, ob er will oder nicht, an einer Ausspielung teilnehmen muß, die mit 99% Wahrscheinlichkeit 1 DM erbringt, jedoch mit 1% Wahrscheinlichkeit die Wahlmöglichkeit (a) eröffnet. Dabei wird angenommen, daß jetzt auch tatsächlich dieselbe Wahl wie im Fall (a) getroffen wird. Wenn man durch das Eintreten der 1%-Chance nicht wagemutig geworden ist („Ich habe eine Glückssträhne!"), dann ist das wohl sinnvoll[38]. Nun soll (c) etwas modifiziert werden, indem man den Entscheidungsträger bittet, bereits *vor* der ersten Ausspielung festzulegen, welche Wahl er nach Eintreten des 1%-Ereignisses treffen würde. Es wird wohl kaum jemanden außer ALLAIS (1952, S. 313–330; 1953, S. 538) geben, der jetzt zu einer anderen Entscheidung kommt. Man bekundet also die Präferenz

(d) $$\left(\begin{array}{cc}1\% & 99\% \\ \begin{pmatrix}98\% & 2\% \\ 500\text{ Mio DM} & 0\text{ DM}\end{pmatrix} & 1\text{ DM}\end{array}\right) \prec \left(\begin{array}{cc}1\% & 99\% \\ 100\text{ Mio DM} & 1\text{ DM}\end{array}\right).$$

Rechnet man die Wahrscheinlichkeit für das Erreichen der einzelnen Gewinne nach der Multiplikationsregel für unabhängige Ereignisse aus, stellt man fest, daß in (d) die gleichen Alternativen wie in (b) zur Wahl stehen. Die Tatsache, daß dort eine andere Wahlentscheidung getroffen wurde als hier, deckt die Inkonsistenz auf. Man wäre also gut beraten gewesen, hätte man dem Unabhängigkeitsaxiom vertraut und sich erst einmal über seine eigenen Präferenzen Klarheit verschafft.

Man könnte mit ALLAIS (1953, S. 540) gegen diesen Vorwurf einwenden, daß die Probleme (b) und (d) zu Unrecht als gleich bezeichnet werden, da sich doch immerhin der Ausspielungsmodus unterscheide. Das ist aber die oben schon erwähnte Kritik am Ordnungsaxiom, die wir für nüchterne wirtschaftliche Entscheidungen im Gegensatz zur Spielsituation nicht zulassen können[39]. Im übrigen wird die im Beispiel aufgedeckte Inkonsistenz wohl tatsächlich nicht durch Spielfreunde erklärt, denn unter allen vom Verfasser dazu befragten Personen fand sich nicht eine einzige, die die Wahlprobleme (b) und (d) für unterschiedlich hielt[40]. Die wahre Erklärung

---

38 An die „Glückssträhne" zu glauben bedeutet, die Wahrscheinlichkeiten der zweiten fälschlicherweise vom Ergebnis der ersten Ausspielung abhängig zu machen, und ist als irrational abzutun.

39 Vgl. S. 11 f.

40 Vgl. Fn. 36.

der Inkonsistenz liegt darin, daß in der Präsentationsweise (b) die Bedeutung kleiner Wahrscheinlichkeiten etwas verdeckt wird, während sie in (d) klar hervortritt.

Das Resümee zu Allais' Kritik ist damit das gleiche wie zu der Ellsbergs. Die Menschen sind in Ungewißheitssituationen mitunter überfordert. Sie machen Fehler und verhalten sich damit nicht wie ihr idealisierter Verwandter in dieser Untersuchung. Aber immerhin, der Verwandte ist ihr Idol. Das ist seine Daseinsberechtigung.

3.1.3. Eine Rehabilitation des Prinzips des unzureichenden Grundes

Ausgerüstet mit dem Unabhängigkeitsaxiom sind wir jetzt in der Lage, den Münzwurftrick zu verallgemeinern. Das Ergebnis dieser Verallgemeinerung besteht in dem berühmten *Prinzip des unzureichenden Grundes*, das auf Jacob BERNOULLI (1713, S. 88f.) und LAPLACE (1814, S. IV u. VII) zurückgeht[41]. Angewandt auf unser Problem muß nach diesem Prinzip einer jeden der alternativ möglichen Zustandsklassen der Welt die gleiche äquivalente Wahrscheinlichkeit zugeordnet werden, wenn man keine Anhaltspunkte dafür hat, das Auftreten einer Zustandsklasse für glaubwürdiger als das einer anderen zu halten.

Ein anschauliches Beispiel vom Wert des Prinzips des unzureichenden Grundes liefert das oben mehrfach verwendete Modell einer Urne mit *bekanntem* Inhalt. Da wir bei völlig gleichmäßig gefertigten Kugeln keinen Grund sehen, warum eine Kugel eher als eine andere gezogen werden sollte, halten wir ihr Erscheinen für gleich wahrscheinlich und schließen folgerichtig vom relativen Anteil der Kugeln in einer Farbe auf die Wahrscheinlichkeit des Erscheinens dieser Farbe. Daß wir damit richtig kalkuliert haben, läßt sich durch Ermittlung des relativen Anteils der Farbe bei wiederholten Ziehungen experimentell überprüfen.

Liegt eine andere Entscheidungssituation vor, wenn wir wissen, daß Kugeln bestimmter Farbe vertreten sind, jedoch das Mischungsverhältnis nicht kennen? Der Unterschied ist wohl nicht grundsätzlicher Natur, denn man kann sich vorstellen, daß man beim Ziehen aus einer unbegrenzten Menge unabhängig voneinander gefüllter Urnen, die je ausschließlich schwarze und weiße Kugeln enthalten, auf die Dauer auch jede Farbe zur Hälfte in seinem Säckchen vertreten findet. Allein, manch einer empfindet doch, daß es einen Unterschied gibt. Wir wollen daher sehen, ob wir nicht auf andere Weise zeigen können, daß es für ein konsistentes Verhalten erforderlich ist, im Falle völliger subjektiver Unkenntnis irgendwelcher Wahrscheinlichkeitsvorstellungen so zu tun, als ob es für die Zustandsklassen der Entscheidungswelt einander gleiche mit Sicherheit bekannte objektive Wahrscheinlichkeiten gäbe.

[41] KEYNES (1921, S. 41f.) hat es *Indifferenzprinzip* genannt. Vgl. zu den folgenden Ausführungen auch SINN (1980).

Wir beziehen uns auf das folgende Entscheidungsproblem: Es gibt die Zustandsklassen der Welt $Z_1, Z_2, \ldots, Z_n$, deren objektive Wahrscheinlichkeiten $w_1^*, w_2^*, \ldots, w_n^*, \sum_{j=1}^n w_j^* = 1$, allesamt völlig unbekannt sind. Es möge kein Grund erkennbar sein, warum das Auftreten einer dieser Zustandsklassen für plausibler als das einer anderen gehalten werden sollte. Der Möglichkeitsbereich des Entscheidungsträgers besteht aus den Handlungen $a_1, a_2, \ldots, a_m$ mit den Ergebnisvektoren $e_1, e_2, \ldots, e_m$,

$$(7) \qquad e_i = \begin{pmatrix} w_1^* & w_2^* \ldots & w_n^* \\ e_{i1} & e_{i2} \ldots & e_{in} \end{pmatrix} \qquad \forall i = 1, 2, \ldots, m.$$

Es liegt also ein Entscheidungsproblem vor, wie es sich durch die mit Tab. 1 dargestellte Ergebnismatrix veranschaulichen läßt. Gezeigt werden soll, daß

$$(8) \qquad \begin{pmatrix} w_1^* & w_2^* \ldots & w_n^* \\ e_{i1} & e_{i2} \ldots & e_{in} \end{pmatrix} \sim \begin{pmatrix} 1/n & 1/n \ldots & 1/n \\ e_{i1} & e_{i2} \ldots & e_{in} \end{pmatrix} \qquad \forall i = 1, 2, \ldots, m.$$

Wir führen zunächst einen Zufallsgenerator mit den Zuständen $Z'_1, Z'_2, \ldots, Z'_n$ ein, die mit den objektiven Wahrscheinlichkeiten $w'_1, w'_2, \ldots, w'_n$ auftreten, und unterscheiden in einer Kunstwelt die Zustandsklassen nach den Zuständen des Generators und den Zustandsklassen der wahren Welt, so daß es dort n^2 verschiedene Zustandsklassen gibt. Auch in der Kunstwelt läßt sich das oben formulierte Entscheidungsproblem darstellen. In der Tab. 6 ist dies in der ersten Zeile nur einmal repräsentativ an Hand der Handlung a_i mit dem Ergebnisvektor e_i geschehen. Unabhängig von den Zuständen des Generators wird unter der wahren Zustandsklasse der Welt Z_j das Ergebnis e_{ij} erzielt.

Tabelle 6

Zufalls-generator	w'_1 Z'_1	w'_2 Z'_2	w'_3 Z'_3		w'_n Z'_n
wahre Welt	$w_1^* \quad w_n^*$ $Z_1 \ldots Z_n$	$w_1^* \, w_2^* \, w_n^*$ $Z_1 \quad Z_2 \ldots Z_n$	$w_1^* \, w_2^* \, w_3^* \quad w_n^*$ $Z_1 \quad Z_2 \quad Z_3 \ldots Z_n$	$\ldots$	$w_1^* \quad w_{n-1}^* \quad w_n^*$ $Z_1 \ldots Z_{n-1} \quad Z_n$
a_i (1)	$e_{i1} \ldots e_{in}$	$e_{i1} \quad e_{i2} \ldots e_{in}$	$e_{i1} \quad e_{i2} \quad e_{i3} \ldots e_{in}$	$\ldots$	$e_{i1} \ldots e_{i\,n-1} \quad e_{in}$
(2)	$e_{i1} \ldots e_{in}$	$e_{in} \quad e_{i1} \ldots e_{i\,n-1}$	$e_{i1} \quad e_{i2} \quad e_{i3} \ldots e_{in}$	$\ldots$	$e_{i1} \ldots e_{i\,n-1} \quad e_{in}$
(3)	$e_{i1} \ldots e_{in}$	$e_{in} \quad e_{i1} \ldots e_{i\,n-1}$	$e_{i\,n-1} \quad e_{in} \quad e_{i1} \ldots e_{i\,n-2}$	$\ldots$	$e_{i1} \ldots e_{i\,n-1} \quad e_{in}$
	$\vdots \quad \vdots$	$\vdots \quad \vdots \quad \vdots$	$\vdots \quad \vdots \quad \vdots \quad \vdots$		$\vdots \quad \vdots \quad \vdots$
a_i (n)	$e_{i1} \ldots e_{in}$	$e_{in} \quad e_{i1} \ldots e_{i\,n-1}$	$e_{i\,n-1} \quad e_{in} \quad e_{i1} \ldots e_{i\,n-2}$	$\ldots$	$e_{i2} \ldots e_{in} \quad e_{i1}$

Bevor wir zu den anderen Zeilen kommen, wollen wir noch $n-1$ neue Ergebnisvektoren

$$(9)\qquad \begin{aligned} e_i^2 &\equiv \begin{pmatrix} w_1^* & w_2^* & \dots & w_n^* \\ e_{in} & e_{i1} & \dots & e_{in-1} \end{pmatrix} \\ e_i^3 &\equiv \begin{pmatrix} w_1^* & w_2^* & \dots & w_n^* \\ e_{in-1} & e_{in} & \dots & e_{in-2} \end{pmatrix} \\ &\vdots \\ e_i^n &\equiv \begin{pmatrix} w_1^* & w_2^* & \dots & w_n^* \\ e_{i2} & e_{i3} & \dots & e_{i1} \end{pmatrix} \end{aligned}$$

definieren, die aus dem ursprünglichen Vektor

$$e_i \equiv e_i^1 = \begin{pmatrix} w_1^* & w_2^* & \dots & w_n^* \\ e_{i1} & e_{i2} & \dots & e_{in} \end{pmatrix}$$

dadurch entstehen, daß die Einzelergebnisse Schritt um Schritt nach rechts verschoben werden und das ganz rechts herausfallende Einzelergebnis links wieder angesetzt wird. Wir nehmen zur Kenntnis, daß man beim paarweisen Vergleich zweier beliebiger Ergebnisse e_i^j und e_i^k durchweg zu dem Urteil, e_i^j sei nicht schlechter als e_i^k, allerdings e_i^k auch nicht schlechter als e_i^j, somit also zu dem Urteil

$$(10)\qquad e_i^k \sim e_i^j \qquad \forall j, k = 1, 2, \dots, n,$$

kommt. Wie anders sollte denn auch die Aussage, man habe in das Eintreten eines Zustandes aus der einen Klasse nicht weniger Vertrauen als in das Eintreten eines Zustandes aus der anderen Klasse, gedeutet werden können, wenn nicht so, daß es einem gleichgültig erscheint, unter welcher Klasse eine bestimmte Ergebnisausprägung auftritt? (10) ist also geradezu als eine notwendige Bedingung dafür anzusehen, daß man nicht die leiseste Ahnung von der Größe der Wahrscheinlichkeiten der Zustandsklassen der Welt hat.

Es sollen nun die anderen Zeilen der Tab. 6 schrittweise bis zur letzten, der n-ten Zeile, die wir einer Handlung a_i^* zuordnen, entwickelt werden. Wir beginnen, indem wir den Ergebnisvektor, der durch die erste Zeile verkörpert wird, als

$$(11)\qquad e_i^1 = \begin{bmatrix} w_2' & & 1-w_2' & \\ & \left(\dfrac{w_1'}{1-w_2'} \right. & \cdot \;\; \dfrac{w_3'}{1-w_2'} \;\; \dots & \left. \dfrac{w_n'}{1-w_2'}\right) \\ e_i^1 & \left(e_i^1 \right. & \cdot \;\; e_i^1 \;\; \dots & \left. e_i^1 \right) \end{bmatrix}$$

schreiben. Das hat den Vorteil, daß wir unmittelbar auf die Formulierung des Unabhängigkeitsaxioms zurückgreifen und das erste Element (unter w'_2) durch e_i^2 aus (9) ersetzen können. Die Rücktransformation von (11) bringt dann die zweite Zeile der Tab. 6. Diese zweite Zeile kann nun ihrerseits als

$$\text{(12)} \qquad \begin{bmatrix} w'_3 & 1-w'_3 \\ & \begin{pmatrix} \dfrac{w'_1}{1-w'_3} & \dfrac{w'_2}{1-w'_3} & \cdot & \dfrac{w'_4}{1-w'_3} & \cdots & \dfrac{w'_n}{1-w'_3} \\ e_i^1 & e_i^2 & \cdot & e_i^1 & \cdots & e_i^1 \end{pmatrix} \\ e_i^1 & \end{bmatrix}$$

geschrieben werden, was ein leichtes Ersetzen des ersten Elements durch e_i^3 aus (9) ermöglicht. Die abermalige Rückwandlung bringt die Zeile 3 der Matrix. In der beschriebenen Weise fahren wir nun fort und substituieren nacheinander $e_i^4, e_i^5, \ldots, e_i^n$ aus (9). Das Resultat ist die Zeile n der Tab. 6, die das Ergebnis der Handlung a_i^* verkörpert. Durch die schrittweise Umwandlungsprozedur wurde sichergestellt, daß sie genauso bewertet wird wie die Ausgangszeile 1.

Die in der Zeile n auftauchenden Ergebniselemente sind die gleichen wie in der Zeile 1. Für die Wahrscheinlichkeiten, mit denen diese Elemente in der Zeile n auftauchen, erhält man:

$$\text{(13)} \qquad \begin{aligned} W(e_{i1}) &= w_1^* w'_1 + w_2^* w'_2 + \ldots + w_{n-1}^* w'_{n-1} + w_n^* w'_n, \\ W(e_{i2}) &= w_2^* w'_1 + w_3^* w'_2 + \ldots + w_n^* w'_{n-1} + w_1^* w'_n, \\ &\vdots \\ W(e_{in}) &= w_n^* w'_1 + w_1^* w'_2 + \ldots + w_{n-2}^* w'_{n-1} + w_{n-1}^* w'_n. \end{aligned}$$

Man beachte, daß wir bislang noch keinerlei Annahmen über die Größe der Wahrscheinlichkeiten $w'_1, w'_2, \ldots, w'_n$, mit denen der Zufallsgenerator seine Zustände $Z'_1, Z'_2, \ldots, Z'_n$ annimmt, benötigt haben. Die Äquivalenz der auseinander entwickelten Zeilen der Tab. 6 gilt in jedem Fall. So sind wir frei, beispielsweise

$$w'_1 = w'_2 = \ldots = w'_n = 1/n$$

zu setzen, was den Vorteil hat, daß wir in (13) diese Wahrscheinlichkeiten ausklammern können. Das Resultat lautet:

$$\text{(14)} \qquad \begin{aligned} W(e_{i1}) &= 1/n\,(w_1^* + w_2^* + \ldots + w_n^*) = 1/n, \\ W(e_{i2}) &= 1/n\,(w_1^* + w_2^* + \ldots + w_n^*) = 1/n, \\ &\vdots \\ W(e_{in}) &= 1/n\,(w_1^* + w_2^* + \ldots + w_n^*) = 1/n. \end{aligned}$$

Da eine völlig analoge Argumentation auch für alle anderen in (7) beschriebenen Ergebnisvektoren $e_1, e_2, \ldots, e_{i-1}, e_{i+1}, \ldots, e_m$ des ursprünglichen Entscheidungsproblems vollzogen werden kann, ist die in (8) behauptete Äquivalenz bewiesen.

Damit kommen wir zu der folgenden Aussage, die das Prinzip des unzureichenden Grundes rehabilitiert: Bei völlig unbekannten Wahrscheinlichkeiten für die Zustandsklassen der Welt muß der Entscheidungsträger Ergebnisvektoren so bewerten,

(1) als trete jede Zustandsklasse mit der gleichen Wahrscheinlichkeit auf und

(2) als sei diese Wahrscheinlichkeit eine mit Sicherheit bekannte objektive Größe.

Mit anderen, von LUCE und RAIFFA (1957, S. 286–298, bes. S. 291, 296) allerdings nicht als überzeugend empfundenen Axiomensystemen haben auch CHERNOFF (1954) und MILNOR (1954) das Prinzip des unzureichenden Grundes zu fundieren versucht[42]. Diese Ansätze haben in den Axiomen und der Beweistechnik mit dem unsrigen wenig gemein. Hervorzuheben ist, daß beide Ansätze die Annahme benötigen, daß die Elemente der Ergebnismatrix bereits in Form eines von Neumann-Morgenstern-Nutzens vorliegen[43]. Zwar werden auch wir später zu diesem Nutzenkonzept kommen, doch bedarf es dazu eines weiteren Axioms, das von seiten der lexikographischen Präferenztheorie bestritten wird und deshalb nicht unnötigerweise ins Spiel kommen sollte[44].

3.1.4. Äquivalente Wahrscheinlichkeiten für Fallstudien

Bei in der Praxis vorkommenden Entscheidungsproblemen ist es häufig angebracht, Fallstudien vorzunehmen, bei denen „Fälle“, „Unterfälle“, „Unterfälle von Unterfällen“ usw. unterschieden werden. Wie es in der Abb. 2 dargestellt ist, kommt man dann mit Hilfe eines Baumdiagramms zu den Zustandsklassen der Welt.

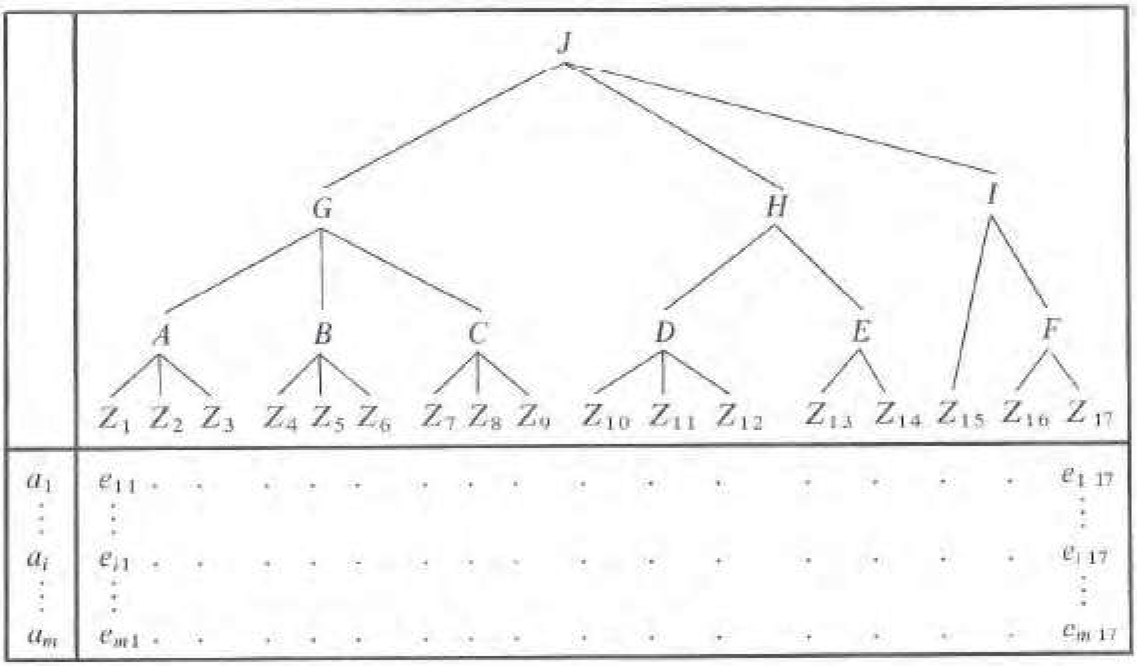
---

[42] Chernoffs Ansatz wird in dem Buch von CHERNOFF und MOSES (1959) nicht wieder aufgegriffen.

[43] Siehe CHERNOFF (1954, S. 422f.) und MILNOR (1954, S. 49). Die Nutzenwerte sind nicht etwa einfache Maßzahlen, die heterogene Ergebnisse gleichnamig machen sollen wie etwa bei KRELLE (1968, S. 122; vgl. auch S. 144f.), so daß noch eine echte von Neumann-Morgenstern-Nutzenfunktion darauf anzuwenden ist, sondern tatsächlich Werte dieser Funktion selbst. Das zeigt sich z. B. darin, daß MILNOR (S. 57) annimmt, daß die Addition einer Konstanten zu jedem Element einer Spalte der Ergebnismatrix die Präferenzordnung über die Zeilen unverändert läßt und daß in Chernoffs Definitionen 7 und 8 und seinem Postulat 8 die mathematische Nutzenerwartung verwendet wird, was beim „Maßzahlnutzen“ nur im Falle der Risikoneutralität möglich wäre.

[44] Gemeint ist das archimedische Axiom. Vgl. Kap. II C 2.1.

Abbildung 2



Wenn man hier nichts über irgendwelche Wahrscheinlichkeiten weiß, ist es natürlich nicht sinnvoll, die Zustandsklassen der Welt $Z_1, Z_2, \ldots, Z_{17}$ unbesehen für gleichwahrscheinlich zu halten. Statt dessen sollte man sich an jeder Astgabelung Klarheit darüber verschaffen, wie plausibel die einzelnen aus ihr entspringenden Äste unter der Bedingung sind, daß man an den Gabeln bereits angekommen ist. Wenn man dann an keiner der Astgabeln sagen kann, ob ein bestimmter aus ihr entspringender Ast mit größerem Vertrauen erwartet wird als ein beliebiger anderer, dann liegt es nahe, die Wahrscheinlichkeitssumme von 1 gleichmäßig auf alle aus einer Gabelung entspringenden Äste zu verteilen. Nach diesem Verfahren würde man den Zustandsklassen der Welt für das Beispiel der Abb. 2 die folgenden Wahrscheinlichkeiten zuweisen:

Nr. der Zustandsklasse	1	2	...	9	10	11	12	13	14	15	16	17
Wahrscheinlichkeit	$\frac{1}{27}$	$\frac{1}{27}$	...	$\frac{1}{27}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$

Nach dem Multiplikationssatz für Wahrscheinlichkeiten würde man also die Wahrscheinlichkeit einer bestimmten Zustandsklasse ermitteln, indem man das Produkt der Wahrscheinlichkeiten aller Äste bildet, über die man gehen muß, um vom Stamm zu dem Astende zu gelangen, das gerade diese Zustandsklasse bezeichnet.

Das Interessante an diesem Ergebnis wäre, daß man mit Hilfe des Prinzips des unzureichenden Grundes nicht nur zu Gleichverteilungen, sondern auch zu anders gestalteten Verteilungen käme. Der kritische Punkt dieser Überlegung ist allerdings, ob durch unsere Axiomatik abgedeckt wird, daß alle an einer bestimmten Gabel entspringenden Äste die gleiche Wahrscheinlichkeit erhalten müssen. Daß dies tatsächlich so ist, soll andeutungsweise gezeigt werden.

Zur Erleichterung der Darstellung beziehen wir uns auf die Abb. 3 und unterstellen, daß eine ganz bestimmte Handlung a_i gewählt wird. In diesem Fall gibt es unter einer jeden Gabelung einen Zufallsvektor von Einzelergebnissen, der bildlich gesprochen durch den gesamten „Strauch" unter dieser Gabelung repräsentiert wird. Wir benennen diesen Zufallsvektor mit dem Buchstaben, der die Gabelung kennzeichnet. Es sei noch vereinbart, daß die Äste, von denen die Rede ist, nie über eine Gabelung hinausragen.

Wir beginnen mit der Betrachtung der Gabelung A und dem zugehörigen Zufallsvektor A, bestehend aus den Elementen e_{i1}, e_{i2} und e_{i3}. In unmittelbarer Anwendung des im vorigen Abschnitt erzielten Ergebnisses können wir den Ästen unter A einander gleiche objektive Wahrscheinlichkeiten zuordnen, ohne an der Bewertung des Vektors A etwas zu ändern. Analog verfahren wir mit den Gabelungen B bis F. Die auf diese Weise mit objektiven Wahrscheinlichkeiten versehenen Ergebnisvektoren nennen wir $A', B', \ldots, F'$. Ohne schon zu diesem Zeitpunkt die veränderten Vektoren in das Baumdiagramm zu integrieren, gehen wir jetzt zu Gabelung G über und fassen den Vektor G als aus den Elementen A, B und C bestehend auf. Damit sind wir wieder bei einem Problem von der Art, wie es im vorigen Abschnitt betrachtet wurde, denn dort haben wir ja keine inhaltlichen Einschränkungen über das, was unter dem Einzelergebnis e_{ij} zu verstehen war, gemacht. Somit kommen wir zu gleichen Wahrscheinlichkeiten (je 1/3) für die Äste unter G. Analog verhält es sich mit den Ästen unter H und I. Der nächste Schritt besteht nun darin, daß wir unter Verwendung des Unabhängigkeitsaxioms im Vektor G die Elemente A, B und C schrittweise durch A', B' und C' ersetzen:

$$\begin{aligned}
&\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ & & \\ A & B & C \end{bmatrix} \sim \begin{bmatrix} 1/3 & 2/3 \\ & \begin{pmatrix} 1/2 & 1/2 \\ B & C \end{pmatrix} \\ A & \end{bmatrix} \sim \begin{bmatrix} 1/3 & 2/3 \\ & \begin{pmatrix} 1/2 & 1/2 \\ B & C \end{pmatrix} \\ A' & \end{bmatrix} \\
(15) \quad &\sim \begin{bmatrix} 1/3 & 2/3 \\ & \begin{pmatrix} 1/2 & 1/2 \\ A' & C \end{pmatrix} \\ B & \end{bmatrix} \sim \begin{bmatrix} 1/3 & 2/3 \\ & \begin{pmatrix} 1/2 & 1/2 \\ A' & C \end{pmatrix} \\ B' & \end{bmatrix} \sim \begin{bmatrix} 1/3 & 2/3 \\ & \begin{pmatrix} 1/2 & 1/2 \\ A' & B' \end{pmatrix} \\ C & \end{bmatrix} \\
&\sim \begin{bmatrix} 1/3 & 2/3 \\ & \begin{pmatrix} 1/2 & 1/2 \\ A' & B' \end{pmatrix} \\ C' & \end{bmatrix} \sim \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ & & \\ A' & B' & C' \end{bmatrix} \equiv G'.
\end{aligned}$$

Der transformierte Vektor heißt G'. Analog wird aus H der Vektor H' und aus I der Vektor I'. In dieser Weise fährt man nun fort. Man ordnet zunächst den Ästen unter J gleiche Wahrscheinlichkeiten zu und ersetzt dann G, H und I durch G', H' und I'. Damit ist die oben aufgestellte Behauptung, daß das Prinzip des unzureichenden Grundes auch bei Fallstudien in Form von Baumdiagrammen gilt, für das Beispiel der Abb. 3 nachgewiesen. Wir sparen uns die Schreibarbeit einer Verallgemeinerung auf beliebig komplexe Baumdiagramme und stellen fest:

Wenn in Fallstudien über die Zustandsklassen der Welt kein Unterfall glaubwürdiger als ein anderer erscheint, dann ist jedem Unterfall eine äquivalente objektive Wahrscheinlichkeit von der Höhe des Kehrwertes der Anzahl der Unterfälle zuzuordnen. Die Wahrscheinlichkeit einer Zustandsklasse ist dann das Produkt der (bedingten) Wahrscheinlichkeiten aller Fälle und Unterfälle, die unterschieden werden müssen, um die Klasse zu definieren.

3.1.5. Kritik am Prinzip des unzureichenden Grundes

Unsere Ergebnisse sind weit davon entfernt, als allgemein akzeptiert gelten zu können. So bestreitet Krelle (1961, S. 99 u. 106; 1968, S. 180f. u. S. 189f.) insbesondere den im vorletzten Abschnitt unter Punkt (2) angeführten Teilaspekt mit Nachdruck. Selbst wenn im Falle völlig unbekannter Wahrscheinlichkeiten gleiche subjektive Wahrscheinlichkeiten zu verwenden seien, so meint er, sei es gleichwohl rational, wenn man bei seinen Entscheidungen eine besondere Ungewißheitsaversion walten lasse. Doch da Krelle (1968, S. 133) mit seinem „Reduktions"- und seinem „Substitutionsaxiom" ebenfalls das (schwache) Unabhängigkeitsaxiom unterstellt, kann aus seiner Warte an unseren Überlegungen nur noch das Urteil (10) bestritten werden. Das aber wird schwerlich gelingen, es sei denn, man würde das bei Krelle (1968, S. 123–125) auch benötigte Ordnungsaxiom und damit unseren ganzen Ansatz verwerfen.

Eine andere Kritik wird häufig gegen das *klassische* Prinzip des unzureichenden Grundes vorgebracht, und wir müssen prüfen, inwieweit sie auch unser Ergebnis trifft. Eine Münze werde zweimal geworfen. Wie groß ist die Wahrscheinlichkeit, beide Male „Zahl" zu erhalten? Unterscheidet man die Ereignisse „Zahl, Zahl" und „nicht: Zahl, Zahl", dann ist die gesuchte Wahrscheinlichkeit 1/2. Unterteilt man die möglichen Fälle in „Zahl, Zahl", „Zahl, Wappen", „Wappen, Zahl" und „Wappen, Wappen", dann ist die Wahrscheinlichkeit 1/4, also gibt es einen Widerspruch. Die richtige Lösung liegt hier auf der Hand. Doch wenn man einmal fragt, wie groß beim zweimaligen Werfen die Wahrscheinlichkeit ist, mindestens einmal „Wappen" zu erhalten, dann kann man leichter in die Irre geführt werden. So argumentiert d'Alembert[45], falls beim ersten Wurf „Wappen" erscheine,

[45] Nach Todhunter (1865, S. 258f., Art. 464) zitiert aus d'Alembert, Croix ou Pile, Encyclopédie ou Dictionnaire Raisonné … 1754.

werde ein zweiter Wurf gar nicht mehr durchgeführt, so daß nur die Fälle „Wappen", „Zahl, Wappen" und „Zahl, Zahl" zu unterscheiden seien. Die gesuchte Wahrscheinlichkeit sei folglich 2/3, statt, wie es richtig ist, 3/4.

Diese Beispiele führen uns zu dem bereits bei VON KRIES (1886, bes. S. 1–23) mit aller Klarheit diskutierten Problem der Klasseneinteilung für die Zustände der Welt. Offenbar benötigt man, um nach dem Prinzip des unzureichenden Grundes objektive Wahrscheinlichkeiten auszurechnen, auch eine „richtige" Klasseneinteilung. Aus der Sicht der klassischen Wahrscheinlichkeitstheorie war das natürlich ein wichtiges, leider nie befriedigend gelöstes Problem. Unsere Überlegungen berührt es indes nur am Rande, denn wir wollen keine objektiven, sondern subjektive Wahrscheinlichkeiten bestimmen, wenn auch letztere in Form äquivalenter objektiver Wahrscheinlichkeiten vorliegen sollen. Um die Sache ganz klar zu legen, *wenn* d'Alembert wirklich keinen Grund sieht, einen seiner drei Fälle für plausibler als die anderen zu halten, dann soll er konsequenterweise bei beliebigen, mit dem Münzwurf verbundenen Entscheidungsproblemen davon ausgehen, daß bei einer fortwährenden Wiederholung des Doppelwurfs in 2/3 aller Fälle „Wappen" erzielt wird.

Ob es freilich sinnvoll ist, alle drei Fälle als gleich plausibel anzusehen, ist eine weitere Frage. Hätte d'Alembert das folgende Baumdiagramm betrachtet, dann hätte er festgestellt, daß kein Ast plausibler als ein anderer ist, und hätte folglich die Wahrscheinlichkeiten 1/2, 1/4 und 1/4 errechnet.

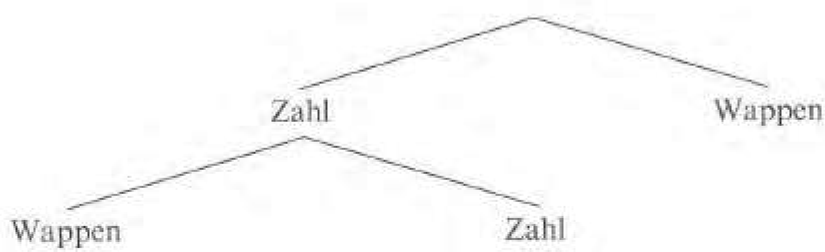


Abbildung 3

Ein eng mit d'Alemberts Fehler verwandtes Problem ist von SAVAGE (1954, S. 65) formuliert worden. Ein Entscheidungsträger kennt mehrere mögliche Aufteilungen von Ereignissen in Zustandsklassen, weiß aber nicht, welches die richtige Aufteilung ist. In diesem Fall scheint das Prinzip des unzureichenden Grundes zu versagen, weil für ein spezielles Ereignis verschiedene Wahrscheinlichkeiten kalkuliert werden können. Betrachten wir das von Savage genannte Beispiel: Zwei Kugeln werden aus einer Urne gezogen, von der bekannt ist, daß sie entweder zwei weiße Kugeln, zwei schwarze Kugeln, oder eine schwarze und eine weiße Kugel enthält. Wenn wir diese drei Möglichkeiten als Zustandsklassen der Welt auffassen, dann ist z. B. die Wahrscheinlichkeit, eine weiße und eine schwarze Kugel zu ziehen, 1/3. Savage hält es jedoch auch für möglich, die Zustandsklassen

„weiß/weiß“, „schwarz/schwarz“, „schwarz/weiß“ und „weiß/schwarz“ zu unterscheiden, so daß die gesuchte Wahrscheinlichkeit 1/2 ist. Glücklicherweise läßt sich das Problem lösen. Wenn Savage wirklich keinen Grund weiß, warum er die eine Zustandsklassenbeschreibung für plausibler als die andere halten soll, so mag er das folgende Baumdiagramm zu Rate ziehen und die Wahrscheinlichkeiten nach der im vorigen Abschnitt angegebenen Regel ermitteln:

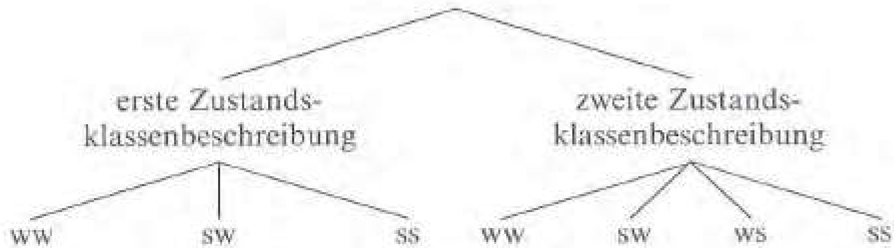


Abbildung 4

Als Resultat erhält er dann die äquivalente objektive Wahrscheinlichkeit von 5/12 für das Ziehen einer weißen und einer schwarzen Kugel.

3.2. *Partiell bekannte Wahrscheinlichkeiten: Die Stufentheorie der Wahrscheinlichkeiten*

Bislang haben wir nur die Fälle mit Sicherheit bekannter und völlig unbekannter Wahrscheinlichkeiten betrachtet. Die Wirklichkeit liegt aber zwischen den Extremen. So wird man in einige Ausprägungen des Ergebnisvektors häufig mehr Vertrauen als in andere setzen, sich aber seines Urteils keineswegs sicher fühlen, wenn man gebeten wird, äquivalente objektive Wahrscheinlichkeiten zu benennen. Man wird statt dessen alternative Wahrscheinlichkeiten mit vielleicht unterschiedlichen Vertrauensgraden für möglich halten und, befragt, ob man diese Vertrauensgrade in sichere Wahrscheinlichkeiten überführen könne, erneut mit den Achseln zucken usw.

Das adäquate Modell für solcherart vielschichtige Unsicherheit liefert die von REICHENBACH (1935, S. 305–322) entwickelte Stufentheorie der Wahrscheinlichkeit. Wir wollen die Grundidee benutzen, um aus den unvermeidlich vagen Glaubwürdigkeitsangaben des Entscheidungsträgers äquivalente objektive Wahrscheinlichkeiten zu berechnen. Im Rahmen der Präferenztheorie stellen unsere Ausführungen eine Verallgemeinerung der Ansätze von TINTNER (1941) und KRELLE (1968, S. 176) sowie einer Notiz von ROBERTS (1963, S. 329, Fn. 5) dar[46]. Unsere Verallgemeinerung zeigt sich in

[46] Parallelen gibt es allerdings auch zu einem Ansatz von SCHNEEWEISS (1964) und allen Ansätzen zur *Bayes-Statistik*, die *a priori*-Verteilungen von Parametern anderer Verteilungen und damit Wahrscheinlichkeiten der 2. Stufe verwenden. Vgl. z. B. HELTEN (1971).

zweierlei Weise. Zum einen werden im Sinne Reichenbachs vielstufige Wahrscheinlichkeiten behandelt. Zum anderen wird zugelassen, daß für die Wahrscheinlichkeiten auf irgendeiner Stufe zwar noch alternative Ausprägungen für möglich gehalten werden, doch hierfür keine Wahrscheinlichkeiten höherer Stufe bekannt sind. Es ist bemerkenswert, daß die amerikanische Schule der Subjektivisten, ob man nun an Savage oder Luce, Raiffa und Schlaifer denkt, diesen Problemen aus dem Wege geht, indem sie unterstellt, subjektive Wahrscheinlichkeiten könnten durch Befragungen, das Anbieten von Wetten und Vergleichsspielen und ähnliches mehr[47] in einem Zuge ermittelt werden.

Zur Erfassung der mehrstufigen Unsicherheit definieren wir die folgenden Größen:

$i_j, j>1$	lfd. Nr. der Ausprägung der Wahrscheinlichkeitsfunktion der Stufe $j-1$
i_1	lfd. Nr. der Zustandsklasse der Welt
$W^j(i_j), j>1$	(sichere) Wahrscheinlichkeit dafür, daß die Wahrscheinlichkeit der Stufe $j-1$ die Ausprägung mit der lfd. Nr. i_j annimmt
$W^1(i_1)$	(sichere) Wahrscheinlichkeit dafür, daß die Zustandsklasse Z_{i_1} auftritt
$W^j_{i_{j+1}}(i_j), j\geq 1$	i_{j+1}-te Ausprägung dieser Wahrscheinlichkeiten, falls sie selbst Zufallsvariablen sind
$\hat{W}^1(Z_{i_1})$	implizit vorhandene (sichere) Wahrscheinlichkeit für das Auftreten der Zustandsklasse Z_{i_1}
$\hat{W}^1_k(Z_{i_1})$	k-te Ausprägung dieser Wahrscheinlichkeit, falls sie eine Zufallsvariable ist
$\tilde{W}(Z_{i_1})$	äquivalente objektive Wahrscheinlichkeit für das Auftreten der Zustandsklasse Z_{i_1}
*	kennzeichnet unbekannte Wahrscheinlichkeiten

Damit läßt sich bereits die Struktur von Wahrscheinlichkeitshierarchien beschreiben. Das soll zunächst unter der Annahme geschehen, daß keine völlig unbekannten Wahrscheinlichkeiten auftauchen.

3.2.1. Völlig bekannte Wahrscheinlichkeitshierarchien

Wahrscheinlichkeiten 1. Stufe

Sind Wahrscheinlichkeiten der ersten Stufe bekannt, dann gibt es eine feste Funktion $W^1(i_1)$, $\sum_{i_1} W^1(i_1)=1$, die den Zustandsklassen der Welt Z_1, Z_2, ... eine Wahrscheinlichkeit zuordnet. Es liegt also der unproblematische Risikofall vor.

---

[47] Vgl. WINKLER (1967a u. b).

Wahrscheinlichkeiten 2. Stufe

Nun wird die Konstanz der Funktion $W^1(.)$ aufgehoben. Es sind mehrere mit $W_1^1(.)$, $W_2^1(.)$, ... bezeichnete Verläufe oder Ausprägungen möglich, die ihrerseits mit den Wahrscheinlichkeiten $W^2(1)$, $W^2(2)$, ..., wobei $\sum_{i_2} W^2(i_2)=1$, auftreten. Damit sind wir bei dem bereits in der Literatur untersuchten Fall der bekannten Wahrscheinlichkeiten 2. Stufe, den TINTNER (1941) wohl etwas zu euphorisch mit Ungewißheit (uncertainty) schlechthin identifiziert. Man beachte, daß die Ausprägungen $W_1^1(.)$, $W_2^1(.)$, ... ganze Wahrscheinlichkeitsverteilungen über den Zustandsklassen der Welt umfassen sollen und nicht etwa für jede einzelne Klasse gesondert definiert werden. Diese Konstruktion schließt nicht aus, daß für einzelne Klassen feste Wahrscheinlichkeiten 1. Stufe bekannt sind, denn in diesem Fall brauchten die Funktionen $W_1^1(.)$, $W_2^1(.)$, ... ja nur über den betreffenden Klassen die gleichen Werte anzunehmen.

Aus den vorliegenden Angaben kann bereits eine implizit vorhandene Wahrscheinlichkeit erster Stufe für jede beliebige Zustandsklasse der Welt berechnet werden, die gleichzeitig die gesuchte äquivalente objektive Wahrscheinlichkeit ist. Da die Wahrscheinlichkeit für das gemeinsame Auftreten der Zustandsklasse Z_{i_1} und der i_2-ten Ausprägung der Wahrscheinlichkeitsverteilung 1. Stufe nach der Multiplikationsregel für Wahrscheinlichkeiten[48] durch $W_{i_2}^1(i_1)\,W^2(i_2)$ angegeben wird, kommt man nach Summation[49] über alle Ausprägungen der Wahrscheinlichkeitsverteilung 1. Stufe zu dem Ergebnis

$$\tilde{W}(Z_{i_1})=\hat{W}^1(Z_{i_1})=\sum_{i_2} W^2(i_2)\,W_{i_2}^1(i_1)\,. \tag{16}$$

Wahrscheinlichkeiten höherer Stufen

Natürlich braucht auch $W^2(.)$ nicht als feste Funktion vorzuliegen. Ein Unsicherheitsproblem 3. Stufe liegt vor, wenn diese Funktion die alternativen Ausprägungen $W_1^2(.)$, $W_2^2(.)$, ... annehmen kann, denen vermöge einer weiteren Funktion die Wahrscheinlichkeiten $W^3(1)$, $W^3(2)$, ... zugewiesen werden. Da wir feststellen, daß die Wahrscheinlichkeit für das gemeinsame Auftreten der i_1-ten Zustandsklasse, i_2-ten Ausprägung der Wahrscheinlichkeitsfunktion 1. Stufe und i_3-ten Ausprägung der Wahrscheinlichkeitsfunktion 2. Stufe in diesem Fall durch $W_{i_2}^1(i_1)\;W_{i_3}^2(i_2)\;W^3(i_3)$ angegeben wird, kommen wir analog zu (16) zu dem Ausdruck

$$\tilde{W}(Z_{i_1})=\hat{W}^1(Z_{i_1})=\sum_{i_3}\sum_{i_2} W^3(i_3)\,W_{i_3}^2(i_2)\,W_{i_2}^1(i_1), \tag{17}$$

---

[48] $W(A\cap B)=W(A)\,W(B/A)$ lautet hier $W[W_{i_2}^1(.)\cap Z_{i_1}]=W^2(i_2)\,W_{i_2}^1(i_1)$.

[49] Nach der Additionsregel $W(A\cup B)=W(A)+W(B)-W(A\cap B)$ und unter Beachtung der Tatsache, daß die Ausprägungen der Wahrscheinlichkeitsverteilung 1. Stufe disjunkt sind.

wenn über alle Ausprägungen der Wahrscheinlichkeitsverteilungen 1. und 2. Stufe summiert wird.

Eine weitere Verallgemeinerung liegt auf der Hand: Sind erst auf der Stufe j objektive Wahrscheinlichkeiten sicher bekannt, ansonsten jedoch nur Wahrscheinlichkeitsverteilungen über die möglichen Gestalten, die die Wahrscheinlichkeitsverteilung der jeweiligen Vorstufe annehmen kann, dann läßt sich mit

$$(18)\qquad \begin{aligned} \hat{W}(Z_{i_1}) &= \hat{W}^1(Z_{i_1}) \\ &= \sum_{i_j}\sum_{i_{j-1}}\cdots\sum_{i_3}\sum_{i_2} W^j(i_j)\, W_{i_j}^{j-1}(i_{j-1})\ldots \\ &\quad \ldots W_{i_4}^3(i_3)\, W_{i_3}^2(i_2)\, W_{i_2}^1(i_1) \end{aligned}$$

eine implizit vorhandene objektive Wahrscheinlichkeit 1. Stufe, die auch automatisch die gesuchte äquivalente objektive Wahrscheinlichkeit ist, jederzeit errechnen.

Kritik

Den Gleichungen (16)–(18) ist gemein, daß mit ihnen die (bei (17) und (18) mehrstufige) Wahrscheinlichkeitsverteilung für die Wahrscheinlichkeit der ersten Stufe durch eine sichere Wahrscheinlichkeit der ersten Stufe ersetzt wird, die gerade so hoch wie der Erwartungswert der Verteilung ist:

$$(19)\qquad \hat{W}(Z_{i_1}) = \hat{W}^1(Z_{i_1}) = E\left[W_{i_2}^1(Z_{i_1})\right].$$

Hat es einen Sinn, diesen Erwartungswert als äquivalente objektive Wahrscheinlichkeit anzusehen? Muß nicht ein Spielraum für subjektive Risikoeinstellungen gelassen werden, etwa in der Weise, daß, falls Z_{i_1} einen erwünschten Zustand bezeichnet, der Optimist $\hat{W}^1(Z_{i_1})$ höher, der Pessimist es jedoch niedriger einschätzt als es in (18) geschieht? Diese Fragen bringen uns wieder in die Nähe des Ellsberg-Paradoxons, denn wenn man sich vorstellt, daß mit Hilfe von (18) bei $j \to \infty$ und geeignet gewählten Streubereichen der Wahrscheinlichkeitsverteilungen der Ungewißheitsfall simuliert wird, dann scheint in dem von Ellsberg beobachteten Verhalten die Bestätigung für eine pessimistische Wahrscheinlichkeitsbewertung zu liegen.

Es bleibt aber dabei, daß ein solches Verhalten bei sorgfältiger Überlegung als unvernünftig bezeichnet werden muß. Dazu mache man sich klar, daß ein Beispiel für die mehrstufige Unsicherheit durch das folgende Urnenmodell gegeben wird: Aus einer Urne, die 50 weiße und 50 schwarze Kugeln enthält, wird eine Zufallsstichprobe im Umfang n_1 gezogen, aus dieser Stichprobe dann eine Unterstichprobe im Umfang n_2, aus dieser eine weitere Unterstichprobe im Umfang n_3 usw. bis schließlich aus einer letzten Unterstichprobe genau eine Kugel gezogen wird. Dabei gelte $n_1 > n_2 > \ldots > 0$.

Macht es für die Glaubwürdigkeit, eine weiße Kugel zu erhalten, einen Unterschied, ob aus der Urne von vornherein eine Kugel gegriffen wird oder ob die Kugel über die beschriebene komplizierte Stichprobenhierarchie ermittelt wird? Sieht man von einer spezifischen Spielvorliebe ab, wie anfangs vereinbart[50], dann ist nicht einzusehen, wo der Unterschied liegen sollte. Wenn man objektive Wahrscheinlichkeiten als Grenzwerte relativer Häufigkeiten auffaßt, gibt es jedenfalls keinen. Insofern liegt bei Entscheidungsproblemen mit mehrstufigen, letztlich aber bekannten Wahrscheinlichkeiten der reine Risikofall vor.

Diese Lösung ist sicher sehr attraktiv, denn sie liefert ein Argument für die Beschränkung der Analyse auf den Risikofall. Doch leider ist das Argument nicht besonders schlagkräftig, denn anders als es ja TINTNER (1941) bereits für die Benutzung von Wahrscheinlichkeiten 2. Stufe behauptet hat, trägt die Lösung der besonderen Problematik der Ungewißheitssituation nur unzureichend Rechnung. Das führt uns zum folgenden Abschnitt.

3.2.2. Teilweise bekannte Wahrscheinlichkeitshierarchien

Mit der Aufgabe, Wahrscheinlichkeiten höherer Stufe zu benennen, wird ein Entscheidungsträger vermutlich schnell überfordert. Es ist z.B. denkbar, daß verschiedene Ausprägungen der Wahrscheinlichkeiten 2. Stufe für möglich gehalten werden, jedoch zwischen den Ausprägungen keine Diskriminierung nach der Glaubwürdigkeit mehr möglich ist.

In solchen Fällen kann wieder das *Prinzip des unzureichenden Grundes* zu Rate gezogen werden: Hält man keine der Ausprägungen für glaubwürdiger als andere, dann muß so getan werden, als seien für alle Ausprägungen gleiche objektive Wahrscheinlichkeiten bekannt.

Das soll für den allgemeinen Fall gezeigt werden. Sagen wir, Wahrscheinlichkeiten der Stufe $j+1$ seien nicht mehr bekannt, obwohl die Wahrscheinlichkeitsfunktion der Stufe j noch r alternative Ausprägungen $W_1^j(.)$, $W_2^j(.)$, ..., $W_r^j(.)$ annehmen kann. Nach (18) hat dann die implizierte Wahrscheinlichkeit 1. Stufe genausoviele alternative Ausprägungen $\hat{W}_1^1(Z_{i_1})$, $\hat{W}_2^1(Z_{i_1})$, ..., $\hat{W}_r^1(Z_{i_1})$, deren Wert durch

$$(20) \qquad \hat{W}_k^1(Z_{i_1}) = W^j(k) \sum_{i_{j-1}} \sum_{i_{j-2}} \dots \sum_{i_3} \sum_{i_2} W_{i_j}^{j-1}(i_{j-1})\, W_{i_{j-1}}^{j-2}(i_{j-2}) \dots W_{i_4}^3(i_3)\, W_{i_3}^2(i_2)\, W_{i_2}^1(i_1), k = 1, 2, \dots, r,$$

angegeben wird. So entpuppt sich das Problem der Unkenntnis von Wahrscheinlichkeiten auf der Stufe $j+1$ als ein Problem unbekannter Wahrscheinlichkeiten 2. Stufe. Es sei nun unterstellt, daß eine bestimmte Handlung a_i gewählt wird. Dann läßt sich der Ergebnisvektor als

[50] Vgl. S. 12.

(21) $$e_i = \begin{pmatrix} W^{*j+1}(1) & W^{*j+1}(2) & \dots & W^{*j+1}(r) \\ {}_1e_i & {}_2e_i & \dots & {}_re_i \end{pmatrix}$$

mit

$${}_ke_i \equiv \begin{pmatrix} \hat{W}_k^1(Z_1) & \hat{W}_k^1(Z_2) & \dots & \hat{W}_k^1(Z_n) \\ e_{i1} & e_{i2} & \dots & e_{in} \end{pmatrix}, \quad k=1,2,\dots,r,$$

schreiben, wobei die mit einem Stern bezeichneten Wahrscheinlichkeiten der Stufe $j+1$ völlig unbekannt sind. Da diese Formulierung jener der Beziehung (8) analog ist, können wir jetzt unmittelbar auf die oben gegebene Fundierung für das Prinzip des unzureichenden Grundes zurückgreifen und

(22) $$W^{*j+1}(k) = \frac{1}{r} \qquad \forall k=1,2,\dots,r,$$

setzen. Da wir dieses Ergebnis statt für die Handlung a_i auch für jede beliebige andere Handlung erhalten, ist das Problem der Ungewißheit bei Wahrscheinlichkeiten höherer Stufe in ein äquivalentes Risikoproblem verwandelt worden: Immer dann, wenn die Wahrscheinlichkeitshierarchie abbricht, weil die Wahrscheinlichkeiten ab einer bestimmten Stufe völlig unbekannt sind, müssen gleiche Wahrscheinlichkeiten für die Ausprägungen der Wahrscheinlichkeiten der nächstniedrigeren Stufe angenommen werden.

3.3. *Ergebnis*

Das Grundproblem der Entscheidungstheorie für die unsichere Welt liegt darin, die verschiedenen Typen der Unsicherheit auf einen gemeinsamen Nenner zu bringen. Ein extremer Typ ist der mit Sicherheit bekannter objektiver Wahrscheinlichkeiten und ein anderer der völlig unbekannter objektiver Wahrscheinlichkeiten für die möglichen Handlungsergebnisse. Wir haben gezeigt, wie sich alle Typen auf den der sicher bekannten objektiven Wahrscheinlichkeit zurückführen lassen, so daß er als Basis für die weitere Analyse dienen kann.

Für den Extremfall völlig unbekannter objektiver Wahrscheinlichkeiten kommt man zu einer einfachen Regel, die das Prinzip des unzureichenden Grundes rehabilitiert: Hat man keinerlei Vorstellung von der Größe der objektiven Wahrscheinlichkeiten für die alternativ möglichen Handlungsergebnisse, dann ist zum einen allen Alternativen die gleiche Wahrscheinlichkeit zuzuordnen und zum anderen die Handlungsbewertung so vorzunehmen, als handele es sich bei diesen Wahrscheinlichkeiten um mit Sicherheit bekannte objektive Werte. Verantwortlich für dieses Resultat sind nur zwei keinesfalls neue, sondern weithin akzeptierte Axiome: Das Ordnungs- und das Unabhängigkeitsaxiom.

In der Praxis werden bei der Analyse der möglichen Handlungsergebnisse unter Unsicherheit häufig Fallstudien durchgeführt. Das Prinzip des unzureichenden Grundes ist dann so anzuwenden, daß allen gleichgeordneten Unterfällen die gleichen Wahrscheinlichkeiten zugeordnet werden. Auf diese Weise entsteht eine Wahrscheinlichkeitsverteilung über den möglichen Zustandsklassen der Welt, die keinesfalls mit einer Gleichverteilung übereinstimmen muß.

Weder die völlige Unkenntnis objektiver Wahrscheinlichkeiten noch ihre sichere Kenntnis sind in der Praxis vorherrschende Erscheinungen. Unterstellt man realistischerweise, daß der Entscheidungsträger zwar zu Wahrscheinlichkeitsangaben in der Lage ist, doch sich dieser Angaben keineswegs sicher ist, kommen Wahrscheinlichkeiten höherer Stufen ins Spiel. Dabei sind zwei Fälle zu unterscheiden:

1. Sind Wahrscheinlichkeitsverteilungen von Wahrscheinlichkeitsverteilungen von ... über den Zustandsklassen der Welt bekannt, dann läßt sich eine implizit vorhandene objektive Wahrscheinlichkeit erster Stufe errechnen. Es liegt ein Entscheidungsproblem bei Risiko vor.
2. Nimmt die Wahrscheinlichkeitsverteilung einer beliebigen Stufe unterschiedliche Ausprägungen an, für die keinerlei Wahrscheinlichkeiten höherer Stufe mehr bekannt sind, dann muß für diese Ausprägungen eine mit Sicherheit bekannte Gleichverteilung unterstellt werden. Auf diese Weise wird ein dem ursprünglichen Ungewißheitsproblem äquivalentes Entscheidungsproblem unter Risiko formuliert.