Choquet-Erwartungsnutzen
und
antizipierter Nutzen

Ein Beitrag zur Entscheidungstheorie
bei einem und mehreren Attributen

DISSERTATION

zur Erlangung des akademischen Grades eines Doktors der Wirtschafts- und Sozialwissenschaften des Fachbereiches Wirtschafts- und Organisationswissenschaften der Universität der Bundeswehr Hamburg.

vorgelegt von

Rainer Dyckerhoff

aus Büdingen

Hamburg 1994

Referent:	Prof. Dr. Karl Mosler
Korreferent:	Prof. Dr. Wolf Krumbholz
Tag der mündlichen Prüfung:	6. Januar 1994

Gedruckt mit Unterstützung der Universität der Bundeswehr Hamburg

Inhaltsverzeichnis

Herrn Professor Dr. K. Mosler danke ich herzlichst für die zahlreichen Anregungen und Hinweise, die entscheidend zu dieser Arbeit beigetragen haben, und für die engagierte Betreuung während meiner Tätigkeit als wissenschaftlicher Mitarbeiter an der Universität der Bundeswehr Hamburg.

Einleitung

Die vorliegende Arbeit befaßt sich mit zwei neueren Modellen der Entscheidungstheorie, dem *Choquet-Erwartungsnutzen* und dem *antizipierten Nutzen*.

Die normative Entscheidungstheorie beschäftigt sich mit der Beschreibung rationalen Entscheidungsverhaltens. Dabei geht man von sogenannten Axiomen aus, d.h. möglichst einfachen und plausiblen Annahmen, die als Grundlage rationalen Handelns angesehen werden können. Es wird untersucht, wie ein Individuum, das gewisse Axiome für sich akzeptiert, sich in komplexen Entscheidungssituationen verhalten sollte. Unter einer Axiomatisierung eines bestimmten entscheidungstheoretischen Modells versteht man das Aufstellen einer Reihe von Axiomen derart, daß die Akzeptanz dieser Axiome zu einem Entscheidungsverhalten gemäß dem betrachteten Modell führt.

Zu Beginn der Entwicklung der Wahrscheinlichkeitstheorie im 17. und frühen 18. Jahrhundert war das gängige Modell das sogenannte Erwartungswert-Prinzip, welches besagt, daß Spiele oder Lotterien nach dem Erwartungswert des Gewinns beurteilt werden sollten. Das sogenannte St. Petersburger Paradoxon zeigte jedoch, daß das Entscheidungsverhalten von Individuen durch die Maximierung des Gewinnerwartungswertes nicht adäquat beschrieben werden kann. Beim St. Petersburger Paradoxon wird das folgende Spiel betrachtet. Eine faire Münze wird so lange geworfen, bis zum ersten Mal Zahl fällt. Wenn Zahl das erste Mal beim n-ten Wurf auftritt, erhält der Spieler 2^n DM. Für welchen Preis ist eine Person bereit, an diesem Spiel teilzunehmen? Man kann zeigen, daß der Erwartungswert des Gewinns unabhängig vom Einsatz unendlich ist. Sicher ist jedoch niemand bereit, für die Teilnahme an einem solchen Spiel einen sehr hohen Betrag wie z.B. 1000.- DM zu bezahlen. Eine Lösung des St. Petersburger Paradoxons wurde von Daniel Bernoulli (1738) vorgeschlagen. Bernoulli erklärte das St. Petersburger Paradoxon mit der Annahme, daß die entscheidenden Faktoren für die Bewertung einer Lotterie nicht die Auszahlungen selbst seien, sondern der subjektiv empfundene Wert oder Nutzen, den die Person von einer Auszahlung hat. Dementsprechend ist das angemessene Kriterium zur Beurteilung einer Lotterie nicht der Erwartungswert der Auszahlung sondern eben der Erwartungswert des Nutzens der möglichen Auszahlungen. Nach Bernoulli kann der Nutzen einer Auszahlung durch den Logarithmus des erreichten Endvermögens beschrieben werden. Unter diesen Annahmen kam Bernoulli zu einer Erklärung des St. Petersburger Paradoxons. Das Prinzip, Geldlotterien nach dem erwarteten Nutzen zu beurteilen, wird als *Erwartungsnutzen-Prinzip* bezeichnet. Zu Ehren Bernoullis ist insbeondere in der deutschen ökonomischen Literatur der Begriff *Bernoulli-Prinzip* verbreitet.

Eine solide axiomatische Grundlage erhielt das Erwartungsnutzen-Prinzip jedoch erst 1944 mit der ersten Auflage von John von Neumann und Oskar Morgenstern „Theory of Games and Economic Behavior“. Hier findet sich erstmals ein Axiomensystem, das ein Verhalten gemäß dem Erwartungsnutzen-Prinzip impliziert. Der Beweis dieser Aussage fehlt dort jedoch noch und findet sich erst in der zweiten Auflage von 1947. Obwohl sich die mathematische Formulierung der Theorien bei von Neumann/Morgenstern und Bernoulli ähneln, besteht inhaltlich ein wichtiger Unterschied. Während Bernoulli Wahrscheinlichkeiten und Nutzen als gegeben annimmt und davon ausgehend die Präferenzen zwischen Lotterien definiert, gehen von Neumann/Morgenstern den umgekehrten Weg. Ausgangspunkt ist bei ihnen eine Präferenz des Entscheidungsträgers auf Lotterien. von Neumann und Morgenstern zeigen, daß unter gewissen Annahmen an die Eigenschaften der Präferenz, die sie als Axiome formulieren, eine Funktion u existiert, so daß die Präferenz durch den Erwartungswert dieser Nutzenfunktion dargestellt wird. Daher ist die Interpretation der Nutzenfunktion in beiden Modellen eine völlig andere. Für Bernoulli hat die Nutzenfunktion eine Interpretation in dem Sinne, daß sie eine tatsächliche Gegebenheit beschreibt, nämlich den Nutzen, den ein bestimmter Entscheidungsträger aus einem gewissen Geldbetrag zieht. In von Neumann/Morgensterns Modell jedoch hat die Nutzenfunktion keine inhaltliche Bedeutung. Sie ist lediglich ein mathematisches Konstrukt, das zur Beschreibung der Präferenzen des Entscheidungsträgers dient.

Einen Schritt weiter in dieser Richtung geht Savage (1954). Während in von Neumann/Morgensterns Modell die Wahrscheinlichkeiten noch als gegeben angenommen werden, geht Savage von einer Formulierung des Entscheidungsproblems aus, in der die verschiedenen möglichen Aktionen des Entscheidungsträgers in Abhängigkeit von sogenannten Umweltzuständen bestimmte Konsequenzen liefern. Ein einfaches Beispiel eines solchen Entscheidungsproblems ist eine Pferdewette. Dabei sind die verschiedenen Aktionen die Wettmöglichkeiten des Entscheidungsträgers, die Umweltzustände die möglichen Ausgänge des Rennens und die Konsequenzen eben die resultierenden Auszahlungen an den Entscheidungsträger. Ausgehend von Axiomen an die Präferenz des Entscheidungsträgers leitet Savage simultan eine Nutzenfunktion (auf den Konsequenzen) und ein Wahrscheinlichkeitsmaß (auf den Umweltzuständen) ab. Da dieses Wahrscheinlichkeitsmaß keine objektiven Wahrscheinlichkeiten beschreibt, sondern lediglich das Vertrauen des Entscheidungsträgers in das Eintreffen eines bestimmten Ereignisses, bezeichnet man ein solches, aus Präferenzen abgeleitetes Wahrscheinlichkeitsmaß als *subjektives Wahrscheinlichkeitsmaß*. Da in Savages Theorie die Präferenzen des Entscheidungsträgers durch den Erwartungswert der Nutzenfunktion bezüglich dieses subjektiven Wahrscheinlichkeitsmaßes bewertet werden, bezeichnet man sein Modell als *subjektives Erwartungsnutzen-Modell.*

Bereits Anfang der fünfziger Jahre jedoch begann sich Kritik an diesen Modellen zu regen. Die von Allais (1953a,b) und Ellsberg (1961) konstruierten Entscheidungsprobleme und nachfolgende empirische Untersuchungen zeigten, daß Verhaltensmuster existieren, die mit dem Erwartungsnutzen-Modell und dem subjektivem-Erwartungsnutzen-Modell nicht adäquat beschrieben werden können. Die Existenz solcher Verhaltensmuster ist für sich allein genommen noch keine ernsthafte Kritik eines normativen Entscheidungsmodells. Entscheidungen, die den Axiomen eines Modells widersprechen, sind damit erklärbar, daß der Entscheidungsträger eine komplexe Entscheidungssituation nicht überblickt oder nicht angemessen analysiert hat. In empirischen Untersuchungen zeigte sich jedoch, daß Personen, die den Axiomen widersprechende Entscheidungen getroffen hatten, auch nach einer genauen Analyse des Problems sowie Diskussion der Axiome nicht bereit waren, ihre Entscheidungen zu revidieren. Solche Beobachtungen stellen nun aber in der Tat eine Herausforderung an eine Theorie dar, die für sich in Anspruch nimmt, rationales Entscheidungsverhalten zu beschreiben.

In der Folgezeit wurden alternative Modelle entwickelt, die in der Lage waren, gewisse häufig beobachtete Verhaltensmuster zu erklären. Grundidee einer Klasse dieser Modelle war, daß Wahrscheinlichkeiten transformiert werden, bevor sie in eine Entscheidung eingehen, d.h. genau wie im Erwartungsnutzen-Prinzip die Gewinne durch die Nutzenfunktion transformiert werden, sollten nun auch die Wahrscheinlichkeiten durch eine sogenannte Verzerrungsfunktion transformiert werden. Viele der entwickelten Modelle konnten zwar Phänomene wie das Allais-Paradox erklären, hatten jedoch an anderer Stelle Defizite. Die erste Axiomatisierung eines Modells, das sowohl der Idee Rechnung trug, die Wahrscheinlichkeiten transformiert in die Bewertung einer Lotterie eingehen zu lassen, als auch vom normativen Standpunkt aus gesehen keine offensichtlichen Unzulänglichkeiten besaß, gelang Quiggin (1982, erste Version 1979). Er bezeichnete sein Modell als „Antizipierter Nutzen“ („Anticipated Utility“). Daneben sind heute auch die Bezeichnungen „Erwartungsnutzen mit rangabhängigen Wahrscheinlichkeiten“ („Expected Utility with Rank Dependent Probabilities“) und „Rangabhängiger Erwartungsnutzen“ („Rank Dependent Expected Utility“) verbreitet. Im folgenden sprechen wir kürzer vom AU-Modell oder einfach von AU. Quiggin formulierte seine Theorie für Entscheidungsprobleme unter Risiko, d.h. bei gegebenen Wahrscheinlichkeiten. Wenig später entwickelte Schmeidler (1989, erste Version 1982) ein analoges Modell für Entscheidungsprobleme unter Unsicherheit, d.h. ohne gegebene Wahrscheinlichkeiten. Schmeidlers Modell und nachfolgende Axiomatisierungen dieses Modells bezeichnet man als „Choquet-Erwartungsnutzen“ („Choquet Expected Utility“) oder kürzer als CEU-Modell.

Die Analogie der beiden Modelle wird vor allem in der mathematischen Formulierung deutlich. Die verbindenden Begriffe sind *Choquet-Kapazität* und *Choquet-Integral.* Choquet-Kapazitäten werden gelegentlich auch als *nichtadditive Wahr-*

scheinlichkeitsmaße bezeichnet, da sie im Gegensatz zu Wahrscheinlichkeitsmaßen nicht mehr additiv sondern nur noch isoton bezüglich Teilmengenbeziehung sind. Integration bezüglich solcher Choquet-Kapazitäten wurde zuerst von Choquet (1953/54) definiert. Die Choquet-Integration stellt eine Verallgemeinerung des Lebesgueschen Integralbegriffes dar. In CEU und in AU wird die Präferenzrelation des Entscheidungsträgers durch das Choquet-Integral einer Nutzenfunktion bezüglich einer Choquet-Kapazität repräsentiert. In AU ist die Choquet-Kapazität durch das transformierte Wahrscheinlichkeitsmaß gegeben, in CEU ist sie die aus den Präferenzen des Entscheidungsträgers abgeleitete subjektive Bewertung der Ereignisse. Daß die Analogie zwischen diesen beiden Modellen über die mathematische Formulierung hinausgeht, zeigte Wakker (1990b). Er bewies, daß CEU, erweitert um ein zusätzliches Axiom und angewendet auf Entscheidungsprobleme unter Risiko, mit AU identisch ist.

Ziel dieser Arbeit ist eine weitgehend parallele Untersuchung dieser beiden Modelle auf der Grundlage der gemeinsamen mathematischen Formulierung.

Zum Aufbau dieser Arbeit. In Kapitel 1 stellen wir die entscheidungstheoretischen Grundlagen dar, die dieser Arbeit zugrunde liegen. Wir geben die mathematische Formulierung von Entscheidungsproblemen unter Unsicherheit und unter Risiko. Es folgt eine kurze Darstellung des Erwartungsnutzen-Modells und des subjektiven Erwartungsnutzen-Modells sowie eine Übersicht über verschiedene Axiomatisierungen des subjektiven Erwartungsnutzen-Modells. Danach gehen wir auf die wichtigsten Kritikpunkte am Erwartungsnutzen-Modell ein. Wir stellen von Neumann/Morgensterns Unabhängigkeitsaxiom und Savages sure-thing principle dar und zeigen wie in gewissen Entscheidungsproblemen (wie z.B. dem Allais- oder Ellsberg-Paradox) gegen diese Axiome verstoßen wird.

Die mathematischen Grundlagen zum Verständnis des CEU- und des AU-Modells werden in Kapitel 2 gelegt. Die grundlegenden Begriffe sind die der Choquet-Kapazität und des Choquet-Integrals. Wir geben die Definitionen von Choquet-Kapazität und Choquet-Integral und wiederholen die wichtigsten Eigenschaften des Choquet-Integrals. Wir formulieren dann den Satz über die Existenz eines assoziierter Wahrscheinlichkeitsmaße. Darunter verstehen wir ein Wahrscheinlichkeitsmaß auf einem schwach geordnetem Raum derart, daß für alle monoton wachsenden Funktionen das Choquet-Integral bezüglich der Kapazität und das Integral bezüglich des assoziierten Wahrscheinlichkeitsmaßes übereinstimmen. Dieser Satz erlaubt es uns, in vielen für die Anwendung wichtigen Situationen das Choquet-Integral als Integral bezüglich eines Wahrscheinlichkeitsmaßes zu betrachten. Mit Hilfe des Begriffs des assoziierten Wahrscheinlichkeitsmaßes machen wir einen Vorschlag für die Definition des Produktes zweier Choquet-Kapazitäten. Wir zeigen, daß das von uns vorgeschlagene Konzept viele schöne Eigenschaften besitzt und eine natürliche Verallgemeinerung des Produktes zweier Maße darstellt. Außerdem gilt, sofern man sich auf geeignete Funktionenklassen beschränkt, der Satz von Fubini auch für Produktkapazitäten. Anschließend zeigen wir die Verbindung zwischen Choquet-Kapazitäten und den sogenannten unscharfen Maßen auf. Im letzten Abschnitt des Kapitels finden sich die Beweise der zentralen Sätze dieses Kapitels.

In Kapitel 3 geben wir zunächst einen kleinen historischen Überblick über die Entwicklung des AU-Modells und seiner Vorgänger. Nachdem die mathematischen Grundlagen in Kapitel 2 gelegt wurden, können wir eine genaue mathematische Formulierung des CEU- und des AU-Modells geben. Wir werden außerdem zeigen, daß das CEU-Modell viele bekannten Entscheidungsregeln wie z.B. das Maximin-Prinzip als Spezialfall enthält. Ein eigener Abschnitt ist dem Zusammenhang zwischen dem CEU-Modell und dem Maximin-Prinzip bei vager Vorinformation gewidmet. Zwar ist weder das CEU-Modell Spezialfall des Maximin-Prinzips bei vager Vorinformation noch umgekehrt, jedoch besteht ein enger Zusammenhang zwischen diesen beiden Modellen. Anschließend erläutern wir kurz, wie in AU und CEU Phänomene wie das Allais- oder das Ellsberg-Paradox erklärt werden können.

Den Ausführungen in Kapitel 4 liegt die Arbeit Dyckerhoff und Mosler (1993) zugrunde. Zentraler Gegenstand ist der Begriff der stochastischen Dominanz von Kapazitäten. Stochastische Dominanz von Wahrscheinlichkeitsmaßen ermöglicht es im Erwartungsnutzen-Modell, Entscheidungen zwischen zwei gegebenen Alternativen zu treffen, wenn nur bestimmte qualitative Eigenschaften der Nutzenfunktion, wie z.B. Monotonie oder Konkavität bekannt sind. Wir verallgemeinern den Begriff der stochastischen Dominanz von Wahrscheinlichkeitsmaßen auf Choquet-Kapazitäten und zeigen, daß viele der bekannten Resultate aus dem Erwartungsnutzen-Modell auch in CEU und AU gültig sind. Untersucht werden verschiedene stochastische Dominanzrelationen. Im zweiten Abschnitt konzentrieren wir uns auf Dominanzrelationen, die durch Ungleichungen zwischen den Kapazitäten auf gewissen Teilmengen des Raumes charakterisiert werden können. Sind die Kapazitäten auf einem schwach geordnetem Raum definiert und existiert ein assoziiertes Wahrscheinlichkeitsmaß, so ist es ebenfalls einfach, Ergebnisse vom Erwartungsnutzenmodell auf AU und CEU zu verallgemeinern. In einem weiteren Abschnitt untersuchen wir stochastische Dominanzrelationen von Produktkapazitäten. Wir zeigen, daß in wichtigen Fällen Dominanz der Produktkapazitäten auf Dominanz der Randkapazitäten zurückgeführt werden kann. Schließlich zeigen wir noch, wie die erhaltenen Resultate über stochastische Dominanz von Choquet-Kapazitäten auf CEU- und AU-Modelle angewendet werden können. Insbesondere gehen wir darauf ein, wie in AU auf der Grundlage stochastischer Dominanz Entscheidungen zwischen Alternativen gefällt werden können, wenn sowohl Nutzen- als auch Verzerrungsfunktion nur unvollständig bekannt sind.

Einen knappen Überblick über Risikoaversion in AU-Modellen und Unsicherheitsaversion in CEU-Modellen gibt Kapitel 5. In EU sind verschiedene Definitionen von Risikoaversion gebräuchlich, die alle zur Konkavität der Nutzenfunktion äquivalent sind. Im Gegensatz dazu sind diese Definitionen von Risikoaversion in AU nicht äquivalent. Das AU-Modell erlaubt somit eine differenziertere Modellierung von Risikoaversion, als es im Erwartungsnutzen-Modell möglich ist. Ebenso erlaubt das CEU-Modell im Gegensatz zum subjektiven Erwartungsnutzen-Modell die Modellierung von Unsicherheitsaversion oder Ambiguitätsaversion.

In Kapitel 6, das weitgehend Dyckerhoff (1994) folgt, präsentieren wir eine Dekompositionstheorie für multivariate Nutzenfunktionen in AU- und CEU-Modellen. Im Erwartungsnutzen-Modell sind viele Bedingungen bekannt, unter denen sich eine multivariate Nutzenfunktion in einfacher Form aus univariaten Nutzenfunktionen zusammensetzt. So sind Nutzenfunktionen gebräuchlich, bei denen sich die multivariate Nutzenfunktion als Summe, Produkt oder als Summe von Produkten univariater Nutzenfunktionen darstellen läßt. Diese Nutzenfunktionen werden als additiv, multiplikativ und multilinear bezeichnet. Wir formulieren zunächst die Bedingungen, unter denen sich im Erwartungsnutzen-Modell Darstellungen dieser Art ergeben. Sodann geben wir notwendige und hinreichende Bedingungen dafür, daß die Nutzenfunktion in einem AU- oder CEU- Modell multiplikativ oder multilinear ist. Wir zeigen außerdem, daß die Bedingungen, die im Erwartungsnutzen-Modell eine additive Darstellung der Nutzenfunktion implizieren, in AU-Modellen auch Einschränkungen für die Verzerrungsfunktion zur Folge haben. So folgt z.B. aus der sogenannten Randverteilungsbedingung, die besagt, daß die Präferenzen des Entscheidungsträgers nur von den Randverteilungen einer Lotterie abhängen, nicht nur, daß die Nutzenfunktion additiv ist, sondern auch, daß die Verzerrungsfunktion die identische Abbildung ist. Die Randverteilungsbedingung impliziert also ein Verhalten gemäß dem Erwartungsnutzen-Prinzip. Zwei Abschnitte dieses Kapitels befassen sich mit multivariater Risikohaltung. Zum einen untersuchen wir bivariate Risikoaversion, wie sie von Richard (1975) definiert wurde, zum anderen die sogenannte Korrelationsaversion nach Epstein und Tanny (1980). Wir geben Charakterisierungen dieser Begriffe in AU-Modellen. Auch hier wird sich zeigen, daß im AU-Modell eine größere Vielzahl von bivariaten Risikohaltungen modelliert werden kann als im Erwartungsnutzen-Modell. Die Beweise der Sätze dieses Kapitels sind wieder im letzten Abschnitt zusammengefaßt.

Kapitel 7 geht der Frage nach, wie Nutzenfunktion und Choquet-Kapazität bzw. Nutzenfunktion und Verzerrungsfunktion bestimmt werden können. Die Standardmethoden zur Bestimmung der Nutzenfunktion im Erwartungsnutzen-Modell sind in AU und CEU nicht mehr anwendbar, da in diesen Modellen zusätzlich Verzerrungsfunktion und Choquet-Kapazität unbekannt sind. Wir betrachten zunächst Schmeidlers Axiomatisierung des CEU-Modells. In ihr sind die Konsequenzen Lotterien mit objektiven Wahrscheinlichkeiten. Dadurch ist es möglich den Prozeß der Bestimmung von Nutzenfunktion und Choquet-Kapazität in zwei Phasen aufzuteilen und beide getrennt zu bestimmen. In Rahmen anderer Axiomatisierungen, die keine objektiven Wahrscheinlichkeiten verwenden, ist diese

Trennung nicht mehr möglich. Wir beschreiben zwei verschiedene Methoden zur Bestimmung von Choquet-Kapazität und Nutzenfunktion. Je nachdem, ob der Konsequenzenraum oder der Zustandsraum reichhaltig sind, kann zumindest eine der beiden Methoden angewendet werden. Danach formulieren wir Verfahren zur Bestimmung von Nutzen- und Verzerrungsfunktion in AU. Aufgrund der Analogie zwischen AU und CEU können die für das CEU-Modell entwickelten Methoden gut auf das AU-Modell übertragen werden. Als letztes präsentieren wir noch einen parametrischen Ansatz zur Bestimmung der Verzerrungsfunktion in AU. Hier wird die Verzerrungsfunktion nach einem kleinste-Quadrate-Ansatz aus einer parametrischen Klasse von Verzerrungsfunktionen gewählt.

In Kapitel 8 beschreiben wir die Implikationen, die sich aus der Verwendung von AU und CEU in Anwendungssituationen ergeben. Es wird gezeigt werden, daß mit AU und CEU ein Entscheidungsverhalten modelliert werden kann, das realistischer ist, als das vom Erwartungsnutzen-Modell implizierte. Als Anwendungen werden wir Prämienprinzipien bei Versicherungen sowie Portfolio-Probleme betrachten.

Anhang A enthält mathematische Grundbegiffe der Ordnungs- und Maßtheorie, in Anhang B sind vier verschiedene Axiomatisierungen von CEU-Modellen dargestellt.

Bezeichnungen

Wir bezeichnen mit $\mathbb{N}$ die Menge der natürlichen Zahlen $\{1, 2, 3, \ldots\}$ und mit $\mathbb{N}_0$ die Menge der natürlichen Zahlen einschließlich der Null. Weiter sei $\mathbb{Z}$ die Menge der ganzen Zahlen, $\mathbb{Q}$ die Menge der rationalen Zahlen und mit $\mathbb{R}$ die Menge der reellen Zahlen. Fügt man zu $\mathbb{R}$ die uneigentlichen Zahlen ∞ und $-\infty$ hinzu, so erhält man die *erweiterte reelle Zahlengerade* $\overline{\mathbb{R}}$. Mit ∞ und $-\infty$ wird in der üblichen Weise gerechnet, d.h. $\infty + x = \infty$ für alle $x \in \mathbb{R}$ und $-\infty + x = -\infty$ für alle $x \in \mathbb{R}$. Der Ausdruck $\infty - \infty$ ist nicht definiert. $\mathbb{R}_+$ bezeichne die Menge aller reellen Zahlen x mit $x \geq 0$ und $\mathbb{R}_-$ die Menge aller reellen Zahlen x mit $x \leq 0$. Bezeichnungen wie $\mathbb{Q}_+$ oder $\mathbb{Z}_-$ sind analog zu verstehen.

Mit den Symbolen $\cup$, $\cap$ und $\setminus$ bezeichnen wir die mengentheoretischen Operationen Vereinigung, Schnitt und Differenz. Das Komplement einer Menge A bezeichnen wir mit A^c. $A \subset B$ bedeutet, daß A Teilmenge von B ist, wobei $A = B$ zugelassen ist. Ist A echte Teilmenge von B, so schreiben wir $A \underset{\neq}{\subset} B$. Die leere Menge bezeichnen wir mit $\emptyset$. Ist Ω eine Menge, so bezeichnen wir mit $\mathcal{P}(\Omega)$ die Potenzmenge von Ω, d.h. die Menge aller Teilmengen von Ω.

Schreiben wir $f : A \to B$, so meinen wir damit, daß f eine Funktion von A nach B ist. Ist $B = \mathbb{R}$, so bezeichnen wir f auch als reelle Funktion. $f : \mathbb{R} \to \mathbb{R}$ heißt *monoton wachsend* (fallend), wenn gilt $x \leq y \Rightarrow f(x) \leq f(y)$ $(x \leq y \Rightarrow f(x) \geq f(y))$. Wir sagen, daß $f : \mathbb{R} \to \mathbb{R}$ *streng monoton wachsend* (fallend) ist, wenn gilt $x < y \Rightarrow f(x) < f(y)$ $(x < y \Rightarrow f(x) > f(y))$.

Weitergehende mathematische Grundbegriffe, insbesondere der Ordnungstheorie und der Maßtheorie, finden sich in Anhang A. Für die verwendeten topologischen Begriffe und Ergebnisse sei auf Standardwerke der Topologie verwiesen, siehe z.B. Willard (1970).

Das Ende eines Beweises wird mit $\blacksquare$ gekennzeichnet.

Kapitel 1

Entscheidungstheoretische Grundlagen

1.1 Entscheidungstheoretische Modelle

In diesem Abschnitt beschreiben wir die Standardmodelle der Entscheidungstheorie, in deren Rahmen sich die von uns betrachteten Modelle bewegen. Das Grundproblem der Entscheidungstheorie kann anschaulich wie folgt beschrieben werden. Ein Entscheidungsträger (im folgenden auch kurz als ET bezeichnet) befindet sich in einer Situation, in der er verschiedene Aktionen ergreifen kann. Jede der möglichen Aktionen kann zu verschiedenen Konsequenzen führen. Dabei können die Konsequenzen monetäre Ergebnisse sein, müssen es jedoch nicht sein. Verschiedene Aktionen werden von dem ET als verschieden „gut“ betrachtet, d.h. er zieht gewisse Aktionen anderen vor. Dieses Problem bezeichnet man als *Entscheidungsproblem unter Unsicherheit*. Ein Entscheidungsproblem unter Unsicherheit kann mathematisch wie folgt modelliert werden.

Entscheidung unter Unsicherheit

Ein Entscheidungsproblem unter Unsicherheit ist durch folgende Parameter charakterisiert.

Zustandsraum. Es sei S eine nichtleere Menge, die mit einer Algebra $\mathcal{A}$ versehen sei. S heißt *Zustandsraum*, die Elemente von S heißen Zustände.

Konsequenzenraum. Es sei $\mathcal{C}$ eine nichtleere Menge, die mit einer Algebra $\mathcal{D}$ versehen sei. $\mathcal{D}$ enthalte alle Einpunktmengen. $\mathcal{C}$ heißt *Konsequenzenraum*, die Elemente von $\mathcal{C}$ heißen Konsequenzen.

Aktionen. Es sei $\mathcal{F} \subset \{f : S \to \mathcal{C} \,|\, f \text{ } \mathcal{A}\text{-}\mathcal{D}\text{-meßbar}\}$ eine nichtleere Menge von $\mathcal{A}$-$\mathcal{D}$-meßbaren Funktionen. $\mathcal{F}$ enthalte alle meßbaren Funktionen, die nur endlich viele Werte annehmen. Die Elemente von $\mathcal{F}$ heißen *Aktionen.*

Präferenz. Es sei $\preceq$ eine binäre Relation auf $\mathcal{F}$. $\preceq$ heißt *Präferenzrelation des ET.*

Wir bezeichnen das Quadrupel $((S, \mathcal{A}), (\mathcal{C}, \mathcal{D}), \mathcal{F}, \preceq)$ als *Entscheidungsproblem unter Unsicherheit.*

Die Interpretation ist wie folgt. Einer der Zustände ist der „wahre“ Zustand. Es ist nicht bekannt, welcher der Zustände der wahre Zustand ist. Eine Aktion ordnet jedem möglichem Zustand eine Konsequenz zu. Ist also $s \in S$ der wahre Zustand, so ist das Ergebnis der Aktion $f \in \mathcal{F}$ die Konsequenz $f(s) \in \mathcal{C}$. Die Relation $\preceq$ beschreibt die Präferenzen des ET. Gilt für zwei Aktionen $f \preceq g$, so zieht der ET g gegenüber f vor, d.h. er betrachtet g als mindestens so gut wie f.

Für Aktionen, die nur endlich viele Konsequenzen liefern, benutzen wir die folgende Notation:

$$f = \begin{bmatrix} x_1 & \dots & x_n \\ A_1 & \dots & A_n \end{bmatrix}$$

bezeichnet die Aktion f mit $f(s) = x_i$, falls $s \in A_i$, $i = 1, \dots n$.

Entscheidung unter Risiko

Von einem *Entscheidungsproblem unter Risiko* spricht man, wenn die beiden folgenden, zusätzlichen Annahmen erfüllt sind.

(i) Es liegen objektive Wahrscheinlichkeiten über das Eintreffen der Zustände vor, d.h. es ist ein (endlich additives) W-Maß P auf $(S, \mathcal{A})$ gegeben.

(ii) Es gilt das sogenannte *Neutralitätsaxiom*: Induzieren zwei Aktionen die gleiche Verteilung auf dem Konsequenzenraum, so sind die beiden Aktionen äquivalent. Formal: $f, g \in \mathcal{F}$, $P^f = P^g \Rightarrow f \sim g$.

Unter diesen beiden Annahmen hängen die Präferenzen des ET nur von den auf den Konsequenzen induzierten Wahrscheinlichkeitsverteilungen ab. Da in einem Entscheidungsproblem unter Risiko der Zustandsraum keine Relevanz für die Entscheidung hat, charakterisiert man ein solches Entscheidungsproblem durch die folgenden Größen.

Konsequenzenraum. Es sei $\mathcal{C}$ eine nichtleere Menge, die mit einer Algebra $\mathcal{D}$ versehen sei. $\mathcal{D}$ enthalte alle Einpunktmengen. $\mathcal{C}$ heißt *Konsequenzenraum*, die Elemente von $\mathcal{C}$ heißen Konsequenzen.

Lotterien. Es sei $\mathcal{P}$ eine Menge von (endlich additiven) Wahrscheinlichkeitsmaßen auf $(\mathcal{C}, \mathcal{D})$. $\mathcal{P}$ enthalte alle Wahrscheinlichkeitsmaße mit endlichem Träger. $\mathcal{P}$ heißt Menge aller *Lotterien*.

Präferenz. Es sei $\preceq$ eine binäre Relation auf $\mathcal{P}$. $\preceq$ heißt *Präferenzrelation des ET*.

Das Tripel $((\mathcal{C}, \mathcal{D}), \mathcal{P}, \preceq)$ heißt dann *Entscheidungsproblem unter Risiko*. In dieser Formulierung wird der zugrunde liegende Zustandsraum nicht mehr erwähnt, da die gesamte für die Entscheidung relevante Information in den Wahrscheinlichkeitsverteilungen auf $\mathcal{C}$ enthalten ist. Wir können jedoch die W-Maße in $\mathcal{P}$ immer als Verteilungen von Zufallsvariablen auf einem gemeinsamen Zustandsraum ansehen. Bei einem Entscheidungsproblemen unter Risiko werden wir deshalb auch gelegentlich davon sprechen, daß der ET eine Zufallsvariable X einer Zufallsvariablen Y vorzieht, wenn wir meinen, daß der ET die Verteilung P^X der Verteilung P^Y vorzieht.

Für Lotterien mit endlichem Träger benutzen wir die folgende Notation:

$$P = \begin{bmatrix} x_1 & \dots & x_n \\ p_1 & \dots & p_n \end{bmatrix}$$

bezeichnet die Lotterie P mit $P(\{x_i\}) = p_i$, $i = 1, \dots n$, d.h. die Lotterie, die mit Wahrscheinlichkeit p_i die Konsequenz x_i liefert.

1.2 Erwartungsnutzen und subjektiver Erwartungsnutzen

Das vorherrschende Modell in der Entscheidungstheorie ist nach wie vor das Erwartungsnutzen-Modell (kurz EU) bzw. das subjektive Erwartungsnutzen-Modell (SEU). Beide Modelle sind dadurch charakterisiert, daß die Präferenzrelation des ET durch den Erwartungswert des Nutzens einer Aktion bezüglich eines additiven Wahrscheinlichkeitsmaßes repräsentiert wird.

Definition 1.1 (EU-Modell) *Sei* $((\mathcal{C}, \mathcal{D}), \mathcal{P}, \preceq)$ *ein Entscheidungsproblem unter Risiko. Existiert eine Funktion* $u : \mathcal{C} \to \mathbb{R}$*, so daß die Präferenzrelation* $\preceq$ *des ET durch das EU-Funktional*

$$EU : \mathcal{P} \to \overline{\mathbb{R}}, \; P \mapsto \int_{\mathcal{C}} u \, dP,$$

repräsentiert wird, so bezeichnen wir den ET als EU-Maximierer *und das Entscheidungsproblem als* EU-Modell*. In diesem Fall heißt* u Nutzenfunktion.

Das erste Axiomensystem eines EU-Modells präsentierten John von Neumann und Oskar Morgenstern 1944 in ihrem Buch „Theory of Games and Economic Behavior". Der Beweis dieser Axiomatisierung wurde jedoch erst in die zweite Auflage von Neumann und Morgenstern (1947) aufgenommen. Da es sich beim EU-Modell um ein Entscheidungsproblem unter Risiko handelt, werden die Wahrscheinlichkeiten als gegeben angenommen. Das Analogon zum EU-Modell ist bei Entscheidungsproblemen unter Unsicherheit das SEU-Modell. Hier wird das W-Maß nicht als gegeben angenommen, sondern zusammen mit der Nutzenfunktion aus der Präferenzrelation des ET abgeleitet. Schon Ramsey (1931) skizzierte Axiome eines SEU-Modells. Die erste vollständige Axiomatisierung eines SEU-Modells stammt jedoch von Savage (1954). Andere Axiomatisierungen finden sich in Anscombe und Aumann (1963), Krantz et al. (1971) und anderen Arbeiten.

Definition 1.2 (SEU-Modell) *Sei* $((S,\mathcal{A}),(\mathcal{C},\mathcal{D}),\mathcal{F},\preceq)$ *ein Entscheidungsproblem unter Unsicherheit. Existiert eine Funktion* $u:\mathcal{C}\to\mathbb{R}$ *und ein (endlich additives) Wahrscheinlichkeitsmaß* P *auf* $(S,\mathcal{A})$*, so daß die Präferenzrelation* $\preceq$ *des ET durch das SEU-Funktional*

$$SEU:\mathcal{F}\to\overline{\mathbb{R}},\ f\mapsto\int_S(u\circ f)\,dP,$$

repräsentiert wird, so bezeichnen wir den ET als SEU-Maximierer *und das Entscheidungsproblem als* SEU-Modell. *In diesem Fall heißt* u Nutzenfunktion *und* P subjektives Wahrscheinlichkeitsmaß.

Unterschiede in den Axiomatisierungen

Wir wollen nun skizzieren, auf welche Art sich die Axiomatisierungen von SEU-Modellen unterscheiden. Bislang wurden über den Zustandsraum und den Konsequenzenraum lediglich vorausgesetzt, daß beide Räume nicht leer sind. In den Axiomatisierungen müssen jedoch weitergehende strukturelle Annahmen gemacht werden. Insbesondere müssen zumindest der Zustandsraum oder der Konsequenzenraum „reichhaltig genug" sein. Wir unterscheiden die folgenden Ansätze:

Savages Ansatz. In Savage (1954) werden zwar explizit keine Voraussetzungen an die Strukturen der Räume gemacht, jedoch impliziert Savages Axiom P6, daß der Zustandsraum beliebig fein unterteilt werden kann. Insbesondere folgt, daß S unendlich sein muß (jedoch nicht notwendigerweise überabzählbar). Über den Konsequenzenraum werden, abgesehen von Nichttrivialitätsannahmen, keine weitergehenden Annahmen gemacht.

Der algebraische Ansatz. Im algebraischen Ansatz definiert man eine algebraische Struktur auf der Menge der Aktionen. Resultate über die Einbettbarkeit von archimedischen geordneten Gruppen in die reellen Zahlen führen dann zu einer Darstellung der Präferenzrelation durch eine numerische Funktion. Kennzeichnend für Axiomatisierungen im Rahmen des algebraischen Ansatz sind i.a. ein

Axiom, das die Lösbarkeit gewisser Gleichungen oder Ungleichungen sichert und ein sogenanntes Archimedisches Axiom. Diese Axiome implizieren insbesondere, daß der Konsequenzenraum reichhaltig genug ist. Der algebraische Ansatz wurde vor allem von Krantz et al. (1971) entwickelt.

Der topologische Ansatz. In diesem Ansatz werden topologische Voraussetzungen an den Konsequenzenraum gestellt. I.a. wird vorausgesetzt, daß $\mathcal{C}$ ein zusammenhängender topologischer[1] Raum und die Präferenzrelation $\preceq$ des ET eine stetige schwache Ordnung ist. Dies impliziert dann, daß auch die Nutzenfunktion u stetig ist. Ist u nicht konstant, so ist das Bild von $\mathcal{C}$ unter u eine zusammenhängende Teilmenge von $\mathbb{R}$ und somit ein Intervall. Insbesondere muß also auch $\mathcal{C}$ überabzählbar sein. Axiomatisierungen von SEU im topologischen Ansatz finden sich z.B. in Wakker (1989a).

Der Ansatz von Anscombe und Aumann. Anscombe und Aumann (1963) gehen davon aus, daß der Konsequenzenraum $\mathcal{C}$ die Menge aller Lotterien mit endlichem Träger auf einer Menge Γ ist. Die Elemente von Γ heißen *Preise*. Dieser Ansatz stellt in gewisser Weise eine Mischung aus Entscheidungsproblemen unter Unsicherheit und Risiko dar, da zwar auf dem Zustandsraum kein W-Maß gegeben ist, die Konsequenzen jedoch Lotterien sind. Es wird also in diesem Ansatz davon ausgegangen, daß *objektive Wahrscheinlichkeiten* verfügbar sind. Diese Lotterien mit bekannter Wahrscheinlichkeitsverteilung werden als „roulette lotteries“ bezeichnet. Die Aktionen hingegen werden als „horse lotteries“ bezeichnet, da keine objektiven Wahrscheinlichkeiten des Eintreffens der einzelnen Ereignisse gegeben sind.

1.3 Kritik an EU und SEU

Eine Axiomatisierung eines bestimmten Modells für Entscheidungsverhalten (z.B. EU oder SEU) postuliert eine Reihe von Axiomen, die als plausible Annahmen über rationales Entscheidungsverhalten angesehen werden können. Es wird dann gezeigt, daß ein ET, der diese Axiome als Grundlagen eines rationalen Verhaltens akzeptiert, sich notwendigerweise gemäß dem betrachteten Entscheidungsmodell verhält.

Das grundlegende Axiom in Axiomatisierungen von EU und SEU ist ein sogenanntes Unabhängigkeitsaxiom. In Savages (1954) Axiomatisierung des SEU-Modells wird dieses Unabhängigkeitsaxiom als *sure-thing principle* bezeichnet. Um das sure-thing principle zu formulieren, benötigen wir die folgende Notation. Sind f, g Aktionen und ist $A \subset S$, so bezeichnen wir mit $f_A g_{-A}$ die Aktion, für

[1] Diese und weitere verwendete topologische Begriffe und Ergebnisse finden sich z.B. in Willard (1970).

die gilt

$$f_A g_{-A}(s) = \begin{cases} f(s), & \text{falls } s \in A, \\ g(s), & \text{falls } s \in A^c. \end{cases}$$

Sure-thing principle: Für alle Aktionen f, g, h, k und alle $A \subset S$ gilt

$$f_A h_{-A} \preceq g_A h_{-A} \iff f_A k_{-A} \preceq g_A k_{-A}.$$

Das sure-thing principle besagt also, daß für die Entscheidung zwischen zwei Aktionen nur die Zustände relevant sind, in denen sich die beiden Aktionen unterscheiden.

In von Neumann und Morgensterns (1947) Axiomatisierung von EU wird das Unabhängigkeitsaxiom wie folgt formuliert.

Unabhängigkeitsaxiom: Für alle Lotterien P_1, P_2, P_3 und alle $\alpha \in [0, 1]$ gilt:

$$P_1 \preceq P_2 \iff \alpha P_1 + (1-\alpha)P_3 \preceq \alpha P_2 + (1-\alpha)P_3.$$

Die Lotterie $\alpha P_1 + (1-\alpha)P_3$ kann als zweistufige Lotterie betrachtet werden. Das Ergebnis der ersten Lotterie ist mit Wahrscheinlichkeit α eine Teilnahme an der Lotterie P_1, mit Wahrscheinlichkeit $1-\alpha$ eine Teilnahme an der Lotterie P_3. Entsprechendes gilt für die zweistufige Lotterie $\alpha P_2 + (1-\alpha)P_3$. Das Unabhängigkeitsaxiom besagt nun, daß ein ET, der die Lotterie P_2 der Lotterie P_1 vorzieht, auch die zweistufige Lotterie $\alpha P_2 + (1-\alpha)P_3$ der Lotterie $\alpha P_1 + (1-\alpha)P_3$ vorzieht.

So plausibel diese beiden Axiome auf den ersten Blick auch scheinen mögen, so hat sich doch gezeigt, daß in realen Entscheidungsproblemem diese Axiome systematisch verletzt werden. In diesem Abschnitt stellen wir die wichtigsten dieser beobachteten Verletzungen der Axiome dar und zeigen, daß das in ihnen offenbarte Verhalten im Rahmen von EU oder SEU nicht erklärbar ist. Für einen Überblick über empirische Untersuchungen vergleiche Abschnitt 3.1.

Eines der wichtigsten sogenannten „Paradoxa“ stellt das sogenannte Ellsberg-Paradox vor, das erstmals von Ellsberg (1961) beschrieben wurde.

Beispiel 1.1 (Ellsberg-Paradox) *In einer Urne befinden sich* 90 *Kugeln, von denen 30 rot sind. Die übrigen 60 Kugeln sind jeweils gelb oder blau. Es ist aber nicht bekannt, wieviele dieser 60 Kugeln blau und wieviele gelb sind. Es wird nun blind eine Kugel aus der Urne gezogen. Dem ET werden zwei Paare von Lotterien vorgelegt, zwischen denen er sich jeweils entscheiden muß. Das erste Paar von Lotterien lautet wie folgt:*

Lotterie 1 *Der ET erhält 100.- DM, falls die gezogene Kugel rot ist, andernfalls nichts.*

Lotterie 2 *Der ET erhält 100.- DM, falls die gezogene Kugel gelb ist, andernfalls nichts.*

Das zweite Paar von Lotterien lautet:

Lotterie 3 *Der ET erhält 100.- DM, falls die gezogene Kugel rot oder blau ist, andernfalls nichts.*

Lotterie 4 *Der ET erhält 100.- DM, falls die gezogene Kugel gelb oder blau ist, andernfalls nichts.*

In dieser Entscheidungssituation zog bei Ellsberg der größte Teil der Befragten Lotterie 1 gegenüber Lotterie 2 und Lotterie 4 gegenüber Lotterie 3 vor.

Dieses Verhaltensmuster wird allgemein als „Ellsberg-Paradox" bezeichnet, da es mit dem SEU-Modell nicht erklärbar ist, wie im folgenden gezeigt wird. Der Zustandsraum ist offensichtlich $S = \{r, g, b\}$, wobei r (g, b) bedeutet: „Eine rote (gelbe,blaue) Kugel wird gezogen". Nach dem SEU-Modell existieren ein W-Maß P auf S und eine Nutzenfunktion auf dem Konsequenzenraum $\mathcal{C} = \mathbb{R}$, so daß der ET die Aktion mit dem größten Erwartungsnutzen wählt. Sei $p_r = P(\{r\})$. p_g und p_b seien analog definiert. Wir setzen noch $u(0) = 0$ und $u(100) = 1$. Dann sind die subjektiven Erwartungsnutzen der Lotterien wie folgt gegeben.

$$\begin{array}{ll} SEU(L_1) = p_r & SEU(L_2) = p_g \\ SEU(L_3) = p_r + p_b & SEU(L_4) = p_g + p_b \end{array}$$

Nach dem SEU-Modell wird also Lotterie 1 genau dann Lotterie 2 vorgezogen, wenn auch Lotterie 3 gegenüber Lotterie 4 vorgezogen wird. Es gibt also *kein* W-Maß, das das beobachtete Verhalten im Rahmen von SEU erklären könnte. Man beachte, daß bei unserer Argumentation nicht vorausgesetzt wurde, daß $p_r = \frac{1}{3}$ ist. An der Darstellung der Lotterie in der folgenden Tabelle erkennt man, daß das Ellsberg-Paradox gegen das sure-thing principle verstößt. Beide Lotterienpaare unterscheiden sich nur in der „blauen" Spalte, sind aber ansonsten gleich.

	rot	gelb	blau
Lotterie 1	100	0	0
Lotterie 2	0	100	0
Lotterie 3	100	0	100
Lotterie 4	0	100	100

Beim Ellsberg-Paradox handelt es sich um ein Entscheidungsproblem unter Unsicherheit. Das vermutlich bekannteste „Paradoxon" bei Entscheidungsproblemen unter Risiko ist das sogenannte Allais-Paradox, das von Allais (1953a,b, 1979) beschrieben wurde.

Beispiel 1.2 (Allais-Paradox) *Einem ET werden die beiden Lotterien*

$$L_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \textit{und} \quad L_2 = \begin{bmatrix} 5 & 1 & 0 \\ 0.1 & 0.89 & 0.01 \end{bmatrix}$$

präsentiert, wobei die Konsequenzen als Auszahlungen an den ET in Millionen DM zu verstehen sind. Nachdem sich der ET zwischen den beiden Lotterien entschieden hat, werden ihm die Lotterien

$$L_3 = \begin{bmatrix} 1 & 0 \\ 0.11 & 0.89 \end{bmatrix} \quad \textit{und} \quad L_4 = \begin{bmatrix} 5 & 0 \\ 0.1 & 0.9 \end{bmatrix}$$

zur Entscheidung vorgelegt. Allais berichtet, daß die Mehrzahl der mit diesen Entscheidungsproblemen konfrontierten Personen sich für L_1 und für L_4 entscheidet.

Diese Beobachtung ist mit dem EU-Modell unvereinbar. Es ist ja

$$\begin{aligned} EU(L_1) - EU(L_2) &= u(1) - [0.1u(5) + 0.89u(1) + 0.01u(0)] \\ &= 0.11u(1) - 0.1u(5) - 0.01u(0) \end{aligned}$$

und

$$\begin{aligned} EU(L_3) - EU(L_4) &= [0.11u(1) + 0.89u(0)] - [0.1u(5) + 0.9u(0)] \\ &= 0.11u(1) - 0.1u(5) - 0.01u(0)\,. \end{aligned}$$

Daher gilt gemäß EU $L_1 \succ L_2$ genau dann, wenn $L_3 \succ L_4$ ist. Daß dieses Verhalten dem Unabhängigkeitsaxiom widerspricht, ist wie folgt zu sehen. Die Lotterien P_1, P_2 und P_3 seien definiert durch

$$P_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 & 5 \\ \frac{1}{11} & \frac{10}{11} \end{bmatrix}, \quad P_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Dann lassen sich die Lotterien L_1 bis L_4 darstellen als

$$\begin{aligned} &L_1 = 0.11P_1 + 0.89P_1, \quad L_2 = 0.11P_2 + 0.89P_1, \\ &L_3 = 0.11P_1 + 0.89P_3, \quad L_4 = 0.11P_2 + 0.89P_3. \end{aligned}$$

Ist $P_1 \preceq P_2$, so ist nach dem Unabhängigkeitsaxiom sowohl $L_1 \preceq L_2$ als auch $L_3 \preceq L_4$. Ist aber $P_2 \preceq P_1$, so folgt entsprechend $L_1 \succeq L_2$ und $L_3 \succeq L_4$.

Zur Verteidigung von EU werden gegen das Allais-Paradox vor allem zwei Argumente ins Feld geführt. Zum einen wird behauptet, daß Personen, die sich gemäß dem Allais-Paradox verhalten, ihre Entscheidung nach einer genauen Analyse des Problems revidieren. Es zeigt sich jedoch, daß auch nach einer solchen Analyse ein großer Teil der befragten Personen ihre Entscheidung $L_1 \succ L_2$, $L_3 \prec L_4$ beibehält (siehe MacCrimmon 1968, Slovic und Tversky 1974, Moskowitz 1974 und MacCrimmon und Larsson 1979). Somit kann das im Allais-Paradox beobachtete Verhalten nicht einfach als irrational oder paradox bezeichnet werden. Ein zweiter Kritikpunkt am Allais-Paradox sind die verwendeten extremen Auszahlungen. Es zeigt sich jedoch, daß das Allais-Paradox kein isoliertes Beispiel ist, sondern daß es ein Spezialfall eines allgemeinen Verhaltensmusters ist, das als *„common consequence effect“* bezeichnet wird. Vergleiche dazu die Untersuchungen von MacCrimmon und Larsson (1979), Hagen (1979) und Kahneman und Tversky (1979), in denen moderatere Auszahlungen verwendet werden als im Allais-Paradox.

Beispiel 1.3 (Der „common consequence effect“) *Als „common consequence effect“ wird das gleichzeitige Auftreten von Präferenzen der folgenden Art verstanden:*

$$\begin{bmatrix} x_1 & x_2 & x_3 \\ p_1 & p_2 & p_3 \end{bmatrix} \prec \begin{bmatrix} x_2 \\ 1 \end{bmatrix} \quad \textit{und} \quad \begin{bmatrix} x_1 & x_3 \\ p_1 + p_2 & p_3 \end{bmatrix} \succ \begin{bmatrix} x_1 & x_2 \\ p_2 & p_1 + p_3 \end{bmatrix}.$$

Dieses Verhaltensmuster zeigt sich oft in empirischen Untersuchungen, wenn p_1 relativ klein ist, $x_2 \prec x_3$ und x_2, x_3 groß gegenüber x_1 sind. Die „gemeinsame Konsequenz“ im ersten Lotterienpaar ist ein Gewinn von x_2 mit Wahrscheinlichkeit p_2, im zweiten Lotterienpaar ist es ein Gewinn von x_1 mit Wahrscheinlichkeit p_2.

Für $p_1 = 0.01$, $p_2 = 0.89$, $p_3 = 0.1$ und $x_1 = 0$, $x_2 = 1\,000\,000$, $x_3 = 5\,000\,000$ erhält man das Allais-Paradox. Kahneman und Tversky (1979) verwendeten $p_1 = 0.01$, $p_2 = 0.66$, $p_3 = 0.33$ und $x_1 = 0$, $x_2 = 2400$, $x_3 = 2500$. Der common consequence effect ist im Rahmen von EU nicht zu erklären.

Ein verwandter Effekt ist der sogenannte „common ratio effect“. Vergleiche dazu MacCrimmon und Larsson (1979) und Kahneman und Tversky (1979).

Beispiel 1.4 (Der „common ratio effect“) *Hierunter versteht man das gleichzeitige Auftreten von Präferenzen der folgenden Art:*

$$\begin{bmatrix} 0 & x_1 \\ 1 - p_1 & p_1 \end{bmatrix} \succ \begin{bmatrix} 0 & x_2 \\ 1 - p_2 & p_2 \end{bmatrix} \quad \textit{und} \quad \begin{bmatrix} 0 & x_1 \\ 1 - \lambda p_1 & \lambda p_1 \end{bmatrix} \prec \begin{bmatrix} 0 & x_2 \\ 1 - \lambda p_2 & \lambda p_2 \end{bmatrix},$$

wobei $0 < x_1 < x_2$, $p_1 > p_2$ und λ klein ist. Hier stehen die Wahrscheinlichkeiten für den Gewinn eines positiven Betrages in beiden Lotteriepaaren im selben Verhältnis („common ratio“).

In ihren Untersuchungen verwendeten Kahneman und Tversky (1979) $p_1 = 1$, $p_2 = 0.8$, $\lambda = 0.25$ und $x_1 = 3000$, $x_2 = 4000$. Auch dieser Effekt ist in EU nicht erklärbar. Setzt man $u(0) = 0$, so sind die beobachteten Präferenzen äquivalent zu

$$p_1 u(x_1) > p_2 u(x_2) \quad \text{und} \quad \lambda p_1 u(x_1) < \lambda p_2 u(x_2),$$

was offensichtlich nicht gleichzeitig gelten kann.

Kapitel 2

Choquet-Kapazitäten und das Choquet-Integral

Grundlegend für die im nächsten Kapitel von uns beschriebenen Alternativen zu EU sind die Begriffe Choquet-Kapazität und Choquet-Integration. In diesem Kapitel wollen wir die mathematischen Grundlagen der Choquet-Integrationstheorie darlegen und die für das Spätere wichtigen Eigenschaften von Choquet-Kapazitäten und Choquet-Integration vorstellen. Für eine ausführliche Darstellung der Theorie der Choquet-Integration vergleiche Denneberg (1992).

Zunächst definieren wir Choquet-Kapazitäten und eine wichtige Teilklasse von Choquet-Kapazitäten, die sogenannten verzerrten Wahrscheinkichkeitsmaße. Im nächsten Abschnitt zeigen wir, wie auch für Choquet-Kapazitäten, die im Gegensatz zu Wahrscheinlichkeitsmaßen nicht mehr additiv sind, in sinnvoller Weise ein Integral definiert werden kann. Dies wird uns zum Begriff des Choquet-Integrals führen, das erstmals von Choquet (1953/54) definiert wurde. Im weiteren werden die wichtigsten Eigenschaften des Choquet-Integrals, insbesondere die komonotone Additivität, dargelegt. Des weiteren werden verwandte Integralbegriffe vorgestellt. In Abschnitt 2.3 wird der Begriff des assoziierten Wahrscheinlichkeitsmaßes eingeführt. Es wird gezeigt, daß Choquet-Integration und Integration bezüglich σ-additiver Maße in gewissem Sinne identisch sind, sofern man sich auf komonotone Klassen von Funktionen beschränkt. Der Begriff des assoziierten Wahrscheinlichkeitsmaßes ermöglicht es ferner in einfacher Weise das Produkt zweier Kapazitäten zu definieren. Beschränkt man sich wieder auf geeignete Funktionenklassen (z.B. isotone oder antitone Funktionen), so ist es auch bei Choquet-Kapazitäten möglich, das Integral bezüglich des Produktes zweier Kapazitäten als Doppelintegral zu berechnen. In einem weiteren Abschnitt werden wir noch kurz die Beziehungen zu den sogenannten unscharfen Maßen (engl. fuzzy measures) beschreiben. Der Beweis des Satzes über die Existenz assoziierter Wahrscheinlichkeitsmaße und die Beweise der Fubini-Sätze sind im letzten Abschnitt dieses Kapitels gesammelt.

2.1 Choquet-Kapazitäten

Definition 2.1 (Choquet-Kapazität) *Sei* Ω *eine Menge und* $\mathcal{A}$ *eine Teilmenge der Potenzmenge* $\mathcal{P}(\Omega)$ *mit* $\emptyset, \Omega \in \mathcal{A}$. *Eine Funktion* $\mu : \mathcal{A} \to \mathbb{R}$ *heißt* Choquet-Kapazität *oder* nichtadditives Wahrscheinlichkeitsmaß, *wenn sie die folgenden Eigenschaften besitzt.*

(i) $\mu(\emptyset) = 0$, $\mu(\Omega) = 1$.

(ii) Für alle $A, B \in \mathcal{A}$ *gilt:* $A \subset B \Rightarrow \mu(A) \leq \mu(B)$.

Das Tripel $(\Omega, \mathcal{A}, \mu)$ *bezeichnen wir als* Kapazitätsraum.

In der älteren Literatur werden Kapazitäten als monotone Mengenfunktionen mit gewissen Stetigkeitseigenschaften definiert, siehe z.B. Choquet (1953/54), Dellacherie (1971). In der neueren Literatur zur Entscheidungstheorie ist jedoch die obige Definition gebräuchlich.

Eine Kapazität μ heißt

- *stetig von unten*, falls für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Mengen aus $\mathcal{A}$ mit $A_n \nearrow A \in \mathcal{A}$ gilt $\lim_{n\to\infty} \mu(A_n) = \mu(A)$,
- *stetig von oben*, falls für jede Folge $(A_n)_{n\in\mathbb{N}}$ von Mengen aus $\mathcal{A}$ mit $A_n \searrow A \in \mathcal{A}$ gilt $\lim_{n\to\infty} \mu(A_n) = \mu(A)$,
- *stetig*, falls μ stetig von unten und von oben ist.

Enthält $\mathcal{A}$ mit je zwei Mengen $A, B \in \mathcal{A}$ auch deren Schnitt $A \cap B$ und Vereinigung $A \cap B$, so heißt μ

- *stark subadditiv*, falls für alle $A, B \in \mathcal{A}$ gilt $\mu(A \cup B) + \mu(A \cap B) \leq \mu(A) + \mu(B)$,
- *stark superadditiv*, falls für alle $A, B \in \mathcal{A}$ gilt $\mu(A \cup B) + \mu(A \cap B) \geq \mu(A) + \mu(B)$,

Statt *stark subadditiv* sind auch die Bezeichnungen *konkav, submodular* und *2-alternierend* gebräuchlich. Entsprechend sind statt *stark superadditiv* auch *konvex, supermodular* und *2-monoton* gebräuchlich.

Ist μ eine Kapazität und enthält das Mengensystem $\mathcal{A}$ mit jeder Menge A auch deren Komplement A^c, so ist die Mengenfunktion $\mu^D : \mathcal{A} \to [0, 1]$ mit $\mu^D(A) = 1 - \mu(A^c)$ für alle $A \in \mathcal{A}$ ebenfalls eine Kapazität. Sie heißt die zu μ *duale Kapazität*. Zwischen μ und μ^D bestehen die folgenden Beziehungen (vgl. Denneberg 1992, Proposition 2.3):

- Ist μ stetig von oben, so ist μ^D stetig von unten.

- Ist μ stetig von unten, so ist μ^D stetig von oben.
- Ist μ stark subadditiv, so ist μ^D stark superadditiv.
- Ist μ stark superadditiv, so ist μ^D stark subadditiv.
- Ist μ additiv, so ist $\mu^D = \mu$.
- $(\mu^D)^D = \mu$, d.h. zweimaliges Dualisieren liefert die ursprüngliche Kapazität.

Eine wichtige Klasse von Kapazitäten sind die sogenannten *verzerrten Wahrscheinlichkeitsmaße*.

Definition 2.2 (Verzerrte Wahrscheinlichkeit, Verzerrungsfunktion) *Sei $(\Omega, \mathcal{A}, P)$ ein Wahrscheinlichkeitsraum und sei $q : [0,1] \to [0,1]$ eine monoton wachsende Funktion mit $q(0) = 0$ und $q(1) = 1$. Dann heißt die Mengenfunktion $q \circ P : \mathcal{A} \to [0,1]$* verzerrtes Wahrscheinlichkeitsmaß *oder auch* verzerrte Wahrscheinlichkeit. *Die Funktion q heißt* Verzerrungsfunktion.

Wegen der Eigenschaften von q ist unmittelbar klar, daß $q \circ P$ eine Kapazität ist. Außerdem gelten die folgenden Aussagen (siehe z.B. Denneberg 1992, Example 2.1):

- Ist q stetig von links, so ist $q \circ P$ stetig von unten.
- Ist q stetig von rechts, so ist $q \circ P$ stetig von oben.
- Ist q stetig, so ist $q \circ P$ stetig.
- Ist q konkav, so ist $q \circ P$ stark subadditiv.
- Ist q konvex, so ist $q \circ P$ stark superadditiv.

Die duale Kapazität einer verzerrten Wahrscheinlichkeitsmaßes ist wieder ein verzerrtes Wahrscheinlichkeitsmaß. Es gilt $(q \circ P)^D = q^D \circ P$, wobei $q^D : [0,1] \to [0,1]$ gegeben ist durch $q^D(x) = 1 - q(1-x)$. Die Verzerrungsfunktion q^D heißt die zu q *duale Verzerrungsfunktion*.

Eine Verzerrungsfunktion q heißt *S-förmig*, wenn es ein $c_0 \in (0,1)$ gibt, so daß q in $[0, c_0]$ konvex und in $[c_0, 1]$ konkav ist oder umgekehrt und es genau ein $c_1 \in (0,1)$ gibt mit $q(c_1) = c_1$.

2.2 Das Choquet-Integral

Integration von Funktionen bezüglich einer Kapazität wurde erstmals von Choquet (1953/54) definiert. Seitdem haben sich viele Autoren mit dem Problem der Integration bezüglich eienr Kapazität beschäftigt, siehe insbesondere Dellacherie (1971), Greco (1977), Denneberg (1990,1992).

Definition 2.3 ($\mathcal{A}$-Meßbarkeit von oben) *Eine Funktion $f : \Omega \to \mathbb{R}$ heißt* von oben $\mathcal{A}$-meßbar*, falls $\{\omega \in \Omega \,|\, f(\omega) \geq t\} \in \mathcal{A}$ für jedes $t \in \mathbb{R}$.*

Statt $\{\omega \in \Omega \,|\, f(\omega) \geq t\}$ schreiben wir im folgenden auch kürzer $\{f \geq t\}$. Ebenso schreiben wir statt $\mu(\{\omega \in \Omega \,|\, f(\omega) \geq t\})$ auch einfacher $\mu(f \geq t)$. Wir geben nun die ursprüngliche Definition des Choquet-Integrals von Choquet (1953/54).

Definition 2.4 (Choquet-Integral) *Sei $(\Omega, \mathcal{A}, \mu)$ ein Kapazitätsraum, und sei $f : \Omega \to \mathbb{R}$ eine von oben $\mathcal{A}$-meßbare Funktion. Dann ist das* Choquet-Integral *von f bezüglich μ wie folgt definiert:*

$$\int_\Omega f d\mu \;=\; \int_0^\infty \mu(f \geq t)dt \;+\; \int_{-\infty}^0 [\mu(f \geq t) - 1]dt, \tag{2.1}$$

falls die Summe auf der rechten Seite definiert ist. Ist die Summe endlich, so heißt f Choquet-integrierbar *bezüglich μ.*

Ist μ ein Wahrscheinlichkeitsmaß, so stimmt das Choquet-Integral bezüglich μ mit dem Lebesgue-Integral bezüglich μ überein. Das Choquet-Integral stellt also eine Verallgemeinerung des Lebesgueschen Integrals dar. Diese Verallgemeinerung kann aber mit der gleichen Berechtigung auf verschiedene Weisen geschehen. Bei der soeben gegebenen Form hängt der Wert des Integrals lediglich von den Werten der Kapazität an Mengen der Form $\{f \geq t\}$ ab. Dieses Integral wird daher als *oberes Choquet-Integral* bezeichnet. Will man sich bei der Integration von f nur auf die Werte der Kapazität an Mengen der Form $\{f \leq t\}$ beziehen, so gelangt man zum *unteren Choquet-Integral*, siehe Gilboa (1989):

$$\int_\Omega^D f d\mu \;=\; \int_0^\infty [1 - \mu(f \leq t)]dt \;+\; \int_{-\infty}^0 \mu(f \leq t)dt. \tag{2.2}$$

Das untere Choquet-Integral kann auch dargestellt werden als oberes Choquet-Integral bezüglich der dualen Kapazität, d.h. es besteht folgende Beziehung:

$$\int^D f d\mu \;=\; \int f d\mu^D \;. \tag{2.3}$$

Das untere Choquet-Integral stellt also keinen grundsätzlich neuen Weg der Verallgemeinerung der Lebesgue-Integration dar.

Wir kommen nun zu einigen Eigenschaften des Choquet-Integrals. Zuvor müssen wir aber noch den in diesem Zusammenhang wichtigen Begriff der Komonotonie einführen.

Definition 2.5 (Komonotonie) *Zwei Funktionen $f, g : \Omega \to \mathbb{R}$ heißen* komonoton*, falls $(f(\omega) - f(\omega'))(g(\omega) - g(\omega')) \geq 0$ für alle $\omega, \omega' \in \Omega$. Eine Menge $\mathcal{F}$ von Funktionen von Ω nach $\mathbb{R}$ heißt* komonoton*, falls je zwei Funktionen aus $\mathcal{F}$ komonoton sind.*

Die folgenden drei Bedingungen sind jeweils äquivalent zur Komonotonie von f und g:

- Für alle $\omega, \omega' \in \Omega$ gilt $f(\omega) < f(\omega') \Rightarrow g(\omega) \leq g(\omega')$.
- Es gibt eine lineare Ordnung $\preceq$ auf Ω, so daß f und g isoton bezüglich $\preceq$ sind.
- Die Level-Mengen $\{f \geq t\}, \{g \geq t\}, t \in \mathbb{R}$, sind linear geordnet bezüglich Teilmengenbeziehung.

Das Choquet-Integral besitzt folgende Eigenschaften (siehe z.B. Denneberg 1992, Proposition 5.1):

$$\int 1_A d\mu = \mu(A), \text{ für alle } A \in \mathcal{A}, \tag{2.4}$$

$$\int \lambda f d\mu = \lambda \int f d\mu, \text{ falls } \lambda \geq 0, \tag{2.5}$$

$$\int -f d\mu = -\int f d\mu^D, \tag{2.6}$$

$$f(\omega) \leq g(\omega) \ \forall \omega \in \Omega \Rightarrow \int f d\mu \leq \int g \, d\mu \tag{2.7}$$

$$\int (f+c) d\mu = \int f d\mu + c, \text{ für alle } c \in \mathbb{R}, \tag{2.8}$$

$$\int (f+g) d\mu = \int f d\mu + \int g \, d\mu, \text{ falls } f \text{ und } g \text{ komonoton.} \tag{2.9}$$

Eigenschaft 2.5 bezeichnet man als *positive Homogenität*, Eigenschaft 2.9 als *komonotone Additivität*. Die komonotone Additivität des Choquet-Integrals wurde erstmals von Dellacherie (1971) gezeigt. Er benutzt statt des Begriffs „komonoton“ die Bezeichnung „même tableau de variation“.

Eine wichtige Rolle in der Theorie des Choquet-Integrals spielt der folgende Satz, das sogenannte „Subadditivitäts-Theorem“, da es Additivitätseigenschaften der Kapazität mit denen des Choquet-Integrals in Beziehung setzt. Der Satz wurde vielfach unter verschiedenen Voraussetzungen bewiesen, siehe z.B. Choquet (1953/54), Topsøe (1978), Anger (1977), Huber (1981), Bassanezi und Greco (1984), Buja (1984), Kindler (1986), Schmeidler (1986) und Denneberg (1992).

Satz 2.1 *Sei μ eine Kapazität auf $(\Omega, \mathcal{P}(\Omega))$.*

(i) μ ist stark subadditiv genau dann, wenn für alle μ-integrierbaren Funktionen $f, g : \Omega \to \mathbb{R}$ gilt

$$\int (f+g) d\mu \ \leq \int f d\mu + \int g \, d\mu.$$

(ii) μ ist stark superadditiv genau dann, wenn für alle μ-integrierbaren Funktionen $f, g : \Omega \to \mathbb{R}$ gilt

$$\int (f+g) d\mu \ \geq \int f d\mu + \int g \, d\mu.$$

Beweis: Denneberg (1992), Theorem 6.3 und Korollar 6.4.

Satz 2.2 (Monotone Konvergenz) *Sei* $(\Omega, \mathcal{A})$ *ein Meßraum und* μ *eine Kapazität.* μ *ist genau dann stetig von unten, wenn für jede isotone Folge* $(f_n)_{n\in\mathbb{N}}$ *von nichtnegativen,* $\mathcal{A}$*-meßbaren Funktionen mit* $f_n \nearrow f$ *gilt:*

$$\lim_{n\to\infty} \int f_n d\mu \;=\; \int f d\mu.$$

Beweis: Denneberg (1992), Theorem 8.1.

Eine weitere Möglichkeit der Verallgemeinerung des Integralbegriffs leitet sich aus folgender Beobachtung ab. Sei $f^+ = \max\{f, 0\}$ der Positiv- und $f^- = \max\{-f, 0\}$ der Negativteil der Funktion, dann gilt für das Choquet-Integral

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu^D.$$

Fordert man für eine Verallgemeinerung des Lebesgue-Integrals weiterhin die Gültigkeit der Beziehung $\int f d\mu = \int f^+ d\mu - \int f^- d\mu$ wie beim Lebesgue-Integral, so gelangt man zum *symmetrischen Choquet-Integral* (Šipoš 1979, Denneberg 1992):

$$\int_\Omega^S f d\mu \;=\; \int_0^\infty \mu(f \geq t) dt \;-\; \int_{-\infty}^0 \mu(f \leq t) dt. \tag{2.10}$$

Es gilt also:

$$\int^S f d\mu \;=\; \int f^+ d\mu - \int f^- d\mu.$$

Man bezeichnet das Choquet-Integral gemäß Definition 2.4 daher auch als asymmetrisches Choquet-Integral. Das symmetrische Choquet-Integral besitzt die folgenden Eigenschaften (siehe z.B. Denneberg 1992, Proposition 7.1):

$$\int^S 1_A d\mu = \mu(A), \text{ für alle } A \in \mathcal{A}, \tag{2.11}$$

$$\int^S \lambda f d\mu = \lambda \int^S f d\mu, \text{ für alle } \lambda \in \mathbb{R}, \tag{2.12}$$

$$f(\omega) \leq g(\omega)\ \forall \omega \in \Omega \Rightarrow \int^S f d\mu \leq \int^S g\, d\mu \tag{2.13}$$

$$\int^S (f+g) d\mu = \int^S f d\mu + \int^S g\, d\mu, \text{ falls } f, g \geq 0 \text{ und komonoton.} \tag{2.14}$$

Das symmetrische Choquet-Integral ist also nicht mehr nur *positiv* homogen, sondern sogar homogen (2.12). Auf der anderen Seite gilt aber die komonotone Additivität nur noch, falls beide Funktionen nichtnegativ sind.

Definition 2.4 erlaubt die Integration aller von oben $\mathcal{A}$-meßbaren Funktionen. Wie Greco (1981) gezeigt hat, ist jedoch die Integration einer weit größeren Klasse von Funktionen möglich. Grundlegend ist dabei der Begriff der *quasi-*$\mathcal{T}$*-Meßbarkeit.*

Definition 2.6 (Quasi-$\mathcal{T}$-Meßbarkeit) *Sei $\mathcal{T}$ eine Teilmenge von $\mathcal{A}$, so daß $\emptyset, \Omega \in \mathcal{T}$. Eine Funktion $f : \Omega \to \mathbb{R}$ heißt* quasi-$\mathcal{T}$-meßbar *wenn für alle $\alpha \in \mathbb{R}$ und alle $\epsilon > 0$ eine Menge $T_{\alpha,\epsilon} \in \mathcal{T}$ existiert, so daß*

$$\Big\{f \geq \alpha\Big\} \subset T_{\alpha,\epsilon} \subset \Big\{f \geq \alpha - \epsilon\Big\}. \tag{2.15}$$

Ist f von oben $\mathcal{T}$-meßbar, so ist f insbesondere auch quasi-$\mathcal{T}$-meßbar. In diesem Fall kann ja $T_{\alpha,\epsilon} = \{f \geq \alpha\}$ gewählt werden. Ist $\mathcal{T}$ eine monotone Klasse, so ist jede quasi-$\mathcal{T}$-meßbare Funktion auch von oben $\mathcal{T}$-meßbar. Ist $\mathcal{T}$ eine σ-Algebra, so ist jede quasi-$\mathcal{T}$-meßbare Funktion auch meßbar im herkömmlichen Sinne. Wir bezeichnen mit $\mathcal{Q}(\mathcal{T})$ die Menge aller quasi-$\mathcal{T}$-meßbaren Funktionen und mit $\mathcal{I}(\mathcal{T})$ die Menge aller Indikatorfunktionen von Mengen aus $\mathcal{T}$. Offensichtlich gilt $\mathcal{I}(\mathcal{T}) \subset \mathcal{Q}(\mathcal{T})$. Liegt umgekehrt eine Indikatorfunktion 1_T in $\mathcal{Q}(\mathcal{T})$, so ist $T \in \mathcal{T}$.

Die Fortsetzung des Choquet-Integrals auf die Klasse aller quasi-$\mathcal{T}$-meßbaren Funktionen beruht auf dem folgenden Satz.

Satz 2.3 (Greco 1981) *Seien μ_1, μ_2 Kapazitäten auf $(\Omega, \mathcal{P}(\Omega))$, und sei $\mathcal{T} \subset \mathcal{P}(\Omega)$. Dann gilt*

$$\int f d\mu_1 \;=\; \int f d\mu_2 \quad \text{für alle } f \in \mathcal{Q}(\mathcal{T})$$

genau dann,wenn

$$\mu_1(T) \;=\; \mu_2(T) \text{ für alle } T \in \mathcal{T}.$$

Dies bedeutet nun aber, daß der Wert des Integrals für eine quasi-$\mathcal{T}$-meßbare Funktion nicht von den Werten der Kapazität auf $\mathcal{P}(\Omega) \setminus \mathcal{T}$ abhängt. Auf dieser Basis läßt sich nun das Choquet-Integral auch für Funktionen in $\mathcal{Q}(\mathcal{T})$ definieren.

Definition 2.7 (Choquet-Greco-Integral) *Sei $(\Omega, \mathcal{A}, \mu)$ ein Kapazitätsraum, $\mathcal{T} \subset \mathcal{A}$ und $f \in \mathcal{Q}(\mathcal{T})$. Dann ist das* Choquet-Greco-Integral *von f bezüglich μ wie folgt definiert:*

$$\int f d\mu \;=\; \int_0^\infty \tilde{\mu}(f \geq t)dt \;+\; \int_{-\infty}^0 [\tilde{\mu}(f \geq t) - 1]dt, \tag{2.16}$$

falls die Summe der Integrale definiert ist. Hierbei ist $\tilde{\mu}$ eine Fortsetzung von μ auf $\mathcal{P}(\Omega)$.

Wegen Satz 2.3 ist das Choquet-Greco-Integral wohldefiniert. Entsprechend können natürlich auch ein unteres Choquet-Greco-Integral und ein symmetrisches Choquet-Greco-Integral für quasi-$\mathcal{T}$-meßbare Funktionen definiert werden. Die Eigenschaften des Choquet-Greco-Integrals stimmen mit denjenigen des entsprechenden Choquet-Integrals überein.

Wenn wir im weiteren vom Choquet-Integral sprechen, soll damit immer das asymmetrische, obere Choquet-Greco-Integral gemeint sein.

Wir betrachten nun noch kurz den Spezialfall, daß Ω endlich, und $\mathcal{A}$ die Potenzmenge von Ω ist. Sei $f_i = f(\omega_i)$, $i = 1, \ldots, n$ und $f_0 = 0$. Weiterhin gelte o.B.d.A $f_1 \leq f_2 \leq \ldots f_n$. In diesem Fall kann das Choquet-Integral wie folgt berechnet werden:

$$\begin{aligned}\int f d\mu &= \sum_{i=1}^{n} (f_i - f_{i-1})\, \mu(\{\omega_i, \ldots, \omega_n\}) \\ &= \sum_{i=1}^{n} f_i [\mu(\{\omega_i, \ldots, \omega_n\}) - \mu(\{\omega_{i+1}, \ldots, \omega_n\})]\,.\end{aligned}$$

Im weiteren werden wir auch oft dem Fall begegnen, daß μ ein verzerrtes Wahrscheinlichkeitsmaß ist. Sind f_i, $i = 0, \ldots n$, wie oben erklärt, und sind $p_i = P(\{\omega_i\})$, $i = 1, \ldots n$, so gilt

$$\begin{aligned}\int f\, d(q \circ P) &= \sum_{i=1}^{n} (f_i - f_{i-1})\, q(\sum_{j=i}^{n} p_j) \\ &= \sum_{i=1}^{n} f_i \left[q(\sum_{j=i}^{n} p_j) - q(\sum_{j=i+1}^{n} p_j) \right]\,.\end{aligned}$$

Sei μ eine Kapazität auf $(\Omega, \mathcal{A})$ und $X : \Omega \to \mathbb{R}$ eine quasi-$\mathcal{A}$-meßbare Funktion. In Analogie zum Erwartungswert einer reellen Zufallsvariablen schreiben wir für das Choquet-Integral von X bezüglich μ im folgenden auch gelegentlich $E_\mu(X)$ und bezeichnen $E_\mu(X)$ als *subjektiven Erwartungswert* von X.

2.3 Existenz von assoziierten Wahrscheinlichkeitsmaßen

Wir verdeutlichen die Grundidee dieses Abschnitts zunächst an folgender Situation. Wir betrachten einen endlichen Raum Ω und eine lineare Ordnung $\leq$ auf Ω, sowie die Menge $\mathcal{F}_1$ aller bezüglich $\leq$ isotonen Funktionen von Ω nach $\mathbb{R}$. Wir nehmen nun o.B.d.A. an, daß $\Omega = \{1, \ldots n\}$ und $\leq$ die gewöhnliche Ordnung ist. Dann kann das Choquet-Integral einer Funktion $f \in \mathcal{F}_1$ dargestellt werden als

$$\int f d\mu = \sum_{i=1}^{n} f(i) [\mu(\{i, \ldots n\}) - \mu(\{i+1, \ldots, n\})].$$

Setzt man

$$p_i = \mu(\{i, \ldots n\}) - \mu(\{i+1, \ldots, n\}),\ i = 1, \ldots n,$$

so wird dadurch ein additives Maß auf Ω definiert und man kann das Choquet-Integral jeder Funktion $f \in \mathcal{F}_1$ schreiben als

$$\int f d\mu = \sum_{i=1}^{n} f(i) p_i.$$

Betrachtet man also nur isotone Funktionen, so ist das Choquet-Integral als Integral bezüglich eines *additiven* Maßes darstellbar.

In diesem Abschnitt werden wir zeigen, daß dieses Ergebnis unter geringen zusätzlichen Voraussetzungen sogar für beliebige Räume Ω gilt. Die Bedeutung dieses Ergebnisses liegt darin begründet, daß man bei Beschränkung auf eine Menge von komonotonen Funktionen (die ja isoton bezüglich einer gemeinsamen Ordnung sind) die Choquet-Integration als ganz gewöhnliche Integration bezüglich eines σ-additiven Maßes ansehen kann. Daher sind in diesem Fall die Ergebnisse der Integrationstheorie bezüglich σ-additiver Maße auf die Choquet-Integrationstheorie übertragbar. Beispiele für solche Anwendungen werden wir in den Abschnitten 2.4, 4.3, 4.4 und 8.2.1 kennenlernen.

Im folgenden sei Ω eine Menge, $\leq$ eine schwache Ordnung auf Ω, $\mathcal{U}$ das System aller oberen[1] Mengen und $\mathcal{S}$ die von $\mathcal{U}$ erzeugte Algebra sowie μ eine Kapazität auf $(\Omega, \mathcal{U})$. Scarsini (1992) hat gezeigt, daß es ein endlich additives Wahrscheinlichkeitsmaß m auf $(\Omega, \mathcal{S})$ gibt, so daß

$$\int f d\mu = \int f dm \quad \text{für alle isotonen Funktionen } f,$$

wobei beide Integrale als Choquet-Integrale zu verstehen sind. Diese Aussage ergibt sich ähnlich wie das einführende Beispiel.

Wir untersuchen nun die Frage, unter welchen Bedingungen dieses endlich additive W-Maß m zu einem σ-additivem Wahrscheinlichkeitsmaß P auf der von den oberen Mengen erzeugten σ-Algebra $\mathcal{A}$ fortgesetzt werden kann, so daß gilt

$$\int f d\mu = \int f dP \quad \text{für alle isotonen Funktionen } f\ . \tag{2.17}$$

Eine zu (2.17) äquivalente Bedingung ist, daß μ und P auf allen oberen Mengen übereinstimmen:

Lemma 2.1 *Sei $(\Omega, \leq)$ ein schwach geordneter Raum, $\mathcal{U}$ das System aller oberen Mengen und $\mathcal{F}_1$ die Menge aller isotonen Funktionen von Ω nach $\mathbb{R}$. Die folgenden Bedingungen sind äquivalent.*

(i) $\int f d\mu = \int f dP$ *für alle* $f \in \mathcal{F}_1$.

[1]Ist $(\Omega, \leq)$ ein mit einer Präordnung versehener Raum, so heißt eine Menge $U \subset \Omega$ *obere Menge*, wenn für alle $x, y \in \Omega$ gilt $x \in U, x \leq y \Rightarrow y \in U$.

(ii) $\mu(U) = P(U)$ *für alle* $U \in \mathcal{U}$.

Beweis: (i) $\Rightarrow$ (ii) erhält man, indem man für f Indikatorfunktionen von oberen Mengen einsetzt.
(ii) $\Rightarrow$ (i): Ist f isoton, so sind die Level-Mengen $\{f \geq t\} \in \mathcal{U}$ für alle $t \in \mathbb{R}$. Aus

$$\begin{aligned}
\int f d\mu &= \int_0^\infty \mu(f \geq t)dt + \int_{-\infty}^0 [\mu(f \geq t) - 1]dt \\
&= \int_0^\infty P(f \geq t)dt + \int_{-\infty}^0 [P(f \geq t) - 1]dt \\
&= \int f dP
\end{aligned}$$

folgt dann die Behauptung. ■

Nach dem vorangehenden Lemma genügt es also, ein Wahrscheinlichkeitsmaß zu finden, das mit μ auf allen oberen Mengen übereinstimmt. Ein notwendige Bedingung für die Existenz eines solchen Maßes ist offensichtlich, daß μ stetig auf den oberen Mengen ist, d.h. daß für jede auf- oder absteigende Folge $(U_n)_{n\in\mathbb{N}}$ von oberen Mengen gelten muß

$$\lim_{n\to\infty} \mu(U_n) = \mu\Big(\lim_{n\to\infty} U_n\Big) \ .$$

Leider ist diese Bedingung nicht hinreichend, wie das folgende Beispiel zeigt, das eine leichte Modifikation eines Beispiels aus Halmos (1974, S. 40) darstellt.

Beispiel 2.1 Sei $\Omega = \mathbb{Q} \cap [0,1]$, $\leq$ die gewöhnliche Ordnung auf $\mathbb{Q}$, und μ eine Kapazität mit $\mu((a,1]) = \mu([a,1]) = 1 - a$ für alle $a \in \mathbb{R} \cap [0,1]$. Dann ist μ offensichtlich stetig auf oberen Mengen. Es gibt aber kein σ-additives W-Maß, das mit μ auf $\mathcal{U}$ übereinstimmt. Gäbe es nämlich so ein Maß P, so wäre

$$P\Big(\{a\}\Big) = P\Big([a,1]\Big) - P\Big((a,1]\Big) = \mu\Big([a,1]\Big) - \mu\Big((a,1]\Big) = 0$$

für jedes $a \in \Omega$. Da Ω abzählbar ist, würde daraus folgen

$$1 \;=\; P\Big(\Omega\Big) \;=\; P\Big(\bigcup_{a\in\Omega}\{a\}\Big) \;=\; \sum_{a\in\Omega} P\Big(\{a\}\Big) \;=\; 0.$$

Um hinreichende Bedingungen für die Existenz eines W-Maßes mit den gewünschten Eigenschaften zu finden, müssen wir eine weitere Voraussetzung an die Ordnungsstruktur des Raumes stellen.

Definition 2.8 (Vollständigkeit bezüglich monotoner Folgen) *Sei* $(\Omega, \leq)$ *ein linear geordneter Raum.* $(\Omega, \leq)$ *heißt* vollständig bezüglich monotoner Folgen, *wenn jede monoton wachsende Folge ein Supremum und jede monoton fallende*

Folge ein Infimum besitzt. $(\Omega, \leq)$ *heißt* vollständig bezüglich beschränkter monotoner Folgen*, wenn jede monoton wachsende, beschränkte Folge ein Supremum und jede monoton fallende, beschränkte Folge ein Infimum besitzt. Ein schwach geordneter Raum* $(\Omega, \leq)$ *heißt* vollständig bezüglich (beschränkter) monotoner Folgen*, wenn der linear geordnete Raum* $(\Omega^*, \leq^*)$ *der Äquivalenzklassen*[2] *diese Eigenschaft besitzt.*

Satz 2.4 besagt, daß Vollständigkeit bezüglich beschränkter, monotoner Folgen neben der Stetigkeit der Kapazität auf oberen Mengen hinreichend für die Existenz des von uns gesuchten W-Maßes ist. Der Beweis von Satz 2.4 findet sich in Abschnitt 2.6.

Satz 2.4 *Sei* $(\Omega, \leq)$ *ein schwach geordneter, bezüglich beschränkter, monotoner Folgen vollständiger Raum,* $\mathcal{U}$ *das System aller oberen Mengen und* $\mathcal{A}$ *die von* $\mathcal{U}$ *erzeugte* σ*-Algebra. Sei weiter* μ *eine Kapazität auf* $(\Omega, \mathcal{U})$*, die stetig auf oberen Mengen ist. Dann gibt es genau ein* σ*-additives Wahrscheinlichkeitsmaß* P *auf* $(\Omega, \mathcal{A})$*, das mit* μ *auf allen oberen Mengen übereinstimmt.*

Definition 2.9 (Assoziiertes Wahrscheinlichkeitsmaß) *In der Situation von Satz 2.4 heißt das eindeutig bestimmte Wahrscheinlichkeitsmaß* P*, das zu* μ *und* $\leq$ assoziierte Wahrscheinlichkeitsmaß*. Wir schreiben dafür auch* $P = P_{\mu,\leq}$.

Bemerkung 1: Die σ-Algebra $\mathcal{A}$ ist die kleinste σ-Algebra bezüglich der alle isotonen Funktionen meßbar sind.

Bemerkung 2: Ist $\Omega = \mathbb{R}$ und $\leq$ die gewöhnliche Ordnung auf $\mathbb{R}$, so ist $\mathcal{A}$ die Borelsche σ-Algebra.

Korollar 2.1 *Für* P *wie in Satz 2.4 gelten die folgenden Aussagen.*

$$\begin{aligned} \int f d\mu &= \int f dP \qquad \textit{für alle isotonen Funktionen } f. \\ \int^D f d\mu &= \int f dP \qquad \textit{für alle antitonen Funktionen } f. \end{aligned}$$

Beweis: Die erste Gleichheit folgt unmittelbar aus Lemma 2.1. Die zweite Gleichheit folgt aus der Tatsache, daß für eine antitone Funktion f die Level-Mengen $\{f \leq \alpha\}$, $\alpha \in \Omega$, obere Mengen sind. ∎

Korollar 2.2 *Sei* $(\Omega, \leq)$ *wie in Satz 2.4,* $\mathcal{L}$ *das System aller unteren Mengen und* $\mathcal{A}$ *die von* $\mathcal{L}$ *erzeugte* σ*-Algebra. Sei* μ *eine Kapazität, die stetig auf den unteren Mengen ist. Dann gibt es genau ein* σ*-additives Wahrscheinlichkeitsmaß* P *auf*

[2] vgl. Anhang A.1

$(\Omega, \mathcal{A})$, das mit μ auf allen unteren Mengen übereinstimmt. Für dieses W-Maß gelten die folgenden Gleichheiten:

$$\begin{aligned} \int f d\mu &= \int f dP \qquad \textit{für alle antitonen } f \\ \int^D f d\mu &= \int f dP \qquad \textit{für alle isotonen } f \end{aligned}$$

Beweis: Definiert man $\leq^D$ auf Ω durch $x \leq^D y \iff y \leq x$, so folgt das Korollar, indem man Satz 2.4 auf $(\Omega, \leq^D)$ anwendet. ∎

Satz 2.4 beruht auf der Tatsache, daß die oberen Level-Mengen der betrachteten Funktionen bezüglich Mengeninklusion linear geordnet sind. Man kann nun auch dies zum Ausgangspunkt nehmen und von einer Klasse von Funktionen ausgehen, deren obere Level-Mengen linear geordnet sind. Nach der Bemerkung im Anschluß an Definition 2.5 ist dies gleichwertig mit der Komonotonie der betrachteten Funktionen. Wir betrachten also im folgenden einen Raum $(\Omega, \mathcal{A})$ und eine komonotone Menge von Funktionen.

Lemma 2.2 *Ist $\mathcal{F}$ eine komonotone Menge von Funktionen auf Ω, so ist die durch*

$$x \leq_{\mathcal{F}} y \iff f(x) \leq f(y) \textit{ für alle } f \in \mathcal{F}$$

definierte Relation $\leq_{\mathcal{F}}$ eine schwache Ordnung auf Ω, und $\mathcal{F}$ ist eine Teilmenge der isotonen Funktionen bezüglich $\leq_{\mathcal{F}}$.

Beweis: Die Transitivität ist klar. Es bleibt die Vollständigkeit zu zeigen. Angenommen $\leq_{\mathcal{F}}$ wäre nicht vollständig. Dann gibt es $x, y \in \Omega$, für die weder $x \leq_{\mathcal{F}} y$ noch $y \leq_{\mathcal{F}} x$ gilt. Also existieren zwei Funktionen $f, g \in \mathcal{F}$, so daß $f(x) > f(y)$ und $g(y) > g(x)$ ist. Dann sind aber f und g nicht komonoton; Widerspruch. Also ist $\prec_{\mathcal{F}}$ auch vollständig. ∎

Haben wir also eine Menge $\mathcal{F}$ von komonotonen Funktionen vorliegen und suchen wir nach einem Wahrscheinlichkeitsmaß, das (2.17) für alle $f \in \mathcal{F}$ erfüllt, so müssen wir nur prüfen, ob die von $\mathcal{F}$ induzierte schwache Ordnung die Voraussetzungen von Satz 2.4 erfüllt. Ist dies der Fall, so schreiben wir für das assoziierte W-Maß statt $P_{\mu,\leq_{\mathcal{F}}}$ auch einfacher $P_{\mu,\mathcal{F}}$.

2.4 Produkte von Kapazitäten

In diesem Abschnitt wird eine Möglichkeit vorgeschlagen, analog zum Produkt zweier σ-additiver Maße das Produkt zweier Kapazitäten zu definieren.

Definition 2.10 *Eine Funktion $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ heißt* komponentenweise komonoton*, wenn sowohl die Menge der Funktionen $f(\omega_1, \cdot)$, $\omega_1 \in \Omega_1$, als auch die Menge der Funktionen $f(\cdot, \omega_2)$, $\omega_2 \in \Omega_2$, komonoton sind.*

Eine Menge $A \subset \Omega_1 \times \Omega_2$ heißt komponentenweise komonoton*, wenn ihre Indikatorfunktion komponentenweise komonoton ist.*

Für festes $\omega_1 \in \Omega_1$ und $A \subset \Omega_1 \times \Omega_2$ heißt $A^{\omega_1} = \{\omega_2 \in \Omega_2 \,|\, (\omega_1, \omega_2) \in A\}$. der ω_1-Schnitt von A. Entsprechend definiert man den ω_2-Schnitt von A. Sei $f : \Omega_1 \times \Omega_2 \to \Omega'$ eine Funktion von $\Omega_1 \times \Omega_2$ nach einer Menge Ω'. Die Funktionen f_{ω_1}, $\omega_1 \in \Omega_1$, und f_{ω_2}, $\omega_2 \in \Omega_2$, sind definiert durch

$$\begin{aligned} f_{\omega_1} : \Omega_2 \to \Omega', &\quad \omega_2 \mapsto f(\omega_1, \omega_2), \\ f_{\omega_2} : \Omega_1 \to \Omega', &\quad \omega_1 \mapsto f(\omega_1, \omega_2). \end{aligned}$$

Die Funktion f_{ω_1} heißt der ω_1-Schnitt von f, entsprechend für ω_2.

Beispiele für komponentenweise komonotone Funktionen sind alle bezüglich der komponentenweisen partiellen Ordnung auf $\mathbb{R}^2$ monoton wachsenden Funktionen. Alle Rechtecke $A_1 \times A_2$, $A_1 \subset \Omega_1$, $A_2 \subset \Omega_2$ sind komponentenweise komonoton, ebenso alle oberen Mengen.

Wir definieren eine Produktkapazität μ durch die Gleichung

$$\begin{aligned} \mu(A) &= \int_{\Omega_1} \int_{\Omega_2} 1_A \, d\mu_2 \, d\mu_1 \\ &= \int_{\Omega_1} \mu_2(A^{\omega_1}) \, d\mu_1, \end{aligned} \tag{2.18}$$

deren Gültigkeit wir jedoch nur für komponentenweise komonotone Mengen A fordern. Offensichtlich ist $\mu(\emptyset) = 0$ und $\mu(\Omega_1 \times \Omega_2) = 1$. Wegen der Monotonie des Choquet-Integrals ist auch klar, daß μ isoton bezüglich Mengeninklusion (eingeschränkt auf komponentenweise komonotone Mengen) ist. Wir definieren daher:

Definition 2.11 (Produktkapazität) *Für $i = 1, 2$ sei μ_i eine Choquet-Kapazität auf einem Meßraum $(\Omega_i, \mathcal{A}_i)$. Dann heißt jede Kapazität μ auf $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$, für die (2.18) für alle komponentenweise komonotonen Mengen $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ gilt,* Produktkapazität *von μ_1 und μ_2. Wir schreiben $\mu \in \mu_1 \otimes \mu_2$.*

Durch obige Definition ist eine Produktkapazität auf allen komponentenweise komonotonen Mengen eindeutig festgelegt. Produktmaße von σ-additiven Maßen sind durch ihre Werte auf den Rechtecken bereits auf der ganzen Produkt-σ-Algebra eindeutig bestimmt. Dies ist bei Produktkapazitäten jedoch nicht der Fall.

Ein Beispiel für eine Produktkapazität ist z.B. die durch

$$\mu(A) = \int_{\Omega_1} \int_{\Omega_2} 1_A \, d\mu_2 \, d\mu_1 \quad \text{für alle } A \in \mathcal{A}_1 \otimes \mathcal{A}_2$$

definierte Kapazität.

Lemma 2.3 *Für $i = 1, 2$ sei μ_i eine Choquet-Kapazität auf einem Meßraum $(\Omega_i, \mathcal{A}_i)$ und μ eine Produktkapazität auf $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$. Dann ist μ^D eine Produktkapazität von μ_1^D und μ_2^D.*

Beweis: Sei $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ eine komponentenweise komonotone Menge. Dann ist auch A^c komponentenweise komonoton und es gilt

$$\begin{aligned}\mu^D(A) &= 1 - \mu(A^c) \\ &= 1 - \int_{\Omega_1} \int_{\Omega_2} 1_{A^c} d\mu_2 d\mu_1.\end{aligned}$$

Wegen (2.8) und (2.6) ist

$$\begin{aligned}1 - \int_{\Omega_1} \int_{\Omega_2} 1_{A^c} d\mu_2 d\mu_1 &= \int_{\Omega_1} \int_{\Omega_2} (1 - 1_{A^c})\, d\mu_2^D d\mu_1^D \\ &= \int_{\Omega_1} \int_{\Omega_2} 1_A d\mu_2^D d\mu_1^D.\end{aligned}$$

Also ist μ eine Produktkapazität von μ_1^D und μ_2^D.

Lemma 2.4 *Für $i = 1, 2$ sei μ_i eine Choquet-Kapazität auf einem Meßraum $(\Omega_i, \mathcal{A}_i)$ und μ eine Produktkapazität auf $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$. Dann gilt für alle Rechtecke $A_1 \times A_2 \in \mathcal{A}_1 \times \mathcal{A}_2$:*

$$\mu(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2).$$

Beweis: Für Indikatorfunktionen von Rechtecken gilt

$$1_{A_1 \times A_2}(\omega_1, \omega_2) = 1_{A_1}(\omega_1) 1_{A_2}(\omega_2).$$

Da außerdem jedes Rechteck komponentenweise komonoton ist, gilt

$$\begin{aligned}\mu(A_1 \times A_2) &= \int_{\Omega_1} \int_{\Omega_2} 1_{A_1}(\omega_1) 1_{A_2}(\omega_2)\; d\mu_2 d\mu_1 \\ &= \int_{\Omega_1} 1_{A_1}(\omega_1) \int_{\Omega_2} 1_{A_2}(\omega_2)\; d\mu_2 d\mu_1 \\ &= \int_{\Omega_1} 1_{A_1}(\omega_1) \mu_2(A_2) d\mu_1 \\ &= \mu_1(A_1) \cdot \mu_2(A_2).\end{aligned}$$

■

Lemma 2.5 *Für $i = 1, 2$ sei μ_i eine Choquet-Kapazität auf einem Meßraum $(\Omega_i, \mathcal{A}_i)$ und μ eine Produktkapazität auf $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$. Sind μ_1 und μ_2 stetig, so ist μ stetig auf komponentenweise komonotonen Mengen.*

Beweis: Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge von komponentenweise komonotonen Mengen, für die $A_n \nearrow A$ gilt. Dann ist auch für jedes $\omega_1 \in \Omega_1$ die Folge $(A_n^{\omega_1})_{n\in\mathbb{N}}$ aufsteigend gegen A^{ω_1}. Daher gilt $\mu_2(A_n^{\omega_1}) \nearrow \mu_2(A^{\omega_1})$. Nach Satz 2.2 ist dann aber auch

$$\int_{\Omega_1} \mu_2(A_n^{\omega_1}) d\mu_1 \nearrow \int_{\Omega_1} \mu_2(A^{\omega_1}) d\mu_1$$

und somit

$$\lim_{n\to\infty} \mu(A_n) = \mu(A),$$

wie behauptet. ■

Voraussetzung (P): Für $i = 1, 2$ seien $(\Omega_i, \leq_i)$ schwach geordnete Räume, die vollständig bezüglich beschränkter, monotoner Folgen seien, $\mathcal{A}_i$ die von den oberen Mengen bezüglich $\leq_i$ erzeugten σ-Algebren und μ_i stetige Kapazitäten auf $\mathcal{A}_i$. Weiter sei $(\Omega, \mathcal{A}) = (\Omega_1 \times \Omega_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$, μ eine Produktkapazität auf $(\Omega, \mathcal{A})$ und $\leq$ die von $\leq_1$ und $\leq_2$ auf Ω induzierte partielle Ordnung.

Satz 2.5 *Unter der Voraussetzung (P) gilt für jede nichtnegative, bezüglich $\leq$ isotone, $\mathcal{A}$-meßbare Funktion $f : \Omega \to \mathbb{R}$:*

$$\int_{\Omega_1\times\Omega_2} f \; d(\mu_1 \otimes \mu_2) = \int_{\Omega_1} \left[\int_{\Omega_2} f d\mu_2 \right] d\mu_1 = \int_{\Omega_2} \left[\int_{\Omega_1} f d\mu_1 \right] d\mu_2.$$

Beweis: Abschnitt 2.6.

Korollar 2.3 *Unter der Voraussetzung (P) gilt für jede nichtnegative, bezüglich $\leq$ antitone, $\mathcal{A}$-meßbare Funktion $f : \Omega \to \mathbb{R}$:*

$$\int_{\Omega_1\times\Omega_2} f \; d(\mu_1 \otimes \mu_2) = \int_{\Omega_1} \left[\int_{\Omega_2} f d\mu_2 \right] d\mu_1 = \int_{\Omega_2} \left[\int_{\Omega_1} f d\mu_1 \right] d\mu_2.$$

Beweis: Wir betrachten auf Ω_i die durch $x \leq_i^D y \iff y \leq_i x$ definierten schwachen Ordnungen $\leq_i^D$, $i = 1, 2$. Wir zeigen, daß auch diese schwachen Ordnungen die Voraussetzung (P) erfüllen.

Zunächst sind $\leq_i$, $i = 1, 2$, natürlich auch schwache Ordnungen. $(\Omega_i, \leq_i)$ ist auch vollständig bezüglich beschränkter, monotoner Folgen. Eine aufsteigende Folge bezüglich $\leq_1^D$ ist absteigend bezüglich $\leq_1$ und hat daher ein Infimum (bezüglich $\leq_1$). Dieses Infimum ist aber gerade das Supremum der Folge bezüglich $\leq_1^D$. Für aufsteigende Folgen ist die Situation völlig analog.

Da die oberen Mengen bezüglich $\leq_1^D$ gerade die unteren Mengen bezüglich $\leq_1$ sind und diese wiederum die Komplemente der oberen Mengen bezüglich $\leq_1$, erzeugen die oberen Mengen bezüglich $\leq_1^D$ und die oberen Mengen bezüglich $\leq_1$ dieselbe σ-Algebra.

Ist f eine bezüglich $\leq$ antitone Funktion, so ist f isoton bezüglich $\leq^D$. Somit sind die Voraussetzungen von Satz 2.5 erfüllt, und die Behauptung folgt. ■

Satz 2.6 (Satz von Fubini für Kapazitäten) *Die Voraussetzung (P) sei erfüllt, und $f : \Omega \to \mathbb{R}$ sei eine $\mathcal{A}$-meßbare, μ-integrierbare Funktion. Ist f bezüglich $\leq$ isoton oder antiton, so gilt:*

$$\int_{\Omega_1 \times \Omega_2} f \; d(\mu_1 \otimes \mu_2) = \int_{\Omega_1} \left[\int_{\Omega_2} f d\mu_2 \right] d\mu_1 = \int_{\Omega_2} \left[\int_{\Omega_1} f d\mu_1 \right] d\mu_2.$$

Beweis: Abschnitt 2.6.

Wir geben nun noch einige Beispiele an, in denen die Vorausetzungen des Satzes von Fubini erfüllt sind.

Beispiel 2.2 *Für $i = 1, 2$ seien Ω_i endliche Räume, $\mathcal{A}_i = \mathcal{P}(\Omega_i)$, μ_i Kapazitäten auf $(\Omega_i, \mathcal{A}_i)$ und f eine beliebige komponentenweise komonotone Funktion. Dann gibt es lineare Ordnungen $\leq_i$, $i = 1, 2$, so daß die Vorausetzungen von Satz 2.6 erfüllt sind.*

Beweis: Ist f komponentenweise komonoton, so gibt es lineare Ordnungen $\leq_1$ und $\leq_2$ auf Ω_1 und Ω_2, so daß f bezüglich der von diesen Ordnungen induzierten partiellen Ordnung auf Ω isoton ist. Wegen der Endlichkeit sind die geordneten Räume vollständig bezüglich beschränkter, monotoner Folgen. Weiter überlegt man sich leicht, daß für $i = 1, 2$ die von den oberen Mengen bezüglich $\leq_i$ erzeugte σ-Algebra alle einelementigen Teilmengen enthält und somit $\mathcal{P}(\Omega_i)$ sein muß. Die Produkt-σ-Algebra ist dann ebenfalls die Potenzmenge von Ω. Weiter ist jede Funktion bezüglich der Potenzmenge meßbar. Außerdem ist natürlich auch jede Kapazität auf einer endlichen Menge stetig, so daß tatsächlich die Voraussetzungen erfüllt sind. ■

Beispiel 2.3 *Seien $\Omega_i = \mathbb{R}$, $\leq_i$ die gewöhnliche Ordnung auf $\mathbb{R}$, und $\mathcal{A}_i = \mathcal{B}$ die Borelsche σ-Algebra auf $\mathbb{R}$, μ_i stetige Kapazitäten auf $(\mathbb{R}, \mathcal{B})$ und $f : \mathbb{R}^2 \to \mathbb{R}$ eine isotone, $\mathcal{B}^2$-meßbare, $\mu_1 \otimes \mu_2$-integrierbare Funktion. Dann sind die Voraussetzungen des Satzes von Fubini erfüllt.*

Beweis: $\mathbb{R}$ versehen mit der gewöhnlichen Ordnung ist vollständig bezüglich beschränkter, monotoner Folgen. Die oberen Mengen in $\mathbb{R}$ sind genau die Intervalle der Form (a, ∞), $[a, \infty)$, $a \in \mathbb{R}$ (sowie natürlich $\emptyset$ und $\mathbb{R}$). Daher ist jede Menge der Form $(a, b]$, $a < b$ in der von den oberen Mengen erzeugten σ-Algebra. Das System dieser Mengen ist aber ein Erzeuger der Borelschen σ-Algebra. Also treffen die Vorausetzungen des Satzes von Fubini zu. ■

Sind die Kapazitäten μ_1 und μ_2 nicht stetig, so ist der Satz von Fubini i.a. nicht mehr gültig. Wir geben dazu ein Beispiel.

Beispiel 2.4 (Fubini nicht gültig, falls μ_i nicht stetig) *Seien* $(\Omega_i, \mathcal{A}_i) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$, $i = 1, 2$. *Dann ist* $\Omega = \mathbb{N}^2$ *und* $\mathcal{A} = \mathcal{P}(\mathbb{N}^2)$. *Sei* $\preceq_1 = \leq$ *die gewöhnliche Ordnung auf* $\mathbb{N}$ *und* $\preceq_2 = \leq^D$ *die duale Ordnung, d.h. es ist* $a \preceq_2 b \iff a \geq b$. *Seien weiter* $\mu_1 = \mu_2$ *definiert durch*

$$\mu_1(A) = \begin{cases} 1, & \textit{falls } A \textit{ unendlich,} \\ 0, & \textit{sonst.} \end{cases}$$

Dann sind μ_1 *und* μ_2 *Kapazitäten (sogar endlich additiv). Die Menge* $A = \{(a, b) \,|\, a \geq b\}$ *ist eine obere Menge. Ist nämlich* $(a, b) \in A$ *und* $(a, b) \preceq (c, d)$, *so ist* $a \leq c$ *und* $b \geq d$. *Wegen* $(a, b) \in A$ *ist auch* $a \geq b$. *Daraus folgt sofort, daß* $c \geq d$ *ist, d.h.* $(c, d) \in A$. *Nun ist aber*

$$\begin{aligned} \int_{\mathbb{N}} \int_{\mathbb{N}} 1_A(\omega_1, \omega_2)\, d\mu_2(\omega_2) d\mu_1(\omega_1) &= \int_{\mathbb{N}} \mu_2(\{\omega_2 \,|\, \omega_2 \leq \omega_1\})\, d\mu_1(\omega_1) \\ &= \int_{\mathbb{N}} 0\, d\mu_1 \;=\; 0. \end{aligned}$$

Auf der anderen Seite ist

$$\begin{aligned} \int_{\mathbb{N}} \int_{\mathbb{N}} 1_A(\omega_1, \omega_2)\, d\mu_1(\omega_1) d\mu_2(\omega_2) &= \int_{\mathbb{N}} \mu_1(\{\omega_1 \,|\, \omega_1 \geq \omega_2\})\, d\mu_2(\omega_2) \\ &= \int_{\mathbb{N}} 1\, d\mu_2 \;=\; 1. \end{aligned}$$

Also gilt der Satz von Fubini in diesem Fall nicht, obwohl alle Voraussetzungen bis auf die Stetigkeit der Kapazitäten erfüllt sind.

Für Funktionen, die bezüglich der partiellen Ordnung auf dem Produktraum isoton sind, ist also bei Produktkapazitäten der Satz von Fubini gültig. Eine weitere schöne Eigenschaft der Produktkapazitäten ist, daß das Choquet-Integral nicht nur für komonotone Funktionen additiv ist, sondern auch falls beide Funktionen isoton bezüglich der partiellen Ordnung sind. Wir präzisieren dies im folgenden Satz.

Satz 2.7 *Die Voraussetzung (P) sei erfüllt. Die reellen Funktionen* f, g *seien* $\mathcal{A}$*-meßbar,* μ*-integrierbar und bezüglich* $\leq$ *isoton. Dann gilt*

$$\int_\Omega (f + g)\, d\mu \;=\; \int_\Omega f\, d\mu + \int_\Omega g\, d\mu.$$

Beweis: Nach dem Satz von Fubini ist

$$\int_\Omega (f + g)\, d\mu = \int_{\Omega_1} \left[\int_{\Omega_2} (f + g)_{\omega_1}\, d\mu_2 \right] d\mu_1(\omega_1).$$

Da für jedes ω_1 die Funktionen f_{ω_1} und g_{ω_1} isoton bezüglich der schwachen Ordnung $\leq_2$ sind, sind sie insbesondere komonoton. Daher ist

$$\int_\Omega (f + g)\, d\mu = \int_{\Omega_1} \left[\int_{\Omega_2} f_{\omega_1}\, d\mu_2 + \int_{\Omega_2} g_{\omega_1}\, d\mu_2 \right] d\mu_1(\omega_1).$$

Wegen der Isotonie von f und g in ω_1 sind aber die Funktionen $\omega_1 \mapsto \int f_{\omega_1}\, d\mu_2$ und $\omega_1 \mapsto \int g_{\omega_1}\, d\mu_2$ isoton und somit komonoton. Daher ist

$$\begin{aligned}&\int_{\Omega_1} \left[\int_{\Omega_2} f_{\omega_1}\, d\mu_2 + \int_{\Omega_2} g_{\omega_1}\, d\mu_2\right] d\mu_1(\omega_1) \\ &\qquad = \int_{\Omega_1} \left[\int_{\Omega_2} f_{\omega_1}\, d\mu_2\right] d\mu_1(\omega_1) + \int_{\Omega_1} \left[\int_{\Omega_2} g_{\omega_1}\, d\mu_2\right] d\mu_1(\omega_1).\end{aligned}$$

Die letzten beiden Gleichungen ergeben dann sofort die Behauptung. ■

2.5 Unscharfe Maße

Ein weiterer Bereich, in dem Kapazitäten Anwendung finden, ist die Theorie der „unscharfen Mengen“ bzw. der „unscharfen Maße“ (engl. „fuzzy sets“ bzw. „fuzzy measures“). Der Ausgangspunkt der Theorie der unscharfen Maße ist Dempster (1967). Er betrachtet in seinem Artikel die folgende Fragestellung:

$(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$ seien zwei Meßräume, P ein W-Maß auf $\mathcal{A}_1$ und $\Gamma : \Omega_1 \to \mathcal{A}_2$ eine mengenwertige Funktion mit $\Gamma(\omega_1) \neq \emptyset$ für alle $\omega_1 \in \Omega_1$. Im Rahmen der Entscheidungstheorie kann diese Situation folgendermaßen interpretiert werden. Ω_1 ist die Menge aller möglichen Umweltzustände und Γ eine mögliche Aktion des ET. Ist ω_1 der wahre Umweltzustand, so ist ein $\omega_2 \in \Gamma(\omega_1)$ das Ergebnis dieser Aktion. Die Wahrscheinlichkeiten des Eintreffens der verschiedenen Umweltzustände sind gegeben, wohingegen keine Informationen vorliegen, welches Element aus $\Gamma(\omega_1)$ das Ergebnis der Aktion sein wird. Gesucht ist nun eine Mengenfunktion μ mit der Eigenschaft, daß $\mu(A)$ die Wahrscheinlichkeit ist, daß das Ergebnis der Aktion Γ zur Menge A gehört. Ist $\Gamma(\omega)$ für alle $\omega \in \Omega$ einelementig, so ist Γ eine Zufallsvariable und das Bildmaß P^Γ die gesuchte Mengenfunktion.

In der eben beschriebenen allgemeinen Situation läßt sich jedoch keine eindeutige Mengenfunktion μ angeben. Es lassen sich lediglich obere und untere Schranken für die gesuchte Wahrscheinlichkeit angeben. Für jede Teilmenge $A \in \mathcal{A}_2$ definiert Dempster

$$\mu_*(A) = P(\{\omega_1 \in \Omega_1 \,|\, \Gamma(\omega_1) \subset A\})\,, \tag{2.19}$$
$$\mu^*(A) = P(\{\omega_1 \in \Omega_1 \,|\, \Gamma(\omega_1) \cap A \neq \emptyset\})\,. \tag{2.20}$$

Dann ist offensichtlich $\mu_*(A)$ eine untere und $\mu^*(A)$ eine obere Schranke für die gesuchte Wahrscheinlichkeit. Wegen

$$\begin{aligned}\mu^*(A) &= P(\{\omega_1 \in \Omega_1 \,|\, \Gamma(\omega_1) \not\subset A^c\} \\ &= 1 - P(\{\omega_1 \in \Omega_1 \,|\, \Gamma(\omega_1) \subset A^c\} \\ &= 1 - \mu_*(A^c)\end{aligned}$$

sind μ_* und μ^* zueinander dual.

Beispiel 2.5 (Ellsberg-Paradox, vgl. Beispiel 1.1) *Sei* $\Omega_1 = \{r, gb\}$ *und* $\Omega_2 = \{r, g, b\}$. *Dann ist* P *gegeben durch* $P(\{r\}) = \frac{1}{3}$, $P(\{gb\}) = \frac{2}{3}$. *Weiter ist* Γ *gegeben durch* $\Gamma(r) = \{r\}$, $\Gamma(\{gb\}) = \{g, b\}$. *Damit liegt genau die oben beschriebene Situation vor.*

Die oben definierten Mengenfunktionen μ_* und μ^* sind Kapazitäten. Darüberhinaus erfüllen sie die folgenden Eigenschaften. Für alle Mengen $A_1, \ldots, A_n \in \mathcal{A}_2$ gilt:

$$\mu_*\Big(\bigcup_{i=1}^{n} A_i\Big) \geq \sum_{\emptyset \neq I \subset \{1,\ldots n\}} (-1)^{|I|-1} \mu_*\Big(\bigcap_{i \in I} A_i\Big), \tag{2.21}$$

$$\mu^*\Big(\bigcup_{i=1}^{n} A_i\Big) \leq \sum_{\emptyset \neq I \subset \{1,\ldots n\}} (-1)^{|I|-1} \mu^*\Big(\bigcap_{i \in I} A_i\Big), \tag{2.22}$$

Die Untersuchung von Kapazitäten mit den Eigenschaften (2.21) und (2.22) steht im Mittelpunkt der Theorie der unscharfen Maße.

Definition 2.12 (Vertrauensmaß, Plausibilitätsmaß) *Eine Kapazität* μ*, für die zusätzlich (2.21) gilt, heißt* Vertrauensmaß *(engl.* Belief Measure*). Eine Kapazität* μ*, die zusätzlich (2.22) erfüllt, heißt* Plausibilitätsmaß *(engl.* Plausibility Measure*).*

Setzt man $n = 2$, so erhält man aus (2.21) die starke Superadditivität von μ_* und aus (2.22) die starke Subadditivität von μ^*.

Jedes Vertrauensmaß und jedes Plausibilitätsmaß können durch eine Funktion m, das sogenannte *basic assignment* ausgedrückt werden.

Satz 2.8 *Sei* Ω *endlich und* μ *ein Vertrauensmaß auf* $(\Omega, \mathcal{P}(\Omega))$. *Dann gibt es eine Funktion* $m : \mathcal{P}(\Omega) \to \mathbb{R}$ *mit* $m(A) \geq 0$ *für alle* $A \subset \Omega$ *und* $\sum_{A \subset \Omega} m(A) = 1$, *so daß*

$$\mu(A) = \sum_{B \subset A} m(B) \tag{2.23}$$

und

$$\mu^D(A) = \sum_{B \cap A \neq \emptyset} m(B)\,.$$

Die Funktion m *ist dabei gegeben durch*

$$m(A) = \sum_{B \subset A} (-1)^{|A \setminus B|} \mu(B) \tag{2.24}$$

Beweis: Shafer (1976, Kapitel 2, §7).

In der eingangs beschriebenen Situation ist das basic assignment m gerade durch $m(A) = P(\Gamma = A)$ gegeben.

Durch (2.24) wird einem Vertrauensmaß μ eine Funktion m zugeordnet. Umgekehrt ordnet (2.23) jedem basic assignment ein Vertrauensmaß zu. Man kann nun natürlich auch einer Kapazität μ, die kein Vertrauensmaß ist, eine Funktion $m : \mathcal{A} \to \mathbb{R}$ mittels (2.24) zuordnen. In diesem allgemeinen Fall heißt m die *Möbius-Inverse* von μ. Für m gilt dann

$$m(\emptyset) = 0. \tag{2.25}$$

$$\sum_{A \subset \Omega} m(A) = 1. \tag{2.26}$$

Insbesondere ist i.a. $m(A) \geq 0$ nicht für alle $A \in \mathcal{A}$ erfüllt. Ist m die Möbius-Inverse einer Kapazität μ, so gilt (2.23) auch, wenn μ kein Vertrauensmaß ist. Es ist μ nämlich genau dann ein Vertrauensmaß, wenn die Möbius-Inverse von μ nichtnegativ ist (siehe z.B. Chateauneuf und Jaffray (1989), wo noch andere Charakterisierungen von Kapazitäten durch Eigenschaften ihrer Möbius-Inversen gegeben werden).

Interessant ist vor allem die Darstellung des Choquet-Integrals durch die Möbius-Inverse der Kapazität. Die folgenden Sätze, in denen immer vorausgesetzt wird, daß Ω endlich ist, finden sich in Gilboa und Schmeidler (1992a). In Gilboa und Schmeidler (1992b) finden sich Verallgemeinerungen dieser Sätze für den Fall, daß Ω unendlich ist.

Satz 2.9 (Gilboa und Schmeidler 1992a) *Sei Ω endlich, μ eine Kapazität und m ihre Möbius-Inverse. Dann ist für jede Funktion $f : \Omega \to \mathbb{R}$*

$$\int f d\mu = \sum_{A \subset \Omega} \min_{\omega \in A} f(\omega)\, m(A).$$

Ist μ ein Vertrauensmaß, so kann also das Choquet-Integral einer Funktion f als gewichteter Mittelwert aller Minima von f über Teilmengen von Ω dargestellt werden.

Dieser Satz zeigt auch die Additivität des Choquet-Integrals für komonotone Funktionen in einem neuen Licht. Nach Satz 2.9 erhält man nämlich für das Choquet-Integral einer Summe von Funktionen

$$\int (f+g)\, d\mu = \sum_{A \subset \Omega} \min_{\omega \in A} [f(\omega) + g(\omega)]\, m(A).$$

Dies ist dann gleich der Summe der Choquet-Integrale, wenn

$$\min_{\omega \in A} [f(\omega) + g(\omega)] = \min_{\omega \in A} f(\omega) + \min_{\omega \in A} g(\omega) \quad \text{für alle } A \subset \Omega \text{ mit } m(A) \neq 0 \tag{2.27}$$

gilt, d.h. wenn es für jede Menge A mit $m(A) \neq 0$ ein $\omega^\star$ gibt, das f und g simultan auf A mimimiert. Diese Bedingung ist offensichtlich dann erfüllt, wenn die Funktionen f und g komonoton sind. Mit anderen Worten: Sind f und g komonoton, so ist $\int(f+g)d\mu = \int f d\mu + \int g d\mu$ für *alle* Kapazitäten μ.

Auf der anderen Seite ist auch die folgende Frage von Interesse: Sei μ eine gegebene Kapazität. Für welche Paare (f,g) von Funktionen gilt $\int(f+g)d\mu = \int f d\mu + \int g d\mu$? Die Antwort auf diese Frage liefert Gleichung (2.27). Es sind dies nämlich genau die Paare von Funktionen, die (2.27) erfüllen.

Lemma 2.6 (Gilboa und Schmeidler 1992a) *Sei Ω endlich und μ eine Kapazität auf Ω. Dann gibt es Vertrauensmaße μ^+, μ^-, so daß*

$$\mu \;=\; \mu^+ - \mu^- \,.$$

Für jede Funktion $f : \Omega \to \mathbb{R}$ gilt

$$\int f d\mu \;=\; \int f d\mu^+ - \int f d\mu^- \,.$$

Satz 2.10 (Gilboa und Schmeidler 1992a) *Sei Ω endlich und μ eine Kapazität auf Ω. Dann gibt es Mengen Π^+ und Π^- von W-Maßen auf Ω und $\alpha \in [0.5, 1]$, so daß für alle Funktionen $f : \Omega \to \mathbb{R}$ gilt*

$$\int f d\mu \;=\; \alpha \;\min\{\int f dP \,|\, P \in \Pi^+\} - (1-\alpha)\;\min\{\int f dP \,|\, P \in \Pi^-\} \,.$$

2.6 Beweise

Beweis von Satz 2.4

Bei unserem Beweis werden wir uns hauptsächlich auf den Satz 2.11 beziehen. Bevor wir diesen formulieren können, sind noch ein paar Vorbereitungen notwendig.

Definition 2.13 (Kompakte Klasse) *Sei Ω eine beliebige Menge. Eine Familie $\mathcal{K}$ von Teilmengen von Ω heißt* kompakte Klasse*, wenn für jede Folge $(K_n)_{n\in\mathbb{N}}$ in $\mathcal{K}$ mit $\bigcap_{n\in\mathbb{N}} K_n = \emptyset$ eine natürliche Zahl N existiert, so daß $\bigcap_{n=1}^{N} K_n = \emptyset$.*

Bemerkung: Ist $\mathcal{K}$ durchschnittstabil, so ist $\mathcal{K}$ genau dann eine kompakte Klasse, wenn für jede *absteigende* Folge $(K_n)_{n\in\mathbb{N}}$ in $\mathcal{K}$ mit $K_n \searrow \emptyset$ ein N existiert mit $\bigcap_{n=1}^{N} K_n = \emptyset$.

Definition 2.14 (Semi-Algebra) *Sei Ω eine beliebige Menge. Eine Familie $\mathcal{S}$ von Teilmengen von Ω heißt* Semi-Algebra*, wenn sie die folgenden Eigenschaften besitzt:*

(i) $\emptyset$, $\Omega \in \mathcal{S}$.

(ii) *Sind* $A, B \in \mathcal{S}$, *so ist auch* $A \cap B \in \mathcal{S}$.

(iii) *Ist* $A \in \mathcal{S}$, *so ist* A^c *Vereinigung einer endlichen Familie paarweise disjunkter Mengen aus* $\mathcal{S}$.

Satz 2.11 *Seien* $\mathcal{S}$ *eine Semi-Algebra auf* Ω, $\mathcal{K}$ *eine kompakte Klasse auf* Ω *und* $P : \mathcal{S} \to [0,1]$ *eine additive Mengenfunktion mit* $P(\Omega) = 1$ *und der folgenden Approximationseigenschaft:*

> *Für jedes* $A \in \mathcal{S}$ *und* $\epsilon > 0$ *gibt es* $B \in \mathcal{S}$ *und* $K \in \mathcal{K}$, *so daß* $B \subset K \subset A$ *und* $P(A \setminus B) < \epsilon$ *ist.*

Dann kann P *auf genau eine Weise zu einem* σ*-additiven W-Maß auf der von* $\mathcal{S}$ *erzeugten* σ*-Algebra* $\mathcal{A}$ *fortgesetzt werden.*

Beweis: Pfanzagl und Pierlo (1966, S. 10), siehe auch Neveu (1969, S. 44).

Beweis von Satz 2.4: Der Beweis besteht aus zwei Schritten. Wir führen den Beweis zunächst unter der Annahme, daß $(\Omega, \leq)$ linear geordnet[3] ist. Im zweiten Schritt beweisen wir dann den Satz unter der Voraussetzung, daß $\leq$ lediglich eine schwache Ordnung ist.

Schritt 1: Sei $(\Omega, \leq)$ linear geordnet.

Wir bezeichen die dekumulative Verteilungsfunktion von μ mit F, d.h. es ist $F : \Omega \to [0,1]$ gegeben durch $F(x) = \mu((x, \infty])$.

Für jede obere Menge U gilt entweder

$$\inf\{F(x) \,|\, x \notin U\} = \sup\{F(x) \,|\, x \in U\}$$

oder

$$\inf\{F(x) \,|\, x \notin U\} > \sup\{F(x) \,|\, x \in U\}\,.$$

Im zweiten Fall sagen wir, daß F einen Sprung von U^c nach U hat. Da für alle $a \notin U$, $b \in U$ gilt

$$(a, \infty] \supset U \underset{\neq}{\supset} (b, \infty]\,,$$

ist $\mu(U) = \sup\{F(x) \,|\, x \in U\}$, falls F keinen Sprung von U^c nach U hat. Hat F einen Sprung von U^c nach U, so kann $\mu(U)$ jeden Wert zwischen $\sup\{F(x) \,|\, x \in U\}$ und $\inf\{F(x) \,|\, x \notin U\}$ haben.

Da F nur abzählbar viele Sprünge haben kann, gibt es Gewichte α_i, $i \in \mathbb{N}_0$, so daß μ darstellbar ist als $\mu = \sum_{i=0}^{\infty} \alpha_i \mu_i$, wobei die zu μ_0 gehörige dekumulative Verteilungsfunktion F_0 keine Sprünge hat und die μ_i die Form haben

$$\mu(A) = \begin{cases} 1, & \text{falls } U \underset{\neq}{\subset} A, \\ 0, & \text{sonst,} \end{cases} \tag{2.28}$$

[3] Zu ordnungstheoretischen Begriffen und Bezeichnungen vergleiche Anhang A.1.

oder

$$\mu(A) = \begin{cases} 1, & \text{falls } U \subset A, \\ 0, & \text{sonst,} \end{cases} \tag{2.29}$$

für eine obere Menge U. Es reicht also aus, die Behauptung für den Fall zu zeigen, daß μ entweder eine stetige dekumulative Verteilungsfunktion besitzt oder aber vom Typ (2.28) oder (2.29) ist.

Sei

$$\mathcal{J} = \Big\{ U \setminus V \,|\, U, V \in \mathcal{U},\, V \subset U \Big\}$$

Dann ist $\mathcal{J}$ eine Semi-Algebra, die alle oberen Mengen enthält. Wir definieren P auf $\mathcal{J}$ durch $P(U \setminus V) = \mu(U) - \mu(V)$. Sind die Voraussetzungen von Satz 2.11 erfüllt, so kann P in eindeutiger Weise zu einem σ-additiven W-Maß auf der von $\mathcal{J}$ erzeugten σ-Algebra $\mathcal{A}$ fortgesetzt werden. Da $\mathcal{A}$ bereits von der Familie $\mathcal{U}$ aller oberen Mengen erzeugt wird, folgt damit die Behauptung.

Wir müssen also zeigen, daß eine kompakte Klasse $\mathcal{K}$ existiert, so daß für jedes $A \in \mathcal{J}$ und $\epsilon > 0$ eine Menge $B \in \mathcal{J}$ mit $P(A \setminus B) < \epsilon$ und eine Menge $K \in \mathcal{K}$ existieren, so daß $B \subset K \subset A$ gilt.

Fall 1: Wir nehmen an, daß F keine Sprünge hat, d.h. daß $\inf\{F(x) \,|\, x \notin U\} = \sup\{F(x) \,|\, x \in U\}$ für alle oberen Mengen U gilt.

Wir zeigen, daß das Mengensystem $\mathcal{K} = \{[a, b] \,|\, a, b \in \Omega\}$ eine kompakte Klasse ist. Sei also $([a_n, b_n])_{n\in\mathbb{N}}$ eine Folge in $\mathcal{K}$. Da $\mathcal{K}$ durchschnittstabil ist, können wir die Folge $([a_n, b_n])_{n\in\mathbb{N}}$ als absteigend voraussetzen. Dann sind die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ monoton und durch a_1 nach unten und durch b_1 nach oben beschränkt. Somit existieren, da Ω vollständig bezüglich beschränkter, monotoner Folgen ist, $\sup_{n\in\mathbb{N}} a_n$ und $\inf_{n\in\mathbb{N}} b_n$ und es ist

$$\bigcap_{n\in\mathbb{N}} [a_n, b_n] = [\sup_{n\in\mathbb{N}} a_n, \inf_{n\in\mathbb{N}} b_n].$$

Ist dieser Durchschnitt leer, so ist $\sup_{n\in\mathbb{N}} a_n > \inf_{n\in\mathbb{N}} b_n$. Dann gibt es aber ein N, so daß $\sup_{1\le n\le N} a_n > \inf_{1\le n\le N} b_n$ und somit ist $\bigcap_{n=1}^{N} [a_n, b_n] = \emptyset$. $\mathcal{K}$ ist also eine kompakte Klasse.

Sei $A = U \setminus V$. Da wir angenommen haben, daß F keine Sprünge hat, ist $\mu(U) = \sup\{F(x) \,|\, x \in U\}$ und $\mu(V) = \inf\{F(x) \,|\, x \notin V\}$. Daher gibt es zu jedem $\epsilon > 0$ ein $x_\epsilon \in U$ mit $\mu(U) \ge F(x_\epsilon) > \mu(U) - \frac{\epsilon}{2}$ und ein $y_\epsilon \notin V$ mit $\mu(V) \le F(y_\epsilon) < \mu(V) + \frac{\epsilon}{2}$. x_ϵ und y_ϵ können so gewählt werden, daß $x_\epsilon \le y_\epsilon$. Dann ist $U \setminus V \supset [x_\epsilon, y_\epsilon] \supset (x_\epsilon, y_\epsilon]$ und

$$\begin{aligned} P(U \setminus V) - P((x_\epsilon, y_\epsilon]) &= [\mu(U) - \mu(V)] - [F(x_\epsilon) - F(y_\epsilon)] \\ &= [\mu(U) - F(x_\epsilon)] + [F(y_\epsilon) - \mu(V)] \\ &< \epsilon. \end{aligned}$$

Daher besitzen $B = (x_\epsilon, y_\epsilon]$ und $K = [x_\epsilon, y_\epsilon]$ die verlangten Eigenschaften.

Fall 2: Wir nehmen an, daß μ vom Typ (2.28) oder (2.29) ist.

Fall 2a: Sei μ gegeben durch

$$\mu(A) = \begin{cases} 1, & \text{falls } U \subsetneq A, \\ 0, & \text{sonst,} \end{cases}$$

für eine obere Menge U. Wir zeigen zunächst, daß das Mengensystem

$$\mathcal{K} = \{V \setminus U \mid V \in \mathcal{U},\, V \supset U\}$$

eine kompakte Klasse ist.

Sei also $(V_n \setminus U)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{K}$ mit leerem Durchschnitt. Da $\mathcal{K}$ durchschnittstabil ist, können wir die Folge wieder als absteigend voraussetzen. Das bedeutet, daß die Folge $(V_n)_{n \in \mathbb{N}}$ ebenfalls absteigend ist. Ist $V_n \supsetneq U$ für alle $n \in \mathbb{N}$, so ist wegen

$$\Big(\bigcap_{n=1}^{\infty} V_n\Big) \setminus U = \bigcap_{n=1}^{\infty} (V_n \setminus U) = \emptyset$$

$U = \bigcap_{n=1}^{\infty} V_n$. Wegen der Stetigkeit von μ auf oberen Mengen führt dies zum Widerspruch

$$1 = \lim_{n \to \infty} \mu(V_n) = \mu(U) = 0.$$

Also kann $V_n \supsetneq U$ nicht für alle n gelten, d.h. es gibt ein N mit $V_N = U$. Dann ist $V_N \setminus U = \emptyset$ und $\cap_{n=1}^{N} (V_N \setminus U) = \emptyset$. Also ist $\mathcal{K}$ eine kompakte Klasse.

Sei nun wieder $A \in \mathcal{J}$, d.h. $A = V \setminus W$ für obere Mengen V und W. Für $P(A) = 0$ wählt man $B = K = \emptyset$. Ist $P(A) = 1$, so muß $W \subset U \subsetneq V$ sein. Für $B = K = V \setminus U$ ist daher $A \supset K \supset B$ und es gilt $P(B) = \mu(V) - \mu(U) = 1$ und $K \in \mathcal{K}$. B und K besitzen also die in Satz 2.11 geforderte Approximationseigenschaft.

Fall 2b: Sei μ gegeben durch

$$\mu(A) = \begin{cases} 1, & \text{falls } U \subset A, \\ 0, & \text{sonst,} \end{cases}$$

für eine obere Menge U. Wir zeigen, daß das Mengensystem

$$\mathcal{K} = \{U \setminus V \mid V \in \mathcal{U},\, V \subset U\}$$

eine kompakte Klasse ist.

Sei also $(U \setminus V_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{K}$ mit leerem Durchschnitt. Wieder nutzen wir aus, daß $\mathcal{K}$ durchschnittstabil ist und setzen die Folge als absteigend voraus. Dies bedeutet, daß die Folge $(V_n)_{n \in \mathbb{N}}$ aufsteigend ist. Ist $V_n \subsetneq U$ für alle $n \in \mathbb{N}$, so ist wegen

$$U \setminus \Big(\bigcup_{n=1}^{\infty} V_n\Big) = U \cap \Big(\bigcup_{n=1}^{\infty} V_n\Big)^c = U \cap \bigcap_{n=1}^{\infty} V_n^c = \bigcap_{n=1}^{\infty} (U \setminus V_n) = \emptyset$$

$U = \bigcup_{n=1}^{\infty} V_n$. Wegen der Stetigkeit von μ auf oberen Mengen führt dies zum Widerspruch

$$0 = \lim_{n\to\infty} \mu(V_n) = \mu(U) = 1.$$

Also kann $V_n \subsetneq U$ nicht für alle n gelten, d.h. es gibt ein N mit $V_N = U$. Dann ist $U \setminus V_N = \emptyset$ und $\cap_{n=1}^{N}(U \setminus V_n) = \emptyset$. Also ist $\mathcal{K}$ eine kompakte Klasse.

Sei nun $A \in \mathcal{J}$, d.h. $A = V \setminus W$ für obere Mengen V und W. Für $P(A) = 0$ wählt man $B = K = \emptyset$. Ist $P(A) = 1$, so muß $W \subsetneq U \subset V$ sein. Für $B = K = U \setminus W$ ist daher $A \supset K \supset B$ und es gilt $P(B) = \mu(U) - \mu(W) = 1$ und $K \in \mathcal{K}$. Somit besitzen B und K die Approximationseigenschaft aus Satz 2.11.

Wir haben also gezeigt, daß sowohl für Kapazitäten mit stetiger dekumulativer Verteilungsfunktion als auch für Kapazitäten gemäß (2.28) und (2.29) ein W-Maß existiert, das mit der Kapazität auf allen oberen Mengen übereinstimmt. Da jede Kapazität μ als abzählbare Linearkombination solcher Choquet-Kapazitäten darstellbar ist, existiert also für jede Kapazität μ ein W-Maß P, das mit μ auf allen oberen Mengen übereinstimmt.

Schritt 2: $(\Omega, \leq)$ sei schwach geordnet. In diesem Fall ist $(\Omega^*, \leq^*)$ linear geordnet und vollständig bezüglich beschränkter, monotoner Folgen. Wir definieren μ^* auf $\mathcal{U}^* = \{U \subset \Omega^* \,|\, U \text{ obere Menge in } \Omega^*\} = \{U \subset \Omega^* \,|\, \overline{U} \in \mathcal{U}\}$ durch $\mu^*(U) = \mu(\overline{U})$, wobei $\overline{U} = \{x \in \Omega \,|\, x^* \in U\}$. Da U genau dann eine obere Menge (in Ω^*) ist, wenn $\overline{U}$ eine obere Menge (in Ω) ist, ist auch μ^* stetig auf oberen Mengen. Daher existiert ein W-Maß P^* auf der von $\mathcal{U}^*$ erzeugten σ-Algebra $\mathcal{A}^*$, das auf $\mathcal{U}^*$ mit μ^* übereinstimmt. Ist $\mathcal{A}$ die von $\mathcal{U}$ erzeugte σ-Algebra, so ist $\mathcal{A} = \{\overline{A} \,|\, A \in \mathcal{A}^*\}$. Man definiert dann ein W-Maß P auf $\mathcal{A}$ durch $P(\overline{A}) = P^*(A)$. Dann stimmt P auf $\mathcal{U}$ mit μ überein, wie behauptet. ■

Beweis von Satz 2.5

Wir zeigen zunächst, daß die beiden iterierten Integrale gleich sind. Da μ_1 stetig auf oberen Mengen und $(\Omega_1, \leq_1)$ vollständig bezüglich beschränkter, monotoner Folgen ist, existiert nach Satz 2.4 eine W-Maß P_1 auf der σ-Algebra $\mathcal{A}_1$, das mit μ_1 auf $\mathcal{U}_1$ übereinstimmt. Somit gilt

$$\int_{\Omega_1} f \, d\mu_1 \;=\; \int_{\Omega_1} f \, dP_1$$

für alle isotonen Funktionen $f : \Omega_1 \to \mathbb{R}$. Die entsprechende Aussage gilt für Ω_2.

Ist f isoton bezüglich $\leq$, so ist für jedes ω_1 die Funktion $f_{\omega_1} : \omega_2 \mapsto f(\omega_1, \omega_2)$ isoton bezüglich $\leq_2$. Daher ist für alle $\omega_1 \in \Omega_1$:

$$\int_{\Omega_2} f_{\omega_1} d\mu_2 = \int_{\Omega_2} f_{\omega_1} dP_2 \,.$$

Da die Funktion $\omega_1 \mapsto \int_{\Omega_2} f_{\omega_1} dP_2$ isoton bezüglich $\leq_1$ ist, gilt dann

$$\begin{aligned} \int_{\Omega_1} \left[\int_{\Omega_2} f d\mu_2 \right] d\mu_1 &= \int_{\Omega_1} \left[\int_{\Omega_2} f dP_2 \right] d\mu_1 \\ &= \int_{\Omega_1} \left[\int_{\Omega_2} f dP_2 \right] dP_1 \end{aligned} \tag{2.30}$$

Genauso zeigt man, daß auch

$$\int_{\Omega_2} \left[\int_{\Omega_1} f d\mu_1 \right] d\mu_2 = \int_{\Omega_2} \left[\int_{\Omega_1} f dP_1 \right] dP_2 \tag{2.31}$$

ist. Da f $\mathcal{A}$-meßbar ist, kann nach dem Satz von Fubini (für σ-additive Maße) die Reihenfolge der Integration vertauscht werden. Daher ist

$$\int_{\Omega_1} \left[\int_{\Omega_2} f dP_2 \right] dP_1 = \int_{\Omega_2} \left[\int_{\Omega_1} f dP_1 \right] dP_2 .$$

Zusammen mit (2.30) und (2.31) liefert dies den Beweis dafür, daß in der angegebenen Situation auch bei Choquet-Integralen die Reihenfolge der Integration vertauscht werden kann.

Wir müssen nun noch zeigen, daß das Integral bezüglich der Produktkapazität gleich einem der beiden iterierten Integrale ist.

Ist f eine Indikatorfunktion einer oberen Menge, so folgt dies direkt aus der Definition einer Produktkapazität.

Ist f eine Elementarfunktion, d.h. nimmt f nur endlich viele Werte $\alpha_1 < \ldots < \alpha_n$ an, so ist

$$f = \sum_{i=1}^{n} (\alpha_i - \alpha_{i-1}) 1_{\{f \geq \alpha_i\}},$$

wobei $\alpha_0 = 0$ gesetzt wurde. Da die Differenzen $\alpha_i - \alpha_{i-1}$ sämtlich nichtnegtiv sind und die Indikatorfunktionen $1_{\{f \geq \alpha_i\}}$, $i = 1, \ldots n$, paarweise komonoton sind, folgt die Behauptung aus der positiven Homogenität und der komonotonen Additivität des Choquet-Integrals.

Ist f schließlich eine beliebige $\mathcal{A}$-meßbare Funktion, so gibt es eine Folge $(f_n)_{n \in \mathbb{N}}$ von isotonen Elementarfunktionen, die monoton gegen f konvergieren. Da μ nach Lemma 2.5 stetig auf oberen Mengen ist, gilt für isotone Funktionen der Satz von der monotonen Konvergenz. Da nach dem Vorhergehendem für isotone Elementarfunktionen das Integral bezüglich der Produktkapazität mit den iterierten Integralen übereinstimmt, gilt dies auch für beliebige $\mathcal{A}$-meßbare, isotone Funktionen. ∎

Beweis von Satz 2.6

Wir beweisen den Satz zunächst für den Fall, daß f isoton ist. Es ist

$$\int_\Omega f\ d\mu = \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu^D.$$

Nach Satz 2.5 kann das Integral über f^+ als Doppelintegral geschrieben werden. Da μ^D nach Lemma 2.3 eine Produktkapazität von μ_1^D und μ_2^D ist und wegen Korollar 2.3 (f^- ist antiton !) gilt dasselbe für das Integral bezüglich f^-. Somit ist

$$\int_\Omega f\ d\mu = \int_{\Omega_1} \left[\int_{\Omega_2} f^+_{\omega_1} d\mu_2\right] d\mu_1(\omega_1) - \int_{\Omega_1} \left[\int_{\Omega_2} f^-_{\omega_1} d\mu_2^D\right] d\mu_1^D(\omega_1).$$

Wegen (2.6) ist dies äquivalent zu

$$\int_\Omega f\ d\mu = \int_{\Omega_1} \left[\int_{\Omega_2} f^+_{\omega_1} d\mu_2\right] d\mu_1(\omega_1) + \int_{\Omega_1} \left[-\int_{\Omega_2} f^-_{\omega_1} d\mu_2^D\right] d\mu_1(\omega_1). \tag{2.32}$$

Um dies als Integral (bezgl. μ_1) zu schreiben, müssen wir untersuchen, für welche $\omega_1 \in \Omega_1$ die Summe $\int f^+_{\omega_1} d\mu_2 - \int f^-_{\omega_1} d\mu_2^D$ überhaupt definiert ist, d.h. für welche ω_1 nicht beide Integrale gleich ∞ werden.

Da f integrierbar ist, ist sowohl $\int f^+ d\mu$ als auch $\int f^- d\mu^D$ endlich. Da f^+ und f^- nichtnegativ sind ist nach Satz 2.5

$$\int_{\Omega_1} \left[\int_{\Omega_2} f^+_{\omega_1} d\mu_2\right] d\mu_1(\omega_1) < \infty \quad \text{und} \quad \int_{\Omega_1} \left[\int_{\Omega_2} f^-_{\omega_1} d\mu_2^D\right] d\mu_1^D(\omega_1) < \infty.$$

Daher muß $\mu_1(\{\omega_1 \mid \int f^+_{\omega_1} d\mu_2 = \infty\}) = 0$ und $\mu_1^D(\{\omega_1 \mid \int f^-_{\omega_1} d\mu_2^D = \infty\}) = 0$ sein. Dabei ist die zweite Aussage gleichbedeutend mit $\mu_1(\{\omega_1 \mid \int f^-_{\omega_1} d\mu_2^D < \infty\}) = 1$. Da f^+ isoton und f^- antiton ist, sind die Mengen $\{\omega_1 \mid \int f^+_{\omega_1} d\mu_2 = \infty\}$ und $\{\omega_1 \mid \int f^-_{\omega_1} d\mu_2^D < \infty\}$ obere Mengen. Obere Mengen auf einem schwach geordneten Raum sind aber linear geordnet bezüglich Mengeninklusion, so daß $\{\omega_1 \mid \int f^+_{\omega_1} d\mu_2 = \infty\} \subset \{\omega_1 \mid \int f^-_{\omega_1} d\mu_2^D < \infty\}$ oder die umgekehrte Inklusion gelten muß. Wegen der Isotonie von μ_1 muß aber $\{\omega_1 \mid \int f^+_{\omega_1} d\mu_2 = \infty\} \subset \{\omega_1 \mid \int f^-_{\omega_1} d\mu_2^D < \infty\}$ sein. Daher sind die Mengen $\{\omega_1 \mid \int f^+_{\omega_1} d\mu_2 = \infty\}$ und $\{\omega_1 \mid \int f^-_{\omega_1} d\mu_2^D = \infty\}$ disjunkt, und $\int f^+_{\omega_1} d\mu_2 - \int f^-_{\omega_1} d\mu_2^D$ ist überall definiert. Außerdem sind $\int f^+_{\omega_1} d\mu_2$ und $-\int f^-_{\omega_1} d\mu_2^D$ isoton in ω_1 und daher komonoton. Somit wird (2.32) zu

$$\begin{aligned}\int_\Omega f\ d\mu &= \int_{\Omega_1} \left[\int_{\Omega_2} f^+_{\omega_1} d\mu_2 - \int_{\Omega_2} f^-_{\omega_1} d\mu_2^D\right] d\mu_1(\omega_1) \\ &= \int_{\Omega_1} \left[\int_{\Omega_2} f_{\omega_1} d\mu_2\right] d\mu_1(\omega_1).\end{aligned}$$

Daß auch

$$\int_\Omega f\ d\mu = \int_{\Omega_2} \left[\int_{\Omega_1} f_{\omega_2} d\mu_1\right] d\mu_2(\omega_2).$$

gilt, zeigt man völlig analog. Somit ist die Behauptung für den Fall, daß f isoton ist, gezeigt.

Ist f antiton, so ist f isoton bezüglich $\leq^D$, wobei $\leq^D$ wie im Beweis von Korollar 2.3 definiert ist. Wie dort zeigt man, daß $\leq^D$ die Voraussetzung (P) erfüllt. Damit folgt dann die Behauptung sofort aus dem soeben Gezeigten. ∎

Kapitel 3

Antizipierter Nutzen und Choquet-Erwartungsnutzen

Wie wir in Abschnitt 1.3 gesehen haben, existieren eine Reihe von Verhaltensweisen, die den Axiomen von EU und SEU widersprechen und die nicht einfach als irrational bezeichnet werden können. So waren viele der Personen, die in den empirischen Untersuchungen von Slovic und Tversky (1974), Moskowitz (1974) und MacCrimmon und Larsson (1979) gegen die Axiome von EU oder SEU verstießen, auch dann nicht bereit, ihre Entscheidung zu ändern, wenn ihnen dargelegt wurde, wo und wie sie gegen die Axiome verstoßen hatten. Aus diesem Grunde wurde in den letzten Jahren eine wachsende Anzahl von Modellen entwickelt, die diese Phänomene erklären können (vgl. Fishburn 1988, Camerer und Weber 1992, Kischka und Puppe 1992). Wir konzentrieren uns hier auf eine Klasse von Theorien, denen gemeinsam ist, daß sie EU und SEU verallgemeinern, indem sie die Präferenzrelation des Entscheidungsträgers durch das Choquet-Integral einer Nutzenfunktion bezüglich einer Choquet-Kapazität repräsentieren.

Bevor wir die beiden von uns betrachteten Modelle – antizipierter Nutzen für Entscheidungen unter Risiko und Choquet-Erwartungsnutzen für Entscheidungen unter Unsicherheit – vorstellen, geben wir einen kurzen historischen Überblick. Sodann gehen wir kurz auf Axiomatisierungen dieser Modelle ein. Es wird gezeigt werden, daß nicht nur SEU und EU als Spezialfälle in diesen Modellen enthalten sind, sondern auch viele andere Entscheidungsregeln. CEU und AU stehen in enger Beziehung zum Π-Maximin-Prinzip bei vager Vorinformation. Dies wird Gegenstand eines weiteren Abschnittes sein. Wir werden zeigen, daß in bestimmten Situationen das Choquet-Erwartungsnutzen-Modell und das Π-Maximin-Prinzip zur selben Entscheidung führen. Schließlich werden wir darlegen, wie die in Abschnitt 1.3 betrachteten Phänomene im Rahmen von AU und CEU erklärt werden können.

3.1 Historisches

Bereits 1953 wurde durch das Allais-Paradox (vgl. Beispiel 1.2) deutlich, daß in Entscheidungsproblemen Verhaltensweisen existieren, die mit dem Erwartungsnutzenprinzip nicht erklärbar sind. Weitere empirische Untersuchungen (vgl. Ellsberg 1961, MacCrimmon 1968, Slovic und Tversky 1974, MacCrimmon und Larsson 1979, Kahneman und Tversky 1979) zeigten, daß das Allais-Paradox kein Einzelfall ist, sondern daß auch in realen Entscheidungssituationen Verhalten zu beobachten ist, das nicht im Einklang mit EU steht. Zur Erklärung dieser Phänomene nahmen einige Autoren an, daß Wahrscheinlichkeiten „verzerrt" aufgenommen werden (Edwards 1955, 1962, Fellner 1961). Dies bedeutet, daß Wahrscheinlichkeiten erst transformiert in die Bewertung einer Aktion eingehen, ähnlich wie ja auch in EU die möglichen Ergebnisse einer Aktion mit einer Nutzenfunktion bewertet werden. In der Folgezeit entwickelten Handa (1977), Karmarkar (1978, 1979) und Kahneman und Tversky (1979) Theorien, die dieser Annahme Rechnung trugen.

Als Aktionen betrachten wir die Menge aller reellen Zufallsvariablen mit endlichem Träger. Gemäß EU existiert eine Funktion u, so daß der ET die Aktion X wählt, für die

$$EU(X) = \sum_{i=1}^{n} u(x_i)p_i$$

maximal wird, wobei x_i, $i = 1, \dots n$, die möglichen Ergebnisse und p_i, $i = 1, \dots n$, die entsprechenden Wahrscheinlichkeiten sind. Handa (1977) schlug vor, die Präferenzen des ET durch das Funktional

$$V_H(X) = \sum_{i=1}^{n} x_i q(p_i)$$

zu repräsentieren, wobei $q : [0,1] \to \mathbb{R}$ eine monoton wachsende Funktion mit $q(0) = 0$ ist. Kahneman und Tversky (1979) betrachten nur Lotterien mit höchstens zwei von Null verschiedenen Ergebnissen x_1 und x_2, wobei die Ergebnisse Veränderungen gegenüber augenblicklichem Vermögen darstellen. Auf der Menge dieser Lotterien ist das Präferenzfunktional von Kahnemann und Tversky gegeben durch

$$V_{KT}(X) = \begin{cases} u(x_1)q(p_1) + u(x_2)q(p_2), & \text{falls } x_1x_2 \leq 0 \text{ oder } p_1 + p_2 < 1, \\ u(x_2) + q(p_1)(u(x_1) - u(x_2)), & \text{falls } x_1x_2 > 0 \text{ und } p_1 + p_2 = 1, \end{cases}$$

wobei $u(0) = q(0) = 0$, $q(1) = 1$ und u und q monoton wachsend sind. Das von Karmarkar (1978) vorgeschlagenen SWU-Modell („Subjectively Weighted Utility") verwendet das folgende Funktional

$$V_K(X) = \sum_{i=1}^{n} u(x_i) \frac{q(p_i)}{\sum_{j=1}^{n} q(p_j)},$$

wobei $q : [0,1] \to [0,1]$ mit

$$q(p) = \frac{p^\alpha}{p^\alpha + (1-p)^\alpha} \quad \text{mit } 0 < \alpha < \infty.$$

Zwar können mit diesen Modellen viele der beobachteten Phänomene erklärt werden, jedoch tritt ein neues Problem auf. Um dies zu erläutern benötigen wir den Begriff der stochastischen Dominanz erster Ordnung. Eine Lotterie X dominiert eine Lotterie Y stochastisch von erster Ordnung, wenn die Wahrscheinlichkeit, in der Lotterie X einen Betrag von mindestens t zu erhalten, für alle t mindestens so groß ist wie die entsprechende Wahrscheinlichkeit für die Lotterie Y. Es ist naheliegend anzunehmen, daß ein ET in dieser Situation die Lotterie X der Lotterie Y vorziehen wird. Man sagt von einem Präferenzfunktional mit dieser Eigenschaft, daß es stochastische Dominanz respektiert. Fishburn (1978) und Quiggin (1982) haben gezeigt, daß die eben beschriebenen Modelle stochastische Dominanz *nicht* respektieren. Der Grund hierfür ist grob gesprochen, daß in diesen Modellen nur individuelle Wahrscheinlicheiten transformiert werden. Quiggin war der erste der erkannte, daß man die Wahrscheinlichkeiten nicht einzeln betrachten darf, sondern die gesamte Verteilung betrachten muß. In dem von ihm entwickelten Modell werden nicht einzelne Wahrscheinlichkeiten sondern die Verteilungsfunktion transformiert. Das entsprechende Präferenzfunktional für Lotterien mit endlichem Träger lautet in Quiggins Modell

$$V_Q(X) = \sum_{i=1}^{n} u(x_i) \left[q(\sum_{j=1}^{i} p_j) - q(\sum_{j=1}^{i-1} p_j) \right],$$

Hierbei sind die Ergebnisse $x_i, i = 1, \dots n$ so geordnet, daß $u(x_1) \leq \dots u(x_n)$ gilt. Die Funktion $q : [0,1] \to [0,1]$ ist monoton wachsend und erfüllt $q(0) = 0$ und $q(1) = 1$.

3.2 Axiomatische Grundlagen von AU und CEU

In diesem Abschnitt stellen wir nun die Grundzüge des antizipierten Nutzen-Modells und des Choquet-Erwartungsnutzen-Modells dar. Beiden Modellen ist gemeinsam, daß der Entscheidungsträger versucht, den „erwarteten Nutzen“ zu maximieren, wobei dieser „erwartete Nutzen“ nicht wie im Erwartungsnutzenmodell bezüglich eines Wahrscheinlichkeitsmaßes berechnet wird, sondern als Choquet-Integral bezüglich einer Kapazität. Der Unterschied zwischen beiden Modellen liegt darin, daß das AU-Modell Entscheidungen unter Risiko, d.h. bei gegebenem Wahrscheinlichkeitsmaß auf der Menge der Umweltzustände, beschreibt; das CEU-Modell hingegen Entscheidungen unter Unsicherheit.

Definition 3.1 (CEU-Modell) *Sei* $((S, \mathcal{A}), (\mathcal{C}, \mathcal{D}), \mathcal{F}, \preceq)$ *ein Entscheidungsproblem unter Unsicherheit. Existiert eine Nutzenfunktion* $u : \mathcal{C} \to \mathbb{R}$ *und eine Kapazität* $\mu : \mathcal{A} \to [0, 1]$*, so daß die Präferenzrelation* $\preceq$ *des ET durch das CEU-Funktional*

$$CEU : \mathcal{F} \to \overline{\mathbb{R}}, \; f \mapsto \int_S (u \circ f) \, d\mu,$$

repräsentiert wird, so bezeichnen wir den ET als CEU-Maximierer *und das Entscheidungsproblem als* CEU-Modell.

Es existiert eine Vielzahl von Axiomatisierungen von CEU-Modellen, die sich jeweils in den Voraussetzungen an den Zustandsraum und den Konsequenzenraum unterscheiden. Schmeidler (1989) gibt eine Axiomatisierung von CEU im Rahmen des Anscombe-Aumann-Modells, das sowohl subjektive als auch objektive Wahrscheinlichkeiten benutzt. Die objektiven Wahrscheinlichkeiten sind bei Schmeidler nach wie vor additiv, d.h. die Nutzenfunktion u auf dem Konsequenzenraum (= Menge aller Lotterien auf Γ) ist der Erwartungsnutzen der Lotterie. Lediglich die subjektive Bewertung der Ereignisse ist eine Choquet-Kapazität. Die Axiomatisierung von Gilboa (1987) im Rahmen des Savageschen Ansatzes kommt ohne objektive Wahrscheinlichkeiten aus, setzt aber voraus, daß der Zustandsraum S unendlich ist. Wakker (1989a,b) macht die Voraussetzungen, daß S endlich ist und daß der Konsequenzenraum $\mathcal{C}$ ein zusammenhängender, separabler topologischer Raum ist. Wakker (1993) verallgemeinert diese Ergebnisse für beliebige Zustandsräume. Ergebnisse für den etwas allgemeineren algebraischen Ansatz finden sich in Nakamura (1990), Wakker (1991) und Chew und Karni (1991). Allen diesen Axiomensystemen ist gemeinsam, daß Savages surething principle fallengelassen wird. In Anhang B werden vier Axiomensysteme für CEU explizit dargestellt.

Definition 3.2 (AU-Modell) *Sei* $((\mathcal{C}, \mathcal{D}), \mathcal{P}, \preceq)$ *ein Entscheidungsproblem unter Risiko. Existiert eine Nutzenfunktion* $u : \mathcal{C} \to \mathbb{R}$ *und eine Verzerrungsfunktion* $q : [0, 1] \to [0, 1]$*, so daß die Präferenzrelation* $\preceq$ *des ET durch das AU-Funktional*

$$AU : \mathcal{P} \to \overline{\mathbb{R}}, \; P \mapsto \int_{\mathcal{C}} u \, d(q \circ P),$$

repräsentiert wird, so bezeichnen wir den ET als AU-Maximierer *und das Entscheidungsproblem als* AU-Modell.

Das AU-Modell wurde erstmals von Quiggin (1982) entwickelt. In seiner Arbeit wird eine leicht unterschiedliche Darstellung des AU-Funktionals gegeben. Die Verzerrungsfunktion f in Quiggin (1982) ist die duale Verzerrungsfunktion der von uns benutzten Verzerrungsfunktion q. Außerdem setzt Quiggin (1982) noch voraus, daß $q(0.5) = 0.5$ gilt. Dies bedeutet, daß einfache 50-50-Lotterien entsprechend dem EU-Modell bewertet werden. Diese Bedingung wurde jedoch als zu einschränkend empfunden und in späteren Darstellungen (Chew 1989, erste

Version 1985, Chew, Karni und Safra 1987, Segal 1989,1993, Wakker 1994) fallengelassen. Eine Axiomatisierung, die an der Forderung $q(0.5) = 0.5$ festhält ist Quiggin und Wakker (1992). Wakker (1990b) zeigte, daß jede Axiomatisierung eines CEU-Modells zu einer Axiomatisierung eines AU-Modells führt. Ein Spezialfall des AU-Modells ist Yaaris „duale Theorie“ (Yaari 1987). Man erhält die duale Theorie aus dem AU-Modell, wenn die Nutzenfunktion u linear ist. Allen Axiomatisierungen von AU-Modellen ist gemein, daß das Unabhängigkeitsaxiom des von Neumann-Morgenstern-Modells abgeschwächt wird. Statt der Bezeichnung „Antizipierter Nutzen“ („Anticipated Utility“) finden sich in der Literatur auch die Bezeichnungen „Erwartungsnutzen mit rangabhängigen Wahrscheinlichkeiten“ („Expected Utility with Rank Dependent Probabilities“) und „Rangabhängiger Erwartungsnutzen“ („Rank Dependent Expected Utility“). Einen schönen Überblick über die Entwicklung des AU-Modells und über dessen Eigenschaften findet man in Quiggin (1993).

Das CEU-Modell enthält viele der bekannten Entscheidungskriterien als Spezialfälle.

Beispiel 3.1 (Maximin-Prinzip, Wald 1950) *Das Maximin-Prinzip besagt, daß ein ET aus der Menge der möglichen Aktionen diejenige Aktion wählt, für die der minimale Nutzen (über alle Umweltzustände) maximal wird. Ist in einem CEU-Modell*

$$\mu(A) = \begin{cases} 1, & \text{falls } A = S, \\ 0, & \text{sonst,} \end{cases}$$

so ist

$$\int_S (u \circ f)\, d\mu = \inf\{u(f(s)) \mid s \in S\}\,.$$

Ein CEU-Maximierer mit der angegebenen Choquet-Kapazität entscheidet sich also gemäß dem Maximin-Prinzip.

Beispiel 3.2 (Optimismus-Pessimismus-Kriterium, Hurwicz 1951) *Beim Optimismus-Pessimismus-Kriterium wird eine Aktion durch einen gewichteten Durchschnitt von minimal und maximal möglichem Nutzen bewertet. Das Optimismus-Pessimismus-Kriterium bedient sich dabei eines* Pessimismusparameters $\alpha \in [0,1]$*, der den Grad des Pessimismus des ET beschreibt. Der ET wählt aus der Menge der möglichen Aktionen diejenige Aktion, für die*

$$OP_\alpha(f) = \alpha \inf\{u(f(s)) \mid s \in S\} + (1-\alpha)\sup\{u(f(s)) \mid s \in S\}$$

maximal wird. Ist in einem CEU-Modell die Kapazität μ gegeben durch

$$\mu(A) = \begin{cases} 1, & \text{falls } A = S, \\ 0, & \text{falls } A = \emptyset, \\ 1-\alpha, & \text{sonst,} \end{cases}$$

so ist

$$\int_S (u \circ f)\, d\mu = \alpha \inf\{u(f(s)) \mid s \in S\} + (1-\alpha) \sup\{u(f(s)) \mid s \in S\}\,.$$

Ein CEU-Maximierer mit der angegebenen Kapazität entscheidet sich also nach dem Optimismus-Pessimismus-Kriterium.

Beispiel 3.3 (Bayes-Prinzip, SEU-Modell) *Das Bayes-Prinzip entspricht dem SEU-Modell, d.h. ein ET entscheidet sich nach dem Bayes-Prinzip, wenn ein W-Maß P auf $(S, \mathcal{A})$ existiert, so daß der ET aus der Menge der möglichen Aktionen diejenige Aktion wählt, für die der Erwartungswert des Nutzens (bezüglich P) maximal wird. Ist in einem CEU-Modell $\mu = P$ für ein W-Maß P, so ist*

$$\int_S (u \circ f)\, d\mu = E_P(u \circ f)\,.$$

Ein CEU-Maximierer mit der angegebenen Kapazität entscheidet sich also gemäß dem Bayes-Prinzip.

Beispiel 3.4 (Kriterium von Hodges und Lehmann 1952) *Das Kriterium von Hodges und Lehmann stellt eine Kombination aus dem Maximin-Prinzip und dem Bayes-Prinzip dar. Beim Hodges-Lehmann-Prinzip werden die Aktionen durch das Funktional*

$$HL_{P,\lambda}(f) = (1-\lambda) \inf\{u(f(s)) \mid s \in S\} + \lambda E_P(u \circ f)\,,$$

bewertet, wobei $\lambda \in [0,1]$ ist und P ein W-Maß auf $(S, \mathcal{A})$. λ heißt Vertrauensparameter *und kann als Vertrauen darauf interpretiert werden, daß P wirklich die Wahrscheinlichkeiten der Zustände beschreibt. Ist in einem CEU-Modell*

$$\mu(A) = \begin{cases} 1, & \text{falls } A = S, \\ \lambda P(A), & \text{sonst,} \end{cases}$$

so ist

$$\int_S (u \circ f) d\mu = (1-\lambda) \inf\{u(f(s)) \mid s \in S\} + \lambda E_P(u \circ f)\,.$$

Somit ist auch das Hodges-Lehmann-Prinzip ein Spezialfall des CEU-Modells.

Beispiel 3.5 (Vage Vorinformation, Π-Maximin-Prinzip) *Ist in einem Entscheidungsproblem bekannt, daß das W-Maß P, das die Wahrscheinlichkeiten der Zustände beschreibt, zu einer Menge Π von W-Maßen auf $(S, \mathcal{A})$ gehört, so bezeichnen wir Π als* vage Vorinformation. *Ein ET handelt nach dem Π-Maximin-Prinzip, wenn er aus der Menge der möglichen Aktionen diejenige Aktion f wählt, für die*

$$\inf\{E_P(u \circ f) \mid P \in \Pi\}$$

maximal wird. Auch dieses Prinzip steht in enger Verbindung zum CEU-Modell. Allerdings ist das Π-Maximin-Prinzip kein Spezialfall des CEU-Modells. Hat die

Menge Π *jedoch gewisse Eigenschaften, so kann auch dieses Modell als Spezialfall des CEU-Modells dargestellt werden. Der Zusammenhang zwischen dem* Π*-Maximin-Prinzip und dem CEU-Modell wird in Abschnitt 3.3 ausführlich dargestellt.*

Diese Beispiele sind auch in Hinblick auf die Interpretation der Choquet-Kapazität in einem CEU-Modell interessant, da sie zeigen, daß $\mu(A)$ nicht ohne weiteres als die Wahrscheinlichkeit eines Ereignisses oder das Vertrauen des ET in das Eintreten dieses Ereignisses interpretiert werden kann. Ein ET, der nach dem Maximin-Prinzip entscheidet, ist sicher nicht davon überzeugt, daß jedes Ereignis $A \neq S$ die Wahrscheinlichkeit 0 besitzt, sondern er will sich lediglich gegen den schlimmstmöglichen Fall absichern. Die Werte der Kapazität können also lediglich als Gewichte interpretiert werden, die den möglichen Ergebnissen beigemessen werden.

Wir kommen nun noch kurz zur Frage der Eindeutigkeit der Choquet-Kapazität und der Nutzenfunktion. Dazu benötigen wir die folgende Definition:

Definition 3.3 (Null-Ereignis, universelles Ereignis) *Sei* $((S,\mathcal{A}), (\mathcal{C},\mathcal{D}), \mathcal{F}, \preceq)$ *ein CEU-Modell.*

(i) Ein Ereignis $A \in \mathcal{A}$ *heißt* Null-Ereignis*, wenn für alle* $x, y, z \in \mathcal{C}$ *mit* $z \preceq x, y$ *gilt*

$$\begin{bmatrix} z & x \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} z & y \\ A^c & A \end{bmatrix}.$$

(ii) Ein Ereignis $A \in \mathcal{A}$ *heißt* universelles Ereignis*, wenn für alle* $x, y, z \in \mathcal{C}$ *mit* $x, y \preceq z$ *gilt*

$$\begin{bmatrix} x & z \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} y & z \\ A^c & A \end{bmatrix}.$$

Sowohl die Konsequenzen einer Aktion auf einem Null-Ereignis als auch die Konsequenzen außerhalb eines universellen Ereignisses besitzen keine Relevanz für die Bewertung einer Aktion. Man überlegt sich leicht, daß in einem CEU-Modell mit mindestens zwei verschieden präferierten Konsequenzen genau dann ein Ereignis A existiert, das weder Null-Ereignis noch universelles Ereignis ist, wenn ein Ereignis A existiert, für das $0 < \mu(A) < 1$ gilt.

Definition 3.4 (Triviales Modell, degeneriertes Modell) *Ein CEU-Modell heißt* trivial*, wenn* $f \sim g$ *für alle* $f, g \in \mathcal{F}$ *gilt. Ein CEU-Modell heißt* degeneriert*, wenn jedes Ereignis entweder Null-Ereignis oder universelles Ereignis ist. Ein AU-Modell heißt* degeneriert *oder auch* trivial*, wenn* $P \sim Q$ *für alle* $P, Q \in \mathcal{P}$ *gilt.*

Bei AU-Modellen sind also die Begriffe *degeneriert* und *trivial* synonym. Wir fassen die Eindeutigkeitsaussagen der Axiomatisierungen von CEU- und AU-Modellen im folgenden Satz zusammen. Für einen Beweis vergleiche die oben zitierten Axiomatisierungen.

Satz 3.1 (Eindeutigkeitsaussagen) *Die folgenden Eindeutigkeitsaussagen gelten für CEU- und AU-Modelle gemäß Definition 3.1 und Definition 3.2.*

- *In einem trivialem CEU- oder AU-Modell ist μ bzw. q beliebig und u konstant.*
- *In einem degeneriertem, nichttrivialem CEU-Modell ist μ eindeutig und u eindeutig bis auf positive Transformationen.*
- *In einem nichtdegeneriertem, nichttrivialem CEU- oder AU-Modell ist μ bzw. q eindeutig und u eindeutig bis auf positive affine Transformationen.*

3.3 CEU und das Maximin-Prinzip bei vager Vorinformation

Das CEU-Modell steht in enger Beziehung zum sogenannten Π-Maximin-Prinzip. Diesem Modell zufolge handelt der ET als EU-Maximierer. Allerdings ist das Wahrscheinlichkeitsmaß P auf der Menge der Umweltzustände nicht exakt bekannt, sondern es ist nur bekannt, daß P zu einer Menge Π von Wahrscheinlichkeitsmaßen gehört. Auf diese Situation wird nun das Maximin-Prinzip angewendet, d.h. für jede Aktion des ET wird der bezüglich Π minimale Erwartungsnutzen bestimmt und dann die Aktion gewählt, die diesen maximiert. Die Präferenz wird also durch das Funktional

$$\begin{aligned} V_\Pi : \mathcal{F} &\to \overline{\mathbb{R}} \\ f &\mapsto \min\{\int_S (u \circ f)\, dP \mid P \in \Pi\} \end{aligned}$$

repräsentiert.

Entscheidungsprobleme mit vager Vorinformation wurden unter anderem untersucht von Fishburn (1964, 1965), Schneeweiß (1964), Kofler (1974), Bühler (1975), Kofler und Menges (1976) und Weber (1987). Zur Verwendung vager Vorinformation in der statistischen Qualitätskontrolle vergleiche z.B. Krumbholz (1982). Eine Axiomatisierung des Maximinprinzips bei vager Vorinformation findet sich in Gilboa und Schmeidler (1989).

Ist die Menge der Umweltzustände endlich, so kann die Menge Π z.B. durch lineare Ungleichungen über die Wahrscheinlichkeiten $p_i = P(\{s_i\})$, $i = 1, \ldots, n$, gegeben sein. Wir sprechen dann von *linearer partieller Information* (LPI) über die Wahrscheinlichkeiten. Im einfachsten Fall ist die Information von der Form „Zustand s_i ist mindestens so wahrscheinlich wie Zustand s_j“, d.h. $p_i \geq p_j$, für gewisse Zustände s_i, s_j. Eine andere einfache Möglichkeit ist, wenn bekannt ist, daß die Wahrscheinlichkeiten der Zustände in bestimmten Intervallen liegen, d.h. wenn die Information von der Form $p_i \in [L_i, U_i]$, $i = 1, \ldots, n$, ist. In diesem Fall spricht man auch davon, daß *Intervallwahrscheinlichkeiten* gegeben sind.

Wir wollen diese Situation zunächst an einem Beispiel verdeutlichen und greifen dazu wieder auf das Ellsberg-Paradox zurück.

Beispiel 3.6 (Ellsberg-Paradox, vgl. Beispiel 1.1) *Wir betrachten wieder das Ellsberg-Paradox. Da sich in der Urne eine genau definierte Anzahl von roten, gelben und blauen Kugeln befindet, gibt es ein „objektives“ W-Maß P, das die Ziehung einer Kugel aus der Urne beschreibt. Dieses W-Maß ist allerdings nicht exakt bekannt. Wir wissen nur, daß das W-Maß P zur Menge*

$$\Pi = \{P | P \ \textit{W-Maß auf } S, \ P(\{r\}) = \frac{1}{3}, P(\{g, b\}) = \frac{2}{3}\}$$

gehört. Entscheidung nach dem Π-Maximin-Prinzip bedeutet nun, daß für jede der Aktionen der bezüglich Π minimale Erwartungsnutzen berechnet wird und dann anschließend die Aktion gewählt wird, für die dieser maximal wird.

Der Zusammenhang zwischen CEU und Π-Maximin-Prinzip wird durch das Folgende motiviert.

Definition 3.5 *Sei $(\Omega, \mathcal{A})$ ein Meßraum und Π eine Menge von Wahrscheinlichkeitsmaßen auf $(\Omega, \mathcal{A})$. Dann wird die Mengenfunktion*

$$\begin{aligned} \mu^* \ : \ & \mathcal{A} \to \mathbb{R}, \\ & A \mapsto \mu^*(A) = \sup\{P(A) \,|\, P \in \Pi\}, \end{aligned}$$

als obere Wahrscheinlichkeit *und die Mengenfunktion*

$$\begin{aligned} \mu_* \ : \ & \mathcal{A} \to \mathbb{R}, \\ & A \mapsto \mu_*(A) = \inf\{P(A) \,|\, P \in \Pi\}, \end{aligned}$$

als untere Wahrscheinlichkeit *bezeichnet.*

Man überlegt sich leicht, daß μ^* und μ_* Kapazitäten sind. Ferner sind obere und untere Wahrscheinlichkeit dual zueinander, d.h. es gilt $\mu_*^D = \mu^*$.

Im Ellsberg-Beispiel wäre somit auch das folgende Vorgehen denkbar:

(i) Berechnung der unteren Wahrscheinlichkeit μ_* zu Π.

(ii) Berechnung des Choquet-Erwartungsnutzen $CEU(f) = \int u \circ f d\mu_*$ für jede Aktion f.

(iii) Entscheidung für die Aktion f, die $CEU(f)$ maximiert.

Es stellt sich natürlich die Frage, ob die gerade formulierte Entscheidungsregel und das Π-Maximin-Prinzip zur selben Entscheidung führen. Leider ist dies i.a. nicht der Fall. Der folgende Satz gibt Bedingungen an, unter denen Π-Maximin-Prinzip und CEU-Prinzip zur selben Entscheidung führen. Satz 3.2 wurde unter der zusätzlichen Voraussetzung, daß Ω ein separabler, vollständig metrisierbarer Raum ist, von Huber und Strassen (1973) bewiesen. Schmeidler (1986) beweist einen analogen Satz. Anstatt σ-additiver Maße betrachtet Schmeidler (1986) jedoch die Menge aller *endlich* additiven W-Maße P, für die $P(A) \leq \mu(A)$ für alle $A \in \mathcal{A}$ gilt. Der Beweis von Satz 3.2 findet sich in Abschnitt 3.5.

Satz 3.2 *Sei $(\Omega, \mathcal{A})$ ein Meßraum und $\mu : \mathcal{A} \to [0,1]$ eine Kapazität. Sei weiterhin $\Pi^\mu = \{P : \mathcal{A} \to [0,1] \,|\, P$ ist W-Maß, $P(A) \leq \mu(A)$ für alle $A \in \mathcal{A}\}$ und $\mathcal{L}^1$ die Menge aller Funktionen, die bezüglich μ und μ^D integrierbar sind. Dann sind die folgenden Aussagen äquivalent:*

(i) μ ist stark subadditiv und von oben stetig.

(ii) $\int f d\mu = \max\{\int f dP \,|\, P \in \Pi^\mu\}$ für alle $f \in \mathcal{L}^1$.

(iii) Für jedes $f \in \mathcal{L}^1$ gibt es ein $P_f \in \Pi^\mu$, so daß $\mu(f > t) = P_f(f > t)$ für alle $t \in \mathbb{R}$.

Insbesondere gilt, falls eine der Bedingungen erfüllt ist:

$$\mu(A) = \max\{P(A) \,|\, P \in \Pi^\mu\} \text{ für alle } A \in \mathcal{A}.$$

Beweis: Abschnitt 3.5

Durch Übergang zur dualen Kapazität erhält man sofort das folgende Korollar:

Korollar 3.1 *Sei $(\Omega, \mathcal{A})$ ein Meßraum und $\mu : \mathcal{A} \to [0,1]$ eine Kapazität. Sei weiterhin $\Pi_\mu = \{P : \mathcal{A} \to [0,1] \,|\, P$ ist W-Maß, $P(A) \geq \mu(A)$ für alle $A \in \mathcal{A}\}$ und $\mathcal{L}^1$ die Menge aller Funktionen, die bezüglich μ und μ^D integrierbar sind. Dann sind die folgenden Aussagen äquivalent:*

(i) μ ist stark superadditiv und von unten stetig.

(ii) $\int f d\mu = \min\{\int f dP \,|\, P \in \Pi_\mu\}$ für alle $f \in \mathcal{L}^1$.

(iii) Für jedes $f \in \mathcal{L}^1$ gibt es ein $P_f \in \Pi_\mu$, so daß $\mu(f > t) = P_f(f > t)$ für alle $t \in \mathbb{R}$.

Insbesondere gilt, falls eine der Bedingungen erfüllt ist:

$$\mu(A) = \min\{P(A) \,|\, P \in \Pi_\mu\} \text{ für alle } A \in \mathcal{A}.$$

Ist also die Situation des Korollars gegeben, so kann die Entscheidung nach dem Π-Maximin-Prinzip durchgeführt werden, ohne für jedes W-Maß $P \in \Pi$ den Erwartungswert zu berechnen. Es muß lediglich der Choquet-Erwartungsnutzen bezüglich der Kapazität μ_* berechnet werden.

Die untere Wahrscheinlichkeit einer Menge Π von W-Maßen ist zwar immer superadditiv[1], i.a. jedoch nicht stark superadditiv. Ferner kann es, selbst wenn μ_* stark superadditiv ist, vorkommen, daß die Menge Π eine echte Teilmenge von Π_μ ist. Wir geben für beide Situationen Beispiele, die Huber und Strassen (1973) entnommen sind. In beiden Beispielen gibt es eine Funktion f, für die $\int f d\mu < \min\{\int f dP \mid P \in \Pi_\mu\}$ gilt.

Beispiel 3.7 **(μ nicht stark superadditiv)** *Sei* $\Omega = \{1,2,3,4\}$, *und seien* $P_0 = (0.5, 0.2, 0.2, 0.1)$, $P_1 = (0.6, 0.1, 0.1, 0.2)$, *sowie* Π *die konvexe Hülle von* P_0 *und* P_1. *Für* $A = \{1,2\}$ *und* $B = \{1,3\}$ *ist*

$$\mu_*(A \cup B) + \mu_*(A \cap B) = 0.8 + 0.5 < 0.7 + 0.7 = \mu_*(A) + \mu_*(B) \ ,$$

und somit ist μ_* *nicht stark superadditiv. Für* $f = (2,1,1,0)$ *ist* $\int f d\mu_* = 1.3$ *aber* $\min\{\int f dP | P \in \Pi\} = 1.4$.

Beispiel 3.8 **(μ stark superadditiv, aber $\Pi \subsetneq \Pi_\mu$)** *Seien* $\Omega = \{1,2,3\}$, $P_0 = (\frac{1}{2}, \frac{1}{2}, 0)$ *und* $P_1 = (\frac{4}{6}, \frac{1}{6}, \frac{1}{6})$, *sowie* Π *die konvexe Hülle von* P_0 *und* P_1. *Dann ist*

$$\Pi = \left\{ \left(\frac{3+\lambda}{6}, \frac{3-2\lambda}{6}, \frac{\lambda}{6} \right) \mid \lambda \in [0,1] \right\} \ .$$

Die untere Wahrscheinlichkeit μ_* *ist gegeben durch*

A	$\emptyset$	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	Ω
$\mu_*(A)$	0	1/2	1/6	0	5/6	1/2	1/3	1

und somit stark superadditiv. Die Menge Π_{μ_*} *ist gegeben durch*

$$\Pi_{\mu_*} = \left\{ \left(\frac{3+t}{6}, \frac{3-s-t}{6}, \frac{s}{6} \right) \mid s,t \in [0,1] \right\} \ .$$

Also ist Π *eine echte Teilmenge von* Π_{μ_*}, *wie behauptet. Wir zeigen noch, daß es Funktionen* f *gibt, für die* $\int f d\mu_* < \min\{\int f dP | P \in \Pi\}$ *gilt. Sei dazu* f *gegeben durch* $f = (1, 0, -1)$. *Dann ist* $\int f d\mu_* = \frac{1}{3}$ *aber* $\min\{\int f dP | P \in \Pi\} = \frac{1}{2}$.

Eine einfache und vom praktischen Standpunkt interessante Situation, in der der minimale Erwartungsnutzen durch ein Choquet-Integral ausgedrückt werden kann, liegt vor, wenn Intervallwahrscheinlichkeiten gegeben sind. Wir präzisieren dies im folgenden Satz.

[1]Eine Kapazität μ auf einem Mengensystem $\mathcal{A}$ heißt *superadditiv*, falls für alle $A, B \in \mathcal{A}$ mit $A \cap B = \emptyset$ gilt $\mu(A \cup B) \geq \mu(A) + \mu(B)$.

Satz 3.3 *Sei* $\Omega = \{\omega_1, \ldots, \omega_n\}$, *und sei* Π *die Menge aller W-Maße* P *auf* $\mathcal{P}(\Omega)$ *mit*

$$p_i = P(\{\omega_i\}) \in [L_i, U_i], \quad 1 \leq i \leq n,$$

wobei $0 \leq L_i \leq U_i \leq 1$, $i = 1, \ldots, n$. *Außerdem gelte* $\sum_{i=1}^n L_i \leq 1$ *und* $\sum_{i=1}^n U_i \geq 1$. *Die Mengenfunktion* $\mu : \mathcal{P}(\Omega) \to [0,1]$ *sei definiert durch*

$$\mu(A) = \begin{cases} \sum_{i \in A} L_i, & \text{falls } \sum_{i \in A} L_i + \sum_{i \notin A} U_i \geq 1, \\ 1 - \sum_{i \notin A} U_i, & \text{falls } \sum_{i \in A} L_i + \sum_{i \notin A} U_i < 1. \end{cases}$$

Dann ist μ *eine Kapazität und für alle* $f : \Omega \to \mathbb{R}$ *gilt*

$$\int f d\mu \;=\; \min\left\{ \int f dP \,|\, P \in \Pi \right\}.$$

Beweis: Abschnitt 3.5

3.4 Erklärung der Paradoxa

Wir wollen nun noch kurz darauf eingehen, wie im Rahmen von AU und CEU die in Abschnitt 1.3 betrachteten Phänomene erklärt werden können. Zur Erklärung des common consequence effect und des common ratio effect in AU und der dualen Theorie vgl. auch Segal (1987), Yaari 1987 und Röell (1987).

Beispiel 3.9 (Common consequence effect, vgl. Beispiel 1.3) *Beim common consequence effect wird das gleichzeitige Auftreten der Präferenzen*

$$\begin{bmatrix} x_1 & x_2 & x_3 \\ p_1 & p_2 & p_3 \end{bmatrix} \prec \begin{bmatrix} x_2 \\ 1 \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} x_1 & x_3 \\ p_1 + p_2 & p_3 \end{bmatrix} \succ \begin{bmatrix} x_1 & x_2 \\ p_2 & p_1 + p_3 \end{bmatrix}$$

beobachtet, wobei $x_1 \prec x_2 \prec x_3$ *gilt. In EU sind diese Präferenzen nicht erklärbar (vgl. Beispiel 1.3). Mit* $u(x_1) = 0$, $u(x_3) = 1$ *sind in AU obige Präferenzen äquivalent zu*

$$u(x_2)[q(p_2 + p_3) - q(p_3)] + q(p_3) < u(x_2) \quad \text{und} \quad q(p_3) > u(x_2) q(p_1 + p_3),$$

bzw. zu

$$\frac{q(p_3)}{1 - q(p_2 + p_3) + q(p_3)} < u(x_2) < \frac{q(p_3)}{q(p_1 + p_3)}.$$

Der common consequence effect ist also in AU erklärbar, wenn das Intervall für $u(x_2)$ *nicht leer ist. Dies ist offensichtlich genau dann der Fall, wenn* $q(p_1 + p_3) < 1 - q(p_2 + p_3) + q(p_3)$ *ist, bzw. wenn* $q(p_1 + p_2 + p_3) - q(p_2 + p_3) > q(p_1 + p_3) - q(p_3)$ *ist. Diese Ungleichung wird insbesondere von allen konvexen Verzerrungsfunktionen erfüllt. In AU kann der common consequence effect also dadurch erklärt werden, daß man eine konvexe Verzerrungsfunktion unterstellt. Insbesondere ist in AU auch das Allais-Paradox durch eine konvexe Verzerrungsfunktion erklärbar. Dies wurde bereits in Quiggin (1985) gezeigt.*

Beispiel 3.10 (Common ratio effect, vgl. Beispiel 1.4) *Beim common ratio effect wird das gleichzeitige Auftreteten der Präferenzen*

$$\begin{bmatrix} 0 & x_1 \\ 1-p_1 & p_1 \end{bmatrix} \succ \begin{bmatrix} 0 & x_2 \\ 1-p_2 & p_2 \end{bmatrix} \quad \textit{und} \quad \begin{bmatrix} 0 & x_1 \\ 1-\lambda p_1 & \lambda p_1 \end{bmatrix} \succ \begin{bmatrix} 0 & x_2 \\ 1-\lambda p_2 & \lambda p_2 \end{bmatrix},$$

beobachtet, wobei $\lambda < 1$ und $x_1, x_2 > 0$ sind. Dieses Verhalten ist in EU nicht erklärbar. Im Rahmen von AU sind obige Präferenzen äquivalent zu

$$u(x_1)q(p_1) > u(x_2)q(p_2) \quad \textit{und} \quad u(x_1)q(\lambda p_1) < u(x_2)q(\lambda p_2).$$

Umformung und Kombination der Ungleichungen ergibt

$$\frac{q(\lambda p_1)}{q(\lambda p_2)} < \frac{u(x_2)}{u(x_1)} < \frac{q(p_1)}{q(p_2)}.$$

Der common ratio effect kann also mit jeder Verzerrungsfunktion erklärt werden, für die das oben angegebene Intervall für $u(x_2)/u(x_1)$ nicht leer ist. Dies ist offensichtlich dann erfüllt, wenn $q(\alpha p)/q(p)$ für jedes $\alpha > 1$ monoton wachsend ist. Segal (1987) zeigte, daß diese Bedingung äquivalent dazu ist, daß die Elastizität $(pq'(p))/q(p)$ monoton wachsend ist. Beispiele für solche Funktionen sind z.B. $q(p) = p/(2-p)$ oder $q(p) = 1-(1-p)^\alpha$ für $0 < \alpha < 1$.

Beispiel 3.11 (Ellsberg-Paradox, vgl. Beispiel 1.1) *Im Ellsberg-Paradox werden einem ET die Aktionenpaare*

$$f_1 = \begin{bmatrix} 0 & 100 \\ \{b,g\} & \{r\} \end{bmatrix} \quad \textit{und} \quad \begin{bmatrix} 0 & 100 \\ \{b,r\} & \{g\} \end{bmatrix} = f_2,$$

sowie

$$g_1 = \begin{bmatrix} 0 & 100 \\ \{g\} & \{b,r\} \end{bmatrix} \quad \textit{und} \quad \begin{bmatrix} 0 & 100 \\ \{r\} & \{b,g\} \end{bmatrix} = g_2,$$

präsentiert. Typischerweise lautet die Entscheidung $f_1 \succ f_2$ und $g_1 \prec g_3$. Wählt man die Nutzenfunktion u so, daß $u(0) = 0$ und $u(100) = 1$ gilt, so erhält man für die Choquet-Erwartungsnutzen der Aktionen:

$$\begin{array}{ll} CEU(f_1) = \mu(r), & CEU(f_2) = \mu(g), \\ CEU(g_1) = \mu(b,r), & CEU(g_2) = \mu(b,g) \end{array}$$

Die beobachteten Präferenzen gelten also genau dann, wenn $\mu(r) > \mu(g)$ und $\mu(b,r) < \mu(b,g)$ ist. Da μ als Choquet-Kapazität nicht notwendig additiv sein muß, sind diese Bedingungen in CEU erfüllbar.

3.5 Beweise

Bevor wir mit dem Beweis von Satz 3.2 beginnen, benötigen wir noch einige Vorüberlegungen. Die im Beweis benutzten Sätze von Daniell-Stone und Hahn-Banach sind im Anhang zitiert.

Lemma 3.1 *Sei μ eine stark subadditive Kapazität auf einem Meßraum $(\Omega, \mathcal{A})$. Dann gelten folgende Aussagen:*

(i) Ist μ stetig von oben, so ist μ auch stetig von unten.

(ii) Für alle $A \in \mathcal{A}$ ist $\mu^D(A) \leq \mu(A)$.

Beweis: Teil (i): Sei zunächst μ stetig von oben und $(A_n)_{n\in\mathbb{N}}$ eine Folge von Mengen in $\mathcal{A}$ mit $A_n \nearrow A$. Dann ist $(A \setminus A_n) \searrow \emptyset$, und es gilt $\lim_{n\to\infty} \mu(A\setminus A_n) = 0$. Wegen der Subadditivität ist $\mu(A) \leq \mu(A_n) + \mu(A \setminus A_n)$ für alle $n \in \mathbb{N}$, also $\mu(A_n) \geq \mu(A) - \mu(A \setminus A_n)$ für alle $n \in \mathbb{N}$. Grenzübergang ergibt dann $\lim_{n\to\infty} \mu(A_n) \geq \mu(A)$ und wegen der Monotonie von μ ist $\lim_{n\to\infty} \mu(A_n) = \mu(A)$, d.h. μ ist stetig von unten.

Teil (ii): Sei $A \in \mathcal{A}$. Dann ist wegen der Subadditivität $1 = \mu(\Omega) \leq \mu(A) + \mu(A^c)$ bzw. $\mu(A) \geq 1 - \mu(A^c) = \mu^D(A)$. ■

Die Umkehrung von Teil (i) des Lemmas gilt i.a. nicht, wie das folgende Beispiel zeigt.

Beispiel 3.12 (μ stark subadditiv und stetig von unten, aber nicht stetig von oben) *Sei $\Omega = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(\Omega)$ und μ gegeben durch*

$$\mu(A) = \begin{cases} \frac{1}{2} + \frac{1}{2}\sum_{i\in A} 2^{-i}, & \textit{falls } A \neq \emptyset \\ 0, & \textit{falls } A = \emptyset. \end{cases}$$

Dann ist μ eine von unten stetige, stark subadditive Kapazität, die nicht stetig von oben ist.

Lemma 3.2 *Sei μ eine stark subadditive und von oben stetige Kapazität auf einem Meßraum $(\Omega, \mathcal{A})$. Dann gilt für jede absteigende Folge $(f_n)_{n\in\mathbb{N}}$ von $\mathcal{A}$-meßbaren Funktionen mit $f_n \searrow 0$ und $\int f_i d\mu < \infty$ für mindestens ein $i \in \mathbb{N}$:*

$$\lim_{n\to\infty} \int f_n \, d\mu \; = \; 0.$$

Beweis: Die Behauptung folgt aus Proposition 8.5., Proposition 8.8. und Theorem 8.9. in Denneberg (1992).

Lemma 3.3 *Sei μ eine stark subadditive Kapazität auf einem Meßraum $(\Omega, \mathcal{A})$, und sei*

$$\mathcal{L}^1 = \left\{ f : \Omega \to \mathbb{R} \,|\, f \ \mathcal{A} - \textit{meßbar}, \ \int f d\mu, \ \int f d\mu^D \in \mathbb{R} \right\}$$

Dann ist $\mathcal{L}^1$ ein Vektorverband, der die konstanten Funktionen enthält.

Beweis: Denneberg (1992), Proposition 9.5.

Beweis von Satz 3.2:

(i)$\Rightarrow$(iii): Sei zunächst μ stark subadditiv und stetig von oben. Wir definieren auf $\mathcal{L}^1$ die Funktionale I_μ und I_μ^+ durch $I_\mu(f) = \int f d\mu$ bzw. $I_\mu^+(f) = \int f^+ d\mu$. Dann sind wegen Satz 2.1 I_μ und I_μ^+ positive, sublineare Funktionale auf $\mathcal{L}^1$, und es ist $I_\mu \leq I_\mu^+$.

Sei nun $f \in \mathcal{L}^1$. Wir konstruieren ein Maß P_f, wie in (iii) gefordert. Sei zunächst f strikt positiv. Wir definieren ein Funktional l_f auf $M = \{\lambda f \mid \lambda \in \mathbb{R}\}$ durch $l_f(\lambda f) = \lambda I_\mu(f)$. Dann ist l_f linear, und es gilt $l_f \leq I_\mu^+$ auf M. Nach dem Satz von Hahn-Banach kann l_f zu einem linearen Funktional $L_f : \mathcal{L}^1 \to \mathbb{R}$ mit $L_f \leq I_\mu^+$ fortgesetzt werden. Für beliebiges $g \in \mathcal{L}^1_+$ gilt

$$-L_f(g) = L_f(-g) \leq I_\mu^+(-g) = 0 \; .$$

Somit ist $L_f(g) \geq 0$ für $g \in \mathcal{L}^1_+$, d.h. L_f ist positiv und linear. Da μ stetig von oben ist, gilt nach Lemma 3.2 für jede Folge $(u_n)_{n\in\mathbb{N}} \in \mathcal{L}^1$ mit $u_n \searrow 0$ auch $\lim I_\mu^+(u_n) = 0$ und damit auch $\lim L_f(u_n) = 0$. Außerdem ist nach Lemma 3.3 $\mathcal{L}^1$ ein Vektorverband, der die konstanten Funktionen enthält. Damit sind die Voraussetzungen des Satzes von Daniell-Stone erfüllt. Somit gibt es genau ein Maß P_f auf $\mathcal{A}_{\mathcal{L}^1}$ mit

$$\int g \, dP_f \;=\; L_f(g) \quad \text{für alle } g \in \mathcal{L}^1.$$

Man überlegt sich leicht, daß $\mathcal{A}_{\mathcal{L}^1} = \mathcal{A}$ ist.

Wir zeigen, daß P_f in Π^μ liegt. Zunächst ist für jedes $A \in \mathcal{A}$

$$P_f(A) \;=\; \int 1_A dP_f \;=\; L_f(1_A) \;\leq\; I_\mu(1_A) \;=\; \mu(A) \; .$$

Um zu zeigen, daß $P_f \in \Pi^\mu$ ist, müssen wir also nur noch zeigen, daß $P_f(\Omega) = 1$ ist. Es ist

$$\int f dP_f \;=\; L_f(f) \;=\; l_f(f) \;=\; I_\mu(f) \;=\; \int f d\mu$$

und somit

$$\int_0^\infty P_f(f > t)dt \;=\; \int_0^\infty \mu(f > t)dt \; .$$

Da $P_f(A) \leq \mu(A)$ für alle $A \in \mathcal{A}$ ist, muß $P_f(f > t) = \mu(f > t)$ für fast alle $t \in \mathbb{R}_+$ gelten, da andernfalls das linke Integral echt kleiner als das rechte Integral wäre. Wir zeigen, daß dies sogar für alle $t \in \mathbb{R}$ gilt. Sei also $t \in \mathbb{R}$. Es gibt eine Folge $(t_n)_{n\in\mathbb{N}}$ mit $t_n \searrow t$ und $P_f(f > t_n) = \mu(f > t_n)$ für alle $n \in \mathbb{N}$. Da $\{f > t_n\} \nearrow \{f > t\}$ und da μ nach Lemma 3.1 auch stetig von unten ist, folgt $P_f(f > t) = \mu(f > t)$. Also ist $P_f(f > t) = \mu(f > t)$ für alle $t \in \mathbb{R}$. Insbesondere erhalten wir für $t = 0$, da f strikt positiv ist $P_f(\Omega) = 1$, d.h. P_f liegt in Π^μ. Für strikt positives f ist also (iii) gezeigt.

Sei nun $f \in \mathcal{L}^1$ beliebig. Dann ist $\tilde{f}$ mit $\tilde{f}(\omega) = \frac{\pi}{2} + \arctan(f(\omega))$ strikt positiv, und nach dem eben gezeigten gibt es ein Maß $P_{\tilde{f}} \in \Pi^\mu$ wie in (iii) gefordert. Dann ist aber auch $P_{\tilde{f}}(f > t) = \mu(f > t)$ für alle $t \in \mathbb{R}$. Somit ist $(i) \Rightarrow (iii)$ gezeigt.

***(iii)*⇒*(ii)*:** Zu $f \in \mathcal{L}^1$ wähle P_f wie in (iii). Dann gilt $\int f d\mu = \int f dP_f$, und für alle anderen $P \in \Pi^\mu$ gilt $\int f d\mu \geq \int f dP$. Daher ist $\int f d\mu = \max\{\int f dP \mid P \in \Pi^\mu\}$ wie verlangt.

***(ii)*⇒*(i)*:** Für $A, B \in \mathcal{A}$ gilt

$$\begin{aligned}
\mu(A \cup B) + \mu(A \cap B) &= \int (1_{A \cup B} + 1_{A \cap B})\, d\mu \\
&= \int (1_A + 1_B)\, d\mu \\
&= \max\{\int (1_A + 1_B)\, dP \mid P \in \Pi^\mu\} \\
&= \max\{P(A) + P(B) \mid P \in \Pi^\mu\} \\
&\leq \max\{P(A) \mid P \in \Pi^\mu\} + \max\{P(B) \mid P \in \Pi^\mu\} \\
&= \mu(A) + \mu(B)
\end{aligned}$$

Also ist μ stark subadditiv. Sei jetzt $(A_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{A}$ mit $A_n \searrow A$. Wir setzen in (ii) für f $\sum_{i=1}^\infty 2^{-i} 1_{A_i}$. Dann gibt es ein W-Maß $P \in \Pi^\mu$ mit $\int \sum_{i=1}^\infty 2^{-i} 1_{A_i} d\mu = \int \sum_{i=1}^\infty 2^{-i} 1_{A_i} dP$. Wie im Beweis von $(i) \Rightarrow (ii)$ folgert man, daß $\mu(A_i) = P(A_i)$ für alle $i \in \mathbb{N}$. Damit erhält man

$$\lim_{i \to \infty} \mu(A_i) = \lim_{i \to \infty} P(A_i) = P(A) \leq \mu(A) ,$$

und wegen der Monotonie von μ folgt $\lim_{i \to \infty} \mu(A_i) = \mu(A)$. Somit ist auch $(ii) \Rightarrow (i)$ gezeigt.

Der Zusatz folgt, indem man in (ii) für f Indikatorfunktionen einsetzt. ■

Beweis von Satz 3.3:

Wir zeigen zunächst, daß μ eine Kapazität ist. $\mu(\emptyset) = 0$ und $\mu(\Omega) = 1$ ist klar. Sei nun $A \subset B$. $\mu(A) \leq \mu(B)$ ist klar, falls für A und B jeweils der gleiche Fall aus der Definition von μ zur Anwendung kommt. Wir müssen also nur den Fall untersuchen, in dem $\sum_{i \in A} L_i + \sum_{i \notin A} U_i \geq 1$ und $\sum_{i \in B} L_i + \sum_{i \notin B} U_i < 1$ ist. Dann ist aber

$$\mu(B) = 1 - \sum_{i \notin B} U_i > \sum_{i \in B} L_i \geq \sum_{i \in A} L_i = \mu(A) ,$$

und somit ist auch die Monotonie von μ gezeigt.

Wir zeigen nun, daß $\mu(A) \leq P(A)$ für alle $A \subset \Omega$, $P \in \Pi$. Ist $\mu(A) = \sum_{i \in A} L_i$, so ist alles klar. Im Fall $\mu(A) = 1 - \sum_{i \notin A} U_i$ folgt die Behauptung aus $\mu(A) =$

$1 - \sum_{i \notin A} U_i \leq 1 - P(A^c) = P(A)$ für alle $P \in \Pi$. Hieraus folgt nun aber sofort, daß $\int f d\mu \leq \inf\{\int f dP \mid P \in \Pi\}$.

Wir konstruieren jetzt zu jeder Funktion $f : \Omega \to \mathbb{R}$ ein W-Maß $P_f \in \Pi$, so daß $\int f d\mu = \int f dP_f$. Sei also f beliebig. Wir nehmen o.B.d.A. an, daß $\Omega = \{1, \ldots, n\}$ und $f(1) \leq f(2) \leq \ldots \leq f(n)$. Wir definieren P_f durch

$$P_f(\{i\}) = \mu(\{i, \ldots, n\}) - \mu(\{i+1, \ldots, n\}), \quad i = 1, \ldots, n\,.$$

Zunächst ist zu zeigen, daß P_f in Π liegt. Wir unterscheiden vier Fälle:

Fall 1: $\sum_{k=1}^{i-1} U_k + \sum_{k=i}^{n} L_k \geq 1$ und $\sum_{k=1}^{i} U_k + \sum_{k=i+1}^{n} L_k \geq 1$. Wir erhalten dann

$$P_f(\{i\}) = \sum_{k=i}^{n} L_k - \sum_{k=i+1}^{n} L_k \;=\; L_i\ .$$

Fall 2: $\sum_{k=1}^{i-1} U_k + \sum_{k=i}^{n} L_k < 1$ und $\sum_{k=1}^{i} U_k + \sum_{k=i+1}^{n} L_k < 1$. Dann ist

$$P_f(\{i\}) = \left(1 - \sum_{k=1}^{i-1} U_k\right) - \left(1 - \sum_{k=1}^{i} U_k\right) \;=\; U_i\ .$$

Fall 3: $\sum_{k=1}^{i-1} U_k + \sum_{k=i}^{n} L_k < 1$ und $\sum_{k=1}^{i} U_k + \sum_{k=i+1}^{n} L_k \geq 1$. In diesem Fall ist

$$P_f(\{i\}) = 1 - \sum_{k=1}^{i-1} U_k - \sum_{k=i+1}^{n} L_k\ .$$

Aus $\sum_{k=i}^{n} L_k + \sum_{k=1}^{i-1} U_k < 1$ folgt nun aber

$$P_f(\{i\}) = 1 - \sum_{k=1}^{i-1} U_k - \sum_{k=i+1}^{n} L_k \;> L_i,$$

und aus $\sum_{k=i+1}^{n} L_k + \sum_{k=1}^{i} U_k \geq 1$ folgt

$$P_f(\{i\}) = 1 - \sum_{k=1}^{i-1} U_k - \sum_{k=i+1}^{n} L_k \;\leq U_i\ .$$

Somit ist auch in diesem Fall $P_f(\{i\}) \in [L_i, U_i]$.

Fall 4: $\sum_{k=1}^{i-1} U_k + \sum_{k=i}^{n} L_k \geq 1$ und $\sum_{k=1}^{i} U_k + \sum_{k=i+1}^{n} L_k < 1$. Wegen

$$\Big(\sum_{k=1}^{i} U_k + \sum_{k=i+1}^{n} L_k\Big) - \Big(\sum_{k=1}^{i-1} U_k + \sum_{k=i}^{n} L_k\Big) = U_i - L_i \geq 0$$

kann dieser Fall nicht eintreten.

Aus den vier Fällen ergibt sich, daß $P_f \in \Pi$. Wegen der Definition von P_f ist dann

$$\begin{aligned}
\int f d\mu &= \sum_{i=1}^{n} f(i)[\mu(\{i,\ldots,n\}) - \mu(\{i+1,\ldots,n\})] \\
&= \sum_{i=1}^{n} f(i)P_f(\{i\}) \\
&= \int f dP_f \\
&\geq \min\left\{\int f dP \,|\, P \in \Pi\right\}.
\end{aligned}$$

Da bereits gezeigt wurde, daß $\int f d\mu \leq \inf\{\int f dP \,|\, P \in \Pi\}$ ist, folgt hieraus die Behauptung. ∎

Kapitel 4

Stochastische Dominanz von Kapazitäten

Resultate über stochastische Dominanz zwischen Wahrscheinlichkeitsverteilungen erlauben es in EU und SEU Entscheidungen zwischen Alternativen zu treffen, ohne daß die Nutzenfunktion vollständig bekannt sein muß. Es reicht vielmehr aus zu wissen, daß die Nutzenfunktion ganz bestimmte qualitative Eigenschaften besitzt, wie z.B. Monotonie, Konkavität oder ähnliches. Neben der Entscheidungstheorie wird stochastische Dominanz aber auch in vielen anderen Gebieten angewendet, so z.B. in der Bedienungstheorie, Zuverlässigkeitstheorie, Simulation, stochastischer Programmierung und Finanzplanung. Vergleiche dazu Mosler und Scarsini (1991a).

In diesem Kapitel untersuchen wir stochastische Dominanz von Kapazitäten. Wir verallgemeinern den Begriff der stochastischen Dominanz von W-Maßen auf Choquet-Kapazitäten und untersuchen, welche der Resultate in EU auf den allgemeineren Fall der Dominanz von Kapazitäten übertragen werden können. Im zweiten Abschnitt betrachten wir speziell Mengen-Dominanzrelationen. Dies sind Dominanzrelationen die durch Ungleichungen der Kapazitäten auf bestimmten Teilmengen des Raumes charakterisiert werden können. Bei der Untersuchung von Mengen-Dominanzrelationen spielt der Begriff der Quasi-Meßbarkeit (siehe Definition 2.6) eine zentrale Rolle. Sind die Kapazitäten auf einem schwach geordnetem Raum definiert und sind alle betrachteten Nutzenfunktionen komonoton, so existiert nach Satz 2.4 ein assoziiertes W-Maß. Das bedeutet, daß für komonotone Klassen von Nutzenfunktion die Choquet-Integration als Integration bezüglich eines σ-additiven Maßes dargestellt werden kann. In diesem Fall lassen sich Charakterisierungen von Dominanzrelationen zwischen W-Maßen leicht auf den Fall der Dominanz von Kapazitäten verallgemeinern. Dies werden wir im dritten Abschnitt an einigen Beispielen darlegen. Im vierten Abschnitt zeigen wir, wie mit Hilfe des Satzes von Fubini für Kapazitäten Dominanzrelationen zwischen Produktkapazitäten durch Dominanzrelationen der Randkapazitäten charakterisiert werden können. Schließlich zeigen wir noch, wie unsere Resultate über stochastische Dominanz in AU und CEU angewendet werden können. Wir werden insbesondere zeigen, wie diese Ergebnisse in AU angewendet werden können, wenn neben der Nutzenfunktion auch die Verzerrungsfunktion nicht vollständig spezifiziert ist.

Dieses Kapitel folgt im wesentlichen – bis auf Abschnitt 4.4 – Dyckerhoff und Mosler (1993).

4.1 Grundlagen

Stochastische Dominanz von Kapazitäten wird wie folgt definiert.

Definition 4.1 (Stochastische Dominanz von Kapazitäten) *Seien* μ_1, μ_2 *zwei Kapazitäten auf* $(\Omega, \mathcal{A})$*, und sei* $\mathcal{F} \subset \mathcal{Q}(\mathcal{A})$*. Die Kapazität* μ_1 dominiert μ_2 stochastisch *bezüglich der Familie* $\mathcal{F}$ *(*$\mu_1 \succeq_{\mathcal{F}} \mu_2$*), wenn*

$$\int f d\mu_1 \;\geq\; \int f d\mu_2 \quad \textit{für alle } f \in \mathcal{F}\textit{, für die beide Integrale existieren.}$$

Ist Ω ein Vektorraum und enthält $\mathcal{A}$ mit jeder Menge A auch die Menge $-A = \{-\omega \,|\, \omega \in \mathcal{A}\}$, so läßt sich zu jeder stochastischen Dominanzrelation eine duale Relation definieren. Es sei

$$\mathcal{F}^D = \Big\{g : \Omega \to \mathbb{R} \,|\, g(\omega) = -f(-\omega) \text{ für ein } f \in \mathcal{F} \text{ und alle } \omega \in \Omega\Big\}.$$

$\mathcal{F}^D$ heißt dann die zu $\mathcal{F}$ *duale Funktionenklasse*. Wir geben hierfür einige Beispiele:

- Ist $\mathcal{F}$ die Klasse aller monoton wachsenden Funktionen, so ist $\mathcal{F}^D = \mathcal{F}$.
- Ist $\mathcal{F}$ die Klasse aller konkaven Funktionen, so ist $\mathcal{F}^D$ die Klasse aller konvexen Funktionen.
- Ist $\mathcal{F}$ die Menge aller monoton wachsenden und konkaven Funktionen, so ist $\mathcal{F}^D$ die Klasse aller monoton wachsenden und konvexen Funktionnen.

Wir definieren dann die zu $\succeq_{\mathcal{F}}$ *duale Dominanzrelation* $\succeq_{\mathcal{F}}^D$ durch $\succeq_{\mathcal{F}}^D = \succeq_{\mathcal{F}^D}$. Mit der Bezeichnung $\overline{\mu}(A) = \mu(-A)$ und wegen (2.1) und (2.3) erhält man

$$\mu_1 \succeq_{\mathcal{F}}^D \mu_2 \quad \text{genau dann, wenn} \quad \overline{\mu}_1^D \succeq_{\mathcal{F}} \overline{\mu}_2^D \ .$$

Eine wichtige stochastische Dominanzrelation ist die *stochastische Dominanz erster Ordnung* auf der reellen Achse: Sei $\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B}$ (die Borelsche σ-Algebra) und $\mathcal{F}_1$ die Menge aller monoton wachsenden Funktionen auf $\mathbb{R}$. Diese Dominanzrelation wurde für Wahrscheinlichkeitsmaße bereits von Lehmann (1955) charakterisiert. Eine Charakterisierung dieser Dominanzrelation für Kapazitäten gibt der folgende Satz.

Satz 4.1 *Seien μ_1, μ_2 Kapazitäten auf $(\mathbb{R},\mathcal{B})$ und sei $\mathcal{F}_1 = \{f : \mathbb{R} \to \mathbb{R} | f$ monoton wachsend $\}$. Dann gilt $\mu_1 \succeq_{\mathcal{F}_1} \mu_2$ genau dann, wenn*

$$\begin{array}{rrcll} & \mu_1\Big([t,\infty)\Big) & \geq & \mu_2\Big([t,\infty)\Big) & \textit{für alle } t \in \mathbb{R} \\ \textit{und} & & & & \\ & \mu_1\Big((t,\infty)\Big) & \geq & \mu_2\Big((t,\infty)\Big) & \textit{für alle } t \in \mathbb{R}. \end{array} \tag{4.1}$$

Beweis: Wir zeigen zunächst, daß (4.1) notwendig ist. Sei $\mu_1 \succeq_{\mathcal{F}_1} \mu_2$, d.h., $\int f d\mu_1 \geq \int f d\mu_2$ für alle $f \in \mathcal{F}_1$. Da $1_{[t,\infty)}$ für jedes $t \in \mathbb{R}$ in $\mathcal{F}_1$ liegt und wegen (2.4) erhalten wir

$$\mu_1\left([t,\infty)\right) = \int 1_{[t,\infty)} d\mu_1 \geq \int 1_{[t,\infty)} d\mu_2 = \mu_2\left([t,\infty)\right) .$$

Für Intervalle der Form $]t,\infty]$, $t \in \mathbb{R}$, schließt man analog.

Um zu zeigen, daß (4.1) auch hinreichend ist, nehmen wir an, daß f monoton wachsend ist. Dann sind die Level-Mengen $\{f \geq t\}$, $t \in \mathbb{R}$, obere Mengen. Da die einzigen oberen Mengen in $\mathbb{R}$ die Intervalle der Form $[t,\infty)$ und (t,∞) sind, gilt $\mu_1(f \geq t) \geq \mu_2(f \geq t)$ für alle $t \in \mathbb{R}$. Hieraus folgt sofort

$$\begin{aligned} \int f d\mu_1 &= \int_0^\infty \mu_1(f \geq t)dt + \int_{-\infty}^0 [\mu_1(f \geq t) - 1]dt \\ &\geq \int_0^\infty \mu_2(f \geq t)dt + \int_{-\infty}^0 [\mu_2(f \geq t) - 1]dt \\ &= \int f d\mu_2. \end{aligned}$$

■

4.2 Stochastische Dominanz bezüglich einer Mengenfamilie

Definition 4.2 *Sei Π eine Menge von Kapazitäten auf $(\Omega,\mathcal{A})$. Eine stochastische Dominanzrelation $\succeq_{\mathcal{F}}$ auf Π heißt* stochastische Dominanzrelation bezüglich einer Mengenfamilie *oder kurz* Mengen-Dominanzrelation*, wenn es eine Teilmenge $\mathcal{T} \subset \mathcal{A}$ gibt, so daß*

$$\mu_1 \succeq_{\mathcal{F}} \mu_2 \quad \textit{genau dann, wenn} \quad \mu_1(T) \geq \mu_2(T) \textit{ für alle } T \in \mathcal{T}.$$

Bei σ-additiven Wahrscheinlichkeitsmaßen konnte für viele stochastische Dominanzrelationen gezeigt werden, daß sie Mengen-Dominanzrelationen sind (Mosler und Scarsini 1991b, Bergmann 1991). Die meisten dieser Ergebnisse können auf den Fall der stochastischen Dominanz von Choquet-Kapazitäten verallgemeinert werden.

Verschiedene Klassen von Funktionen können zur selben Dominanzrelation führen. Daher ist es auch interessant, zu einer gegebenen Dominanzrelation jeweils minimale und maximale Funktionenklassen zu finden, die diese Relation repräsentieren.

Wir greifen dazu auf den Begriff der Quasi-$\mathcal{T}$-Meßbarkeit aus Definition 2.6 zurück. Der folgende Satz zeigt den Zusammenhang zwischen Mengen-Dominanz und Quasi-Meßbarkeit auf.

Satz 4.2 *Seien μ_1, μ_2 Kapazitäten auf $(\Omega, \mathcal{A})$, sei $\mathcal{T} \subset \mathcal{A}$ und sei $\mathcal{I}(\mathcal{T}) \subset \mathcal{F} \subset \mathcal{Q}(\mathcal{T})$. Dann gilt*

$$\int f d\mu_1 \geq \int f d\mu_2 \quad \textit{für alle } f \in \mathcal{F} \tag{4.2}$$

genau dann, wenn

$$\mu_1(T) \geq \mu_2(T) \quad \textit{für alle } T \in \mathcal{T}. \tag{4.3}$$

Beweis: Für $T \in \mathcal{T}$ and $f = 1_T$ ist $f \in \mathcal{F}$ und $\int f d\mu_i = \mu_i(T)$, $i = 1, 2$. Somit folgt (4.3) aus (4.2).

Für die umgekehrte Richtung sei $f \in \mathcal{F}$ und somit $f \in \mathcal{Q}(\mathcal{T})$. Für $t \in \mathbb{R}$ and $\epsilon > 0$ existiert ein $T_{t,\epsilon} \in \mathcal{T}$ mit

$$\Big\{f \geq t\Big\} \subset T_{t,\epsilon} \subset \Big\{f \geq t - \epsilon\Big\} .$$

Aus (4.3) folgt, daß

$$\mu_2(f \geq t) \leq \mu_2(T_{t,\epsilon}) \leq \mu_1(T_{t,\epsilon}) \leq \mu_1(f \geq t - \epsilon)$$

and damit

$$\int_0^\infty \mu_2(f \geq t)dt \ \leq \int_0^\infty \mu_1(f \geq t - \epsilon)dt \ = \int_{-\epsilon}^\infty \mu_1(f \geq t)dt\,.$$

Genauso erhält man

$$\int_{-\infty}^0 [\mu_2(f \geq t) - 1]dt \leq \int_{-\infty}^{-\epsilon} [\mu_1(f \geq t) - 1]dt\,.$$

Addition der beiden Gleichungen und Grenzübergang für $\epsilon \to 0$ ergibt

$$\int f d\mu_2 \ \leq \ \int f d\mu_1\,,$$

wie behauptet. ∎

Korollar 4.1 *Seien μ_1, μ_2 Kapazitäten auf $(\Omega, \mathcal{A})$, sei $\mathcal{T} \subset \mathcal{A}$ und sei $\mathcal{I}(\mathcal{T}) \subset \mathcal{F} \subset \mathcal{Q}(\mathcal{T})$. Dann ist*

$$\mu_1 \succeq_{\mathcal{F}} \mu_2 \text{ und } \mu_2 \succeq_{\mathcal{F}} \mu_1$$

genau dann, wenn

$$\mu_1(T) = \mu_2(T) \quad \text{für alle } T \in \mathcal{T}. \tag{4.4}$$

Das Korollar ist eine unmittelbare Folgerung aus obigem Satz. Es zeigt, daß stochastische Dominanzrelationen i.a. nicht antisymmetrisch sind. Eine Dominanzrelation ist lediglich dann antisymmetrisch, wenn die Klasse $\mathcal{T}$ reichhaltig genug ist, d.h. wenn Gleichheit der Kapazitäten auf $\mathcal{T}$ bereits Gleichheit der Kapazitäten auf $\mathcal{A}$ impliziert.

$\mathcal{F} = \mathcal{I}(\mathcal{T})$ ist nicht unbedingt die kleinste Funktionenklasse, für die die Äquivalenz in Satz 4.2 gilt. Wenn jedoch die Voraussetzungen des Satzes erfüllt sind, gilt in jedem Fall $\succeq_{\mathcal{F}} = \succeq_{\mathcal{I}(\mathcal{T})}$. Auf der anderen Seite ist $\mathcal{F} = \mathcal{Q}(\mathcal{T})$ die größte Funktionenklasse, für die die Äquivalenz in Satz 4.2 gilt, wenn μ_1 und μ_2 beliebige Kapazitäten sind. Wie Greco (1981) gezeigt hat, gibt es zu jedem $f \notin \mathcal{Q}(\mathcal{T})$ Kapazitäten μ_1 und μ_2, so daß $\mu_1(T) \geq \mu_2(T)$ für alle $T \in \mathcal{T}$ ist, aber $\int f d\mu_1 < \int f d\mu_2$. Mit anderen Worten:

Korollar 4.2 *Sei $\mathcal{T} \subset \mathcal{A}$, und sei $\mathcal{I}(\mathcal{T}) \subset \mathcal{F} \subset \mathcal{Q}(\mathcal{T})$. Dann ist $\mathcal{Q}(\mathcal{T})$ die Klasse der $\preceq_{\mathcal{F}}$-monotonen Funktionen, d.h. die Menge aller quasi-$\mathcal{A}$-meßbaren Funktionen f, für die gilt:*

$$\mu_1 \preceq_{\mathcal{F}} \mu_2 \quad \Rightarrow \quad \int f d\mu_1 \leq \int f d\mu_1 \ .$$

Im folgenden geben wir einige Beispiele für Mengen-Dominanzrelationen. In jedem Beispiel wird zu einer gegebenen Klasse $\mathcal{T}$ von Mengen die Familie $\mathcal{Q}(\mathcal{T})$ oder zumindest eine Famile $\mathcal{F} \subset \mathcal{Q}(\mathcal{T})$ bestimmt, für die Satz 4.2 zutrifft.

Beispiel 4.1 *Sei $(\Omega, \mathcal{A})$ beliebig und $h : \Omega \to \overline{\mathbb{R}}$ eine gegebene von oben $\mathcal{A}$-meßbare Funktion. Wir betrachten die Familie $\mathcal{T}_h$ der oberen Level-Mengen von h,*

$$\mathcal{T}_h = \Big\{ \{h \geq \alpha\} \Big| \, \alpha \in \overline{\mathbb{R}} \Big\}.$$

Weiter sei

$$\mathcal{F}_h = \Big\{ f : \Omega \to \mathbb{R} \,|\, f = \rho \circ h, \ \rho : h(\Omega) \to \mathbb{R} \text{ monoton wachsend und von rechts stetig} \Big\}.$$

Dann ist

$$\mathcal{F}_h \ = \ \mathcal{Q}(\mathcal{T}_h)$$

und Satz 4.2 gilt mit $\mathcal{T}_h$ und $\mathcal{F}_h$.

Beweis: Sei $\alpha \in \mathbb{R}$, $f \in \mathcal{F}_h$ und $f = \rho \circ h$ mit einer monoton wachsenden und von rechts stetigen Funktion $\rho : h(\Omega) \to \mathbb{R}$. Dann ist $\{\rho \geq \alpha\} = [\beta, \infty)$ für ein $\beta \in \mathbb{R}$ und

$$\Big\{f \geq \alpha\Big\} = \Big\{h \geq \beta\Big\} \in \mathcal{T}_h\,.$$

Somit ist $f \in \mathcal{Q}(\mathcal{T}_h)$, wie behauptet.

Sei nun umgekehrt f in $\mathcal{Q}(\mathcal{T}_h)$. Wir zeigen zuerst, daß aus $h(x) \leq h(y)$ folgt, daß $f(x) \leq f(y)$ ist. Wir nehmen dazu an, daß es ein Paar x, y gibt, so daß $h(x) \leq h(y)$ und $f(x) > f(y)$ ist. Da f quasi-$\mathcal{T}_h$-meßbar ist, existiert für $\alpha = f(x)$ und $\epsilon = ((f(x) - f(y))/2$ ein $\beta \in \overline{\mathbb{R}}$, so daß

$$\Big\{f \geq f(x)\Big\} \subset \Big\{h \geq \beta\Big\} \subset \Big\{f \geq (f(x) + f(y))/2\Big\}\,.$$

Also ist x in $\{h \geq \beta\}$. Da $h(x) \leq h(y)$, ist aber auch y in $\{h \geq \beta\}$. Andererseits ist y nicht in $\{f \geq (f(x) + f(y))/2\}$; Widerspruch.

Wir erhalten also, daß aus $h(x) = h(y)$ auch $f(x) = f(y)$ folgt. Somit gibt es eine reelle Funktion $\rho : h(\Omega) \to \mathbb{R}$ mit $f = \rho \circ h$, und diese Funktion ist monoton wachsend. Wir zeigen nun wieder durch Widerspruchsbeweis, daß ρ auch von rechts stetig ist. Dazu nehmen wir an, daß ρ nicht von rechts stetig ist. Dann gibt es eine monoton fallende Folge $(\beta_n)_{n\in\mathbb{N}}$ in $h(\Omega)$, die gegen $\beta \in h(\Omega)$ konvergiert, so daß $\alpha = \lim_{n\to\infty} \rho(\beta_n) > \rho(\beta)$. Dann gilt für alle ϵ mit $0 < \epsilon < \alpha - \rho(\beta)$

$$\Big\{h \geq \beta_n\Big\} \underset{\neq}{\subset} \Big\{f \geq \alpha\Big\} = \Big\{h > \beta\Big\} = \Big\{f \geq \alpha - \epsilon\Big\} \underset{\neq}{\subset} \Big\{h \geq \beta\Big\}\,.$$

Also gibt es kein γ, für das

$$\Big\{f \geq \alpha\Big\} \subset \Big\{h \geq \gamma\Big\} \subset \Big\{f \geq \alpha - \epsilon\Big\}$$

ist, was im Widerspruch zu $f \in \mathcal{Q}(\mathcal{T}_h)$ steht. ∎

Die Rechtsstetigkeit von ρ ist notwendig, damit $\mathcal{F}_h \subset \mathcal{Q}(\mathcal{T}_h)$ gilt. Spezialfälle ergeben sich z.B., wenn Ω ein normierter Vektorraum mit Norm h ist. Dann ist $\mathcal{T}_h$ die Menge der abgeschlossenen Kugeln um den Ursprung und $\mathcal{F}_h$ ist eine Teilmenge der norm-wachsenden Funktionen. Eine ähnliche Situation entsteht, wenn $\Omega = \Omega' \times \Omega'$ und h eine Metrik auf Ω' ist.

In obigem Beispiel induziert h eine schwache Ordnung auf Ω. Umgekehrt existiert zu jedem schwach geordnetem Raum $(\Omega, \preceq)$ – sofern Ω^* eine abzählbare ordnungsdichte Teilmenge enthält – eine reelle Funktion $h : \Omega \to \mathbb{R}$, die diese schwache Ordnung induziert.

Beispiel 4.2 *Wir betrachten einen mit einer Präordnung versehenen Raum* $(\Omega, \mathcal{A}, \leq)$ *und die Familie* $\mathcal{U} = \{\, U \in \mathcal{A} \,|\, U$ *ist obere Menge*$\}$ *der* oberen Mengen *in* $\mathcal{A}$. *Sei* $\mathcal{F}_1 = \{\, f \in \mathcal{Q}(\mathcal{A}) \,|\, f$ *monoton wachsend*$\}$. *Dann ist*

$$\mathcal{F}_1 = \mathcal{Q}(\mathcal{U})\,. \tag{4.5}$$

Dieses Beispiel stellt eine Verallgemeinerung von Satz 4.1 dar. Setzt man nämlich $\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B}$ und für $\leq$ die gewöhnliche Ordnung auf $\mathbb{R}$, so erhält man genau die Aussage von Satz 4.1. Für das entsprechende Resultat bei Wahrscheinlichkeitsmaßen vergleiche man Lehmann (1955) und Kamae, Krengel und O'Brien (1977).

Beispiel 4.3 *Weitergehende Ergebnisse erhält man unter topologischen Voraussetzungen. Wir betrachten dazu einen topologischen Raum $(\Omega, \mathcal{O})$ mit einer stetigen*[1] *Präordnung $\leq$ und der Borelschen σ-Algebra $\mathcal{B}(\Omega)$. Wir bezeichnen mit*

$$\mathcal{U}_{op} = \Big\{ S \subset \Omega \,|\, S \text{ offen und obere Menge} \Big\},$$

und

$$\mathcal{U}_{cl} = \Big\{ S \subset \Omega \,|\, S \text{ abgeschlossen und obere Menge} \Big\}.$$

Dann ist

$$\mathcal{Q}(\mathcal{U}_{cl}) = \mathcal{Q}(\mathcal{U}_{op}) = \Big\{ f : \Omega \to \mathbb{R} \,|\, f \text{ stetig und monoton wachsend} \Big\},$$

und Satz 4.2 gilt sowohl mit $\mathcal{T} = \mathcal{U}_{cl}$ als auch mit $\mathcal{T} = \mathcal{U}_{op}$.

Beispiel 4.4 *Sei Ω ein Vektorraum, versehen mit einer Präordnung $\leq$, und sei $\mathcal{U}_{cv} = \{ S \,|\, S$ obere Menge und konvex$\}$. Dann ist*

$$\mathcal{Q}(\mathcal{U}_{cv}) = \Big\{ f : \Omega \to \mathbb{R} \,|\, f \text{ monoton wachsend und quasikonkav} \Big\}.$$

Für σ-additive Wahrscheinlichkeitsmaße, vgl. Levhari et. al. (1975).

Beispiel 4.5 *Sei $\Omega = \mathbb{R}^n$, $\leq$ die komponentenweise partielle Ordnung und $<\cdot,\cdot>$ ein Skalarprodukt. Wir betrachten die abgeschlossenen Halbräume $H(a,\beta) = \{ x \in \Omega \mid <a,x> \ \geq \beta\}$, wobei $a \in \mathbb{R}^n$, $\beta \in \mathbb{R}$. Sei $\mathcal{U}_{hs}$ die Menge der oberen Halbräume. Dann ist*

$$\mathcal{Q}(\mathcal{U}_{hs}) = \Big\{ f : \Omega \to \mathbb{R} \,|\, f = \rho(<a,\cdot>),\ \rho : \mathbb{R} \to \mathbb{R} \text{ monoton wachsend} \\ \text{und von rechts stetig, } H(a,0) \text{ obere Menge} \Big\}$$

In Fall σ-additiver Wahrscheinlichkeitsmaße wurde diese Ordnung von Muliere und Scarsini (1989) untersucht.

Beispiel 4.6 *Sei $\Omega = \mathbb{R}^d$ und $\mathcal{G}$ eine Untergruppe der orthogonalen Gruppe $\mathcal{O}^d$. Sei weiter $\mathcal{A}$ invariant unter Transformationen $g \in \mathcal{G}$. Wir definieren auf $\mathbb{R}^d$ eine Relation $\leq_{\mathcal{G}}$ wie folgt: $x \leq_{\mathcal{G}} y$, falls x in der konvexen Hülle der Bahn von*

[1] Eine Präordnung $\preceq$ auf einem topologischem Raum $(\Omega, \mathcal{O})$ heißt *stetig*, wenn für alle $x \in \Omega$ die Mengen $\{y|x \preceq y\}$ und $\{y|y \preceq x\}$ abgeschlossen sind.

y liegt, d.h., $x \in \operatorname{conv}\{z|z = gy,\ g \in \mathcal{G}\}$. Eine Funktion $f : \mathbb{R}^d \to \mathbb{R}$ heißt $\mathcal{G}$-wachsend, wenn sie bezüglich $\leq_{\mathcal{G}}$ isoton ist. Mit $\mathcal{T}_{\mathcal{G}} = \{S \mid 1_S\ \mathcal{G}\text{-wachsend}\}$ erhält man

$$\mathcal{Q}(\mathcal{T}_{\mathcal{G}}) = \Big\{ f : \Omega \to \mathbb{R} \mid f\ \mathcal{G}\text{-wachsend} \Big\} \tag{4.6}$$

In diesem Beispiel ist $\mathcal{T}_{\mathcal{G}}$ die Familie der oberen Mengen bezüglich $\leq_{\mathcal{G}}$. Somit folgt (4.6) aus (4.5). Ein interessanter Spezialfall entsteht, wenn $\mathcal{G} = \mathcal{G}_{Sch} = \{\pi \in \mathcal{O}^d \mid \pi \text{ ist eine Permutationsmatrix }\}$. In diesem Fall ist $\mathcal{Q}(\mathcal{T}_{\mathcal{G}})$ die Menge der Schur-konvexen Funktionen; siehe Marshall and Olkin (1979). Für andere Spezialfälle und Anwendungen mit Wahrscheinlichkeitsmaßen verweisen wir auf die Bücher von Eaton (1987) und Dharmadhikari und Joag-Dev (1988).

Beispiel 4.7 *Sei wieder $\Omega = \mathbb{R}^d$, $\mathcal{G}$ eine Untergruppe von $\mathcal{O}^d$, $\mathcal{A}$ $\mathcal{G}$-invariant und $\mathcal{T}_{\mathcal{G}} = \{T \mid T\ \mathcal{G}\text{-invariant und konvex}\}$. Dann ist*

$$\mathcal{Q}(\mathcal{T}_{\mathcal{G}}) = \Big\{ f : \Omega \to \mathbb{R} \mid f\ \mathcal{G}\text{-invariant und unimodal} \Big\}$$

Dabei heißt eine Funktion f unimodal, wenn die Mengen $\{f \geq \alpha\}$, $\alpha \in \mathbb{R}$ sämtlich konvex sind. Für Anwendungen dieser Dominanzrelation in der Statistik (dort mit Wahrscheinlichkeitsmaßen) vergleiche man Mosler (1987) und Dharmadhikari und Joag-Dev (1988).

Beweis: Ist f $\mathcal{G}$-invariant und unimodal, so ist $\{f \geq \alpha\}$ für jedes $\alpha \in \mathbb{R}$ $\mathcal{G}$-invariant und konvex, und somit ist $f \in \mathcal{Q}(\mathcal{T}_{\mathcal{G}})$.
Sei nun umgekehrt $f \in \mathcal{Q}(\mathcal{T}_{\mathcal{G}})$. Dann existiert für jedes $\alpha \in \mathbb{R}$ und $\epsilon > 0$ eine $\mathcal{G}$-invariante und konvexe Menge $T_{\alpha,\epsilon}$ mit (2.15): $\{f \geq \alpha\} \subset T_{\alpha,\epsilon} \subset \{f \geq \alpha - \epsilon\}$. Somit ist $\{f \geq \alpha\} = \bigcap_{\epsilon > 0} T_{\alpha,\epsilon}$ und als Schnitt konvexer Mengen wieder konvex. Also ist f is unimodal. Wir nehmen nun an, daß f nicht $\mathcal{G}$-invariant ist. Dann gibt es $y \in \Omega$ und $g \in \mathcal{G}$, so daß $f(y) < f(gy)$. Wegen (2.15) mit $\alpha = f(gy), \epsilon = (f(gy) - f(y))/2$ ist aber $gy \in T_{\alpha,\epsilon}$. Da $T_{\alpha,\epsilon}$ aber $\mathcal{G}$-invariant ist, ist auch $y \in T_{\alpha,\epsilon}$ und damit auch in $\{f \geq \alpha - \epsilon\}$. Das führt aber wegen $f(y) > \alpha - 2\epsilon = f(y)$ zum Widerspruch. ■

Beispiel 4.8 *Sei $\Omega = \mathbb{R}^d$, $\leq$ die gewöhnliche partielle Ordnung auf $\mathbb{R}^d$, $\mathcal{T}_{uo} = \{\{x|x \geq z\} \mid z \in \overline{\mathbb{R}}^d\}$ die Familie der oberen Orthanten und $\mathcal{A}$ ein Mengensystem, das $\mathcal{T}_{uo}$ enthält. Dann ist*

$$\mathcal{Q}(\mathcal{T}_{uo}) \supset \Big\{ f : \Omega \to \mathbb{R} \mid f(x) = \rho(\min_i\{\lambda_i x_i\}),\ \lambda_i \geq 0, \\ \rho : \mathbb{R} \to \mathbb{R} \text{ monoton wachsend und von rechts stetig} \Big\}$$

Die Inklusion ist strikt. Dies zeigt das Beispiel $f(x_1, x_2) = \min\{x_1, x_2^2\}$. Mengen-Dominanz von σ-additiven Wahrscheinlichkeitsmaßen bezüglich $\mathcal{T}_{uo}$ wurde von Mosler (1982, 1984), Scarsini (1985, 1988) und Rüschendorf (1981) untersucht.

Beweis: Sei $f(x) = \rho(\min_i\{\lambda_i x_i\})$ für gewisse $\lambda_i \geq 0$ und eine monoton wachsende und von rechts stetige Funktion ρ. Wir definieren die Quasi-Inverse einer monoton wachsenden Funktion wie üblich durch $\rho^{-1}(y) = \inf\{x \mid \rho(x) \geq y\}$. Dann ist

$$\Big\{f \geq \alpha\Big\} = \Big\{x | \min_i\{\lambda_i x_i\} \geq \rho^{-1}(\alpha)\Big\} = \Big\{x | x \geq z\Big\},$$

wobei $z = (z_1, \ldots, z_d)$, $z_i = \rho^{-1}(\alpha)/\lambda_i$. Somit ist $f \in \mathcal{Q}(\mathcal{T}_{uo})$, wie behauptet. ∎

Bemerkung: Sei h eine von oben $\mathcal{A}$-meßbare Funktion. Gilt für alle $\alpha \in \mathbb{R}$

$$\mu_1(h \geq \alpha) \geq \mu_2(h \geq \alpha) \text{ und } \mu_1(h > \alpha) \geq \mu_2(h > \alpha)$$

dann ist auch

$$\mu_1(\rho \circ h \geq \alpha) \geq \mu_2(\rho \circ h \geq \alpha) \text{ und } \mu_1(\rho \circ h > \alpha) \geq \mu_2(\rho \circ h > \alpha)$$

für alle α und für jede monoton wachsende Funktion $\rho : \mathbb{R} \to \mathbb{R}$. Somit ist $\int(\rho \circ h) d\mu_1 \geq \int(\rho \circ h) d\mu_2$ für jede monoton wachsende Funktion $\rho : \mathbb{R} \to \mathbb{R}$. Das bedeutet aber, daß Mengen-Dominanz bezüglich der Level-Mengen einer gegebenen Funktion eine große Klasse von Funktionen liefert, die monoton bezüglich dieser Dominanzrelation sind, nämlich alle monoton wachsenden Transformationen dieser Funktion, bzw. alle Funktionen, die isoton bezüglich der von h auf Ω induzierten schwachen Ordnung sind. Andere Dominanzrelationen auf schwach geordneten Räumne werden im folgenden Abschnitt untersucht.

4.3 Stochastische Dominanz auf schwach geordneten Räumen

Wir untersuchen in diesem Abschnitt Dominanzrelationen auf schwach geordneten Räumen. Sind die Choquet-Kapazitäten stetig auf oberen Mengen und ist $(\Omega, \leq)$ vollständig bezüglich beschränkter, monotoner Folgen, so existiert nach Satz 2.4 ein W-Maß P auf der von den oberen Mengen erzeugten σ-Algebra $\mathcal{A}$, so daß

$$\int f d\mu = \int f dP \quad \text{für alle monoton wachsenden Funktionen } f \tag{4.7}$$

gilt.

Gilt (4.7), so ist es einfach, bekannte Charakterisierungen von stochastischen Dominanzrelationen mit W-Maßen auf Kapazitäten zu übertragen. Ist nämlich $\mathcal{F} \subset \mathcal{F}_1 = \{f : \Omega \to \mathbb{R} \mid f \text{ monoton wachsend}\}$, so ist

$$\int f d\mu_1 \geq \int f d\mu_2 \quad \text{für alle } f \in \mathcal{F}$$

genau dann, wenn

$$\int f dP_1 \geq \int f dP_2 \quad \text{für alle } f \in \mathcal{F} \ ,$$

wobei P_i das zu μ_i bezüglich $\leq$ assoziierte W-Maß bezeichnet, $i = 1, 2$.

Wir beschreiben nun, wie mittels Satz 2.4 Charakterisierungen von stochastischen Dominanzrelationen, die für Wahrscheinlichkeitsmaße bekannt sind, auf Kapazitäten übertragen werden können.

Satz 4.3 *Sei Ω ein Intervall in $\mathbb{R}$, und sei $\mathcal{B} = \mathcal{B}(\Omega)$ die Borelsche σ-Algebra auf Ω. Weiter seien $\mathcal{F} \subset \mathcal{F}_1$ eine Menge von monoton wachsenden Funktionen und μ_1, μ_2 Kapazitäten auf $(\Omega, \mathcal{B})$, die stetig auf oberen Mengen sind. Für Wahrscheinlichkeitsmaße bestehe die folgende Charakterisierung der Dominanzrelation $\leq_{\mathcal{F}}$:*

$$P_1 \preceq_{\mathcal{F}} P_2 \text{ genau dann, wenn } U(P_1, P_2) \, ,$$

wobei $U(P_1, P_2)$ eine Aussage ist, deren Gültigkeit lediglich von den Werten der Wahrscheinlichkeitsmaße auf oberen Mengem abhängt. Dann gilt für die Kapazitäten μ_1, μ_2

$$\mu_1 \preceq_{\mathcal{F}} \mu_2 \text{ genau dann, wenn } U(\mu_1, \mu_2) \ .$$

Beweis: Jedes Intervall in $\mathbb{R}$ ist vollständig bezüglich beschränkter, monotoner Folgen. Somit sind die Voraussetzungen von Satz 2.4 erfüllt. Es existieren also W-Maße P_1, P_2, die auf oberen Mengen mit μ_1 und μ_2 übereinstimmen. Nach Lemma 2.1 ist

$$\int f dP_i = \int f d\mu_i, \quad i = 1, 2, \quad \text{für alle monoton wachsenden } f \, .$$

Damit gelten die folgenden Äquivalenzen:

$$\mu_1 \preceq_{\mathcal{F}} \mu_2 \iff P_1 \preceq_{\mathcal{F}} P_2 \iff U(P_1, P_2) \iff U(\mu_1, \mu_2) \, .$$

∎

Wir illustrieren nun die Ergebnisse dieses Abschnitts mit einigen Beispielen.

Beispiel 4.9 *Sei $\mathcal{F}_2 = \{u : \mathbb{R} \to \mathbb{R} \,|\, u \text{ monoton wachsend und konkav}\}$ und μ_1, μ_2 Kapazitäten auf $(\mathbb{R}, \mathcal{B})$. Sind μ_1, μ_2 stetig auf oberen Mengen, so gilt*

$$\mu_1 \preceq_{\mathcal{F}_2} \mu_2 \iff \int_{-\infty}^{x} [\overline{F}_1(t) - \overline{F}_2(t)] dt \leq 0 \quad \forall x \in \mathbb{R} \, ,$$

wobei $\overline{F}_i(t) = \mu_i([t, \infty))$, $i = 1, 2$, die dekumulative Verteilungsfunktion bezeichnet.

Ein analoges Resultat erhält man für Kapazitäten, die auf unteren Mengen stetig sind, wenn man das untere Choquet-Integral benutzt und $\overline{F}_i$ durch F_i und $\geq$ durch $\leq$ ersetzt.

Beweis: Für Wahrscheinlichkeitsmaße P_1, P_2 gilt

$$P_1 \preceq_{\mathcal{F}_2} P_2 \iff \int_{-\infty}^{x} P_1\Big([t,\infty)\Big) - P_2\Big([t,\infty)\Big)\,dt \;\leq 0 \quad \forall x \in \mathbb{R}\,.$$

Die Aussage auf der rechten Seite hängt lediglich von den Werten der Wahrscheinlichkeitsmaße auf oberen Mengen ab. Die Behauptung folgt daher direkt aus Satz 4.3. ∎

Beispiel 4.10 *Für $k \geq 2$ sei*

$$\mathcal{F}'_k = \Big\{u : \mathbb{R} \to \mathbb{R} \,|\, u(x) = \int_x^{\infty}\int_{t_{k-2}}^{\infty}\cdots\int_{t_1}^{\infty} v(t_0)\,dt_0 \ldots dt_{k-2},$$
$$v \textit{ monoton wachsend und von rechts stetig, } v \leq 0,\ \lim_{x\to\infty} v(x) = 0\Big\}\,.$$

Seien μ_1, μ_2 Kapazitäten auf $(\mathbb{R}, \mathcal{B})$, die stetig auf oberen Mengen sind und bezeichne $\overline{F}_i$ die dekumulative Verteilungsfunktion von μ_i, $i = 1, 2$. Dann gilt

$$\mu_1 \succeq_{\mathcal{F}'_k} \mu_2 \iff \int_{-\infty}^{x}\int_{-\infty}^{t_{k-2}}\cdots\int_{-\infty}^{t_1}[\overline{F}_1(t_0) - \overline{F}_2(t_0)]\,dt_0 \ldots dt_{k-2} \geq 0 \quad \forall x \in \mathbb{R}\,.$$

Die Aussage folgt wieder direkt aus einem bekannten Resultat für W-Maße (Rolski 1976) und Satz 4.3.

Beispiel 4.11 *Für $k \geq 2$ und $b \geq 0$ sei*

$$\mathcal{F}^c_k = \Big\{\, u : [0,b] \to \mathbb{R} \,|\, u \textit{ ist } k\textit{-mal differenzeirbar}$$
$$\textit{und } (-1)^i u^{(i)} \leq 0,\ i = 1, \ldots, k\Big\}\,.$$

Seien μ_1, μ_2 auf oberen Mengen stetige Kapazitäten und sei $m_j(\mu) = \int x^j \mu(dx)$ das j-te obere Moment von μ. Dann sind die folgenden Bedingungen hinreichend für $\mu_1 \succeq_{\mathcal{F}^c_k} \mu_2$:

$$\int_0^{x}\int_0^{t_{k-2}}\cdots\int_0^{t_1}[\overline{F}_1(t_0) - \overline{F}_2(t_0)]\,dt_0 \ldots dt_{k-2} \geq 0 \quad \forall x \in [0,b]\,,$$

und

$$(-1)^j m_j(\mu_1) \;\leq\; (-1)^j m_j(\mu_2), \quad 1 \leq j \leq k-2\,.$$

Die Aussage folgt wieder aus Satz 4.3 und einem Ergebnis in Mosler (1982, S. 70).

Wir betrachten nun noch Nutzenfunktionen, die auf einem partiell geordnetem Raum definiert sind, wie z.B. $\mathbb{R}^n$ mit der gewöhnlichen partiellen Ordnung. Während alle monoton wachsenden Funktionen auf einem linear geordnetem Raum komonoton sind, ist dies auf partiell geordneten Räumen nicht der Fall. Um Satz 2.4 anwenden zu können, muß man sich daher auf komonotone Mengen von Funktionen beschränken. Wir betrachten also nur solche Mengen von Funktionen, bei denen alle Funktionen dieselbe schwache Ordnung induzieren.

Beispiel 4.12 *Sei* $\Omega = \mathbb{R}_+^n$, $\mathcal{B}_+^n$ *die Borelsche* σ*-Algebra auf* $\mathbb{R}_+^n$ *und* $u : \mathbb{R}_+^n \to \mathbb{R}$ *eine stetige, monoton wachsende, quasi-konkave Nutzenfunktion. Weiter seien*

$$\mathcal{F}_1^u = \Big\{ w : \mathbb{R}_+^n \to \mathbb{R}_+ \,|\, w = \rho \circ u,\ \rho : \mathbb{R}_+ \to \mathbb{R}_+ \text{ monoton wachsend und von rechts stetig} \Big\}$$

und

$$\mathcal{F}_2^u = \Big\{ w \in \mathcal{F}_1^u \,|\, w \text{ konkav} \Big\}.$$

Diese Klassen von Nutzenfunktionen wurden von Levy und Levy (1984) sowie, im Rahmen von CEU, von Scarsini (1992) untersucht. Debreu (1976) hat gezeigt, daß es eine Funktion $u^* \in \mathcal{F}_1^u$ *gibt, so daß*

$$\mathcal{F}_2^u = \Big\{ w : \mathbb{R}_+^n \to \mathbb{R}_+ \,|\, w = \rho \circ u^*,\ \rho : \mathbb{R}_+ \to \mathbb{R}_+ \text{ monoton wachsend und von rechts stetig} \Big\}.$$

Die Funktion u^* *heißt auch* minimale konkave Funktion in $\mathcal{F}_1^u$ *(engl.* least concave utility function*). Mit den Bezeichnungen aus Beispiel 4.1 ist somit* $\mathcal{F}_1^u = \mathcal{F}_u$ *und* $\mathcal{F}_2^u = \mathcal{F}_{u^*}$. *Durch* u^* *wird mittels* $x \leq_{u^*} y \iff u^*(x) \leq u^*(y)$ *eine schwache Ordnung auf* Ω *induziert. Seien* μ_1, μ_2 *wieder Kapazitäten auf* $(\mathbb{R}_+^n, \mathcal{B}_+^n)$*, die auf oberen Mengen (bezüglich* $\leq_{u^*}$*) stetig sind. Dann gilt*

$$\mu_1 \preceq_{\mathcal{F}_2^u} \mu_2 \iff \int_0^x \mu_1(u^* \geq t)dt \leq \int_0^x \mu_2(u^* \geq t)dt \quad \forall x \in \mathbb{R}_+ .$$

Beweis: Man erhält dieses Ergebnis aus Satz 2.4 und einem Satz in Levy und Levy (1984). ∎

4.4 Stochastische Dominanz von Produktkapazitäten

In diesem Abschnitt betrachten wir den Fall, daß die Kapazitäten Produktkapazitäten sind. Wir beschränken uns der Einfacheit halber auf den Fall, daß die Produktkapazitäten auf $(\mathbb{R}^2, \mathcal{B}^2)$ definiert sind. Wir haben in Abschnitt 2.4 gesehen, daß in dieser Situation das Choquet-Integral über isotone Funktionen bezüglich der Produktkapazität als iteriertes Integral dargestellt werden kann. Dies ermöglicht es uns in bestimmten Fällen Dominanzrelationen zwischen den Produktkapazitäten durch Dominanzrelationen der Randkapazitäten zu charakterisieren. Für entsprechende Resultate bei Wahrscheinlichkeitsmaßen vergleiche Mosler 1982, S. 42f).

Wir betrachten die folgende Situation:

Voraussetzung (für diesen Abschnitt): Für $i = 1, 2$ seien μ_i, ν_i stetige Kapazitäten auf $(\mathbb{R}, \mathcal{B})$ und $\mu \in \mu_1 \otimes \mu_2$, $\nu \in \nu_1 \otimes \nu_2$ Produktkapazitäten auf $(\mathbb{R}^2, \mathcal{B}^2)$. Weiter sei $\mathcal{F}_i$ eine Menge von isotonen, reellen Funktionen auf $\mathbb{R}$, $i = 1, 2$, und

$$\mathcal{F} = \{f : \mathbb{R}^2 \to \mathbb{R} \,|\, f \ \mathcal{B}^2\text{-meßbar, } \mu\text{- und } \nu\text{-integrierbar,} \\ f_{x_1} \in \mathcal{F}_2, \forall\, x_1 \in \mathbb{R},\ f_{x_2} \in \mathcal{F}_1, \forall\, x_2 \in \mathbb{R}\,\}.$$

Dann sind alle Funktionen in $\mathcal{F}$ isoton bezüglich der gewöhnlichen partiellen Ordnung auf $\mathbb{R}^2$. Nach Beispiel 2.3 sind für die Klasse $\mathcal{F}$ die Voraussetzungen des Satzes von Fubini für Kapazitäten erfüllt.

Wir sagen, daß eine Menge $\mathcal{M}$ von reellen Funktionen *abgeschlossen bezüglich Integration* ist, wenn für jede Menge $\{f_\omega \,|\, \omega \in \Omega\} \subset \mathcal{M}$ und alle Kapazitäten μ die Abbildung

$$x \mapsto \int_\Omega f_\omega(x) d\mu(\omega)$$

wieder in $\mathcal{M}$ liegt.

Satz 4.4 *Unter der Voraussetzung dieses Abschnitts gilt: Ist $\mathcal{F}_1$ oder $\mathcal{F}_2$ abgeschlossen bezüglich Integration, so gilt*

$$\mu_i \preceq_{\mathcal{F}_i} \nu_i,\ i = 1, 2, \quad \Rightarrow \quad \mu_1 \otimes \mu_2 \preceq_{\mathcal{F}} \nu_1 \otimes \nu_2$$

Beweis: Sei $f \in \mathcal{F}$. Wir nehmen o.B.d.A. an, daß $\mathcal{F}_2$ abgeschlossen bezüglich Integration ist. Zunächst ist wegen $\mu_1 \preceq_{\mathcal{F}_1} \nu_1$

$$\int_{\mathbb{R}} f_{x_2} d\mu_1 \le \int_{\mathbb{R}} f_{x_2} d\nu_1 \quad \text{für alle } x_2 \in \mathbb{R}.$$

Aus der Monotonie des Choquet-Integrals folgt dann

$$\int_{\mathbb{R}} \left[\int_{\mathbb{R}} f_{x_2} d\mu_1\right] d\mu_2(x_2) \le \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f_{x_2} d\nu_1\right] d\mu_2(x_2).$$

Da $\mathcal{F}_2$ abgeschlossen bezüglich Integration ist, liegt auch die Funktion

$$x_2 \mapsto \int_{\mathbb{R}} f_{x_2}(x_1) d\nu_1(x_1) = \int_{\mathbb{R}} f_{x_1}(x_2) d\nu_1(x_1)$$

wieder in $\mathcal{F}_2$. Wegen $\mu_2 \preceq_{\mathcal{F}_2} \nu_2$ gilt dann auch

$$\int_{\mathbb{R}} \left[\int_{\mathbb{R}} f_{x_2} d\nu_1\right] d\mu_2(x_2) \le \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f_{x_2} d\nu_1\right] d\nu_2(x_2).$$

Die beiden letzten Ungleichungen zusammen ergeben dann

$$\int_{\mathbb{R}} \left[\int_{\mathbb{R}} f_{x_2} d\mu_1\right] d\mu_2(x_2) \le \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f_{x_2} d\nu_1\right] d\nu_2(x_2).$$

Nun sind aber nach dem Satz von Fubini für Kapazitäten die beiden Doppelintegrale gleich den Integralen bezüglich der Produktkapazitäten. Für alle $f \in \mathcal{F}$ gilt somit

$$\int_{\mathbb{R}^2} f\, d\mu \le \int_{\mathbb{R}^2} f\, d\nu,$$

wie behauptet. ∎

Wir geben noch zwei Beispiele für die Anwendung dieses Satzes.

Beispiel 4.13 *Seien* $\mathcal{F}_i = \{f : \mathbb{R} \to \mathbb{R} \,|\, f \text{ isoton } \}$, $i = 1, 2$. *Dann ist* $\mathcal{F} = \{f : \mathbb{R}^2 \to \mathbb{R} \,|\, f \text{ isoton}\}$. *Sind* μ_i, ν_i, $i = 1, 2$, *stetige Kapazitäten, so gilt*

$$\mu_i \preceq_{\mathcal{F}_i} \nu_i,\ i = 1, 2. \qquad \Longleftrightarrow \qquad \mu_1 \otimes \mu_2 \preceq_{\mathcal{F}} \nu_1 \otimes \nu_2.$$

Beweis: Ist $f(\cdot, \omega_2)$ isoton für jedes $\omega_2 \in \Omega_2$, so ist wegen der Monotonie des Integrals auch $\int_{\Omega_2} f(\cdot, \omega_2) d\mu_2(\omega_2)$ isoton. $\mathcal{F}_1$ ist daher abgeschlossen gegen Integration ($\mathcal{F}_2$ natürlich auch) und mit Satz 4.4 folgt die „$\Rightarrow$“-Richtung.

„$\Leftarrow$“ folgt, da für jedes $f_1 \in \mathcal{F}_1$ die Funktion $f : \Omega \to \mathbb{R}$, $f(\omega_1, \omega_2) = f_1(\omega_1)$ in $\mathcal{F}$ liegt. ∎

Beispiel 4.14 *Seien* $\mathcal{F}_i = \{f : \mathbb{R} \to \mathbb{R} \,|\, f \text{ isoton und konkav } \}$, $i = 1, 2$, *und sei* $\mathcal{F}_{cv} = \{f : \mathbb{R}^2 \to \mathbb{R} \,|\, f \text{ isoton und konkav } \}$. *Sind* μ_i, ν_i, $i = 1, 2$, *stetige Kapazitäten, so gilt*

$$\mu_i \preceq_{\mathcal{F}_i} \nu_i,\ i = 1, 2. \qquad \Longleftrightarrow \qquad \mu_1 \otimes \mu_2 \preceq_{\mathcal{F}_{cv}} \nu_1 \otimes \nu_2.$$

Beweis: Sei $f(\cdot, \omega_2)$ isoton und konkav für jedes $\omega_2 \in \Omega_2$. Im vorherigen Beispiel haben wir gesehen, daß dann auch $\int_{\Omega_2} f(\cdot, \omega_2) d\mu_2(\omega_2)$ isoton ist. Wir müssen uns nur noch überlegen, daß $\int_{\Omega_2} f(\cdot, \omega_2) d\mu_2(\omega_2)$ auch konkav ist. Seien also $x, y \in \mathbb{R}$ und $\lambda \in [0, 1]$. Wegen der Konkavität von f in der ersten Variablen und der Monotonie des Choquet-Integrals ist dann

$$\int_{\mathbb{R}} f(\lambda x + (1-\lambda)y, \omega_2) d\mu_2(\omega_2) \ge \int_{\mathbb{R}} [\lambda f(x, \omega_2) + (1-\lambda) f(y, \omega_2)]\, d\mu_2(\omega_2). \quad (4.8)$$

Da $f(x, \omega_2)$ und $f(y, \omega_2)$ beide isoton in ω_2 sind, sind sie komonoton. Daraus und aus (2.5) folgt

$$\begin{aligned} &\int_{\mathbb{R}} [\lambda f(x, \omega_2) + (1-\lambda) f(y, \omega_2)]\, d\mu_2(\omega_2) \\ &\qquad = \lambda \int_{\mathbb{R}} f(x, \omega_2) d\mu_2(\omega_2) + (1-\lambda) \int_{\mathbb{R}} f(y, \omega_2) d\mu_2(\omega_2). \end{aligned} \quad (4.9)$$

Aus (4.8) und (4.9) zusammen folgt, daß $\int_{\Omega_2} f(\cdot, \omega_2) d\mu_2(\omega_2)$ auch konkav ist. Somit ist $\mathcal{F}_1$ abgeschlossen gegen Integration. Da $\mathcal{F}_{cv} \subset \mathcal{F} = \{f : \mathbb{R}^2 \to \mathbb{R} \,|\, f$ isoton und konkav in jeder Variablen $\}$ ist, folgt mit Satz 4.4 die „$\Rightarrow$“-Richtung.

„$\Leftarrow$“ folgt, da für jedes $f_1 \in \mathcal{F}_1$ die Funktion $f : \Omega \to \mathbb{R}$, $f(\omega_1, \omega_2) = f_1(\omega_1)$ in $\mathcal{F}$ liegt. ∎

4.5 Anwendung stochastischer Dominanz in AU und CEU

Wir legen in diesem Abschnitt dar, wie die oben hergeleiteten Ergebnisse über stochastische Dominanz von Kapazitäten in AU- und CEU-Modellen angewendet werden können.

Sei $((S,\mathcal{A}),(\mathcal{C},\mathcal{D}),\mathcal{F},\preceq)$ ein Entscheidungsproblem unter Unsicherheit und der ET ein CEU-Maximierer mit Kapazität μ und Nutzenfunktion u. Die Kapazität μ sei bekannt, von der Nutzenfunktion u sei nur bekannt, daß sie zu einer Menge $\mathcal{U}$ von Nutzenfunktionen gehört. Eine Aktion f wird dann einer Aktion g vorgezogen, wenn

$$\int (u \circ f)\, d\mu \;\geq\; \int (u \circ g)\, d\mu$$

gilt, bzw.

$$\int u\, d\mu^f \;\geq\; \int u\, d\mu^g, \tag{4.10}$$

wobei μ^f das Bild von μ unter f bezeichnet (vgl. Anhang A.2). Wenn nun $\mu^f \succeq_{\mathcal{U}} \mu^g$ gilt, so ist dies ja gleichbedeutend mit der Gültigkeit von (4.10) *für alle* $u \in \mathcal{U}$. In diesem Fall ist also eine Entscheidung zwischen den beiden Alternativen möglich, obwohl die exakte Form der Nutzenfunktion nicht bekannt ist.

Sei $((\mathcal{C},\mathcal{D}),\mathcal{P},\preceq)$ ein Entscheidungsproblem unter Risiko und der ET ein AU-Maximierer mit Nutzenfunktion u und Verzerrungsfunktion q. Die Verzerrungsfunktion q sei bekannt, von der Nutzenfunktion u sei nur bekannt, daß sie zu einer Menge $\mathcal{U}$ von Nutzenfunktionen gehört. Eine Lotterie P_1 wird dann einer Lotterie P_2 vorgezogen, wenn

$$\int u\, d(q \circ P_1) \;\geq\; \int u\, d(q \circ P_2)\,. \tag{4.11}$$

In diesem Fall muß man also entsprechend versuchen, Resultate über stochastische Dominanz auf die Kapazitäten $q \circ P_1$ und $q \circ P_2$ anzuwenden.

Nachteilig für die Anwendbarkeit wirkt sich natürlich aus, daß in CEU die Kapazität μ, in AU die Verzerrungsfunktion q exakt bekannt sein muß, um die Ergebnisse anwenden zu können. In der Praxis wird z.B. oft nur bekannt sein, daß q zu einer gewissen Klasse von Verzerrungsfunktionen gehört. Interessante Klassen von Verteilungsfunktionen sind unter anderem die folgenden Klassen (vgl. Quiggin 1982, Segal 1987, Chew, Karni und Safra 1987)

- $Q_1 = \Big\{ q : [0,1] \to [0,1] \,|\, q(0) = 0,\ q(1) = 1,\ q$ streng monoton wachsend und stetig $\Big\}$
- $Q_X = \Big\{ q \in Q_1 \,|\, q \text{ konvex} \Big\}$

- $Q_V = \Big\{ q \in Q_1 \,|\, q \text{ konkav} \Big\}$

Wir betrachten zunächst Mengen-Dominanzrelationen zwischen Kapazitäten. Für stochastische Dominanz erster Ordnung auf der reellen Achse findet sich der folgende Satz bereits in Quiggin (1982).

Satz 4.5 *Sei $\mathcal{U}$ eine Klasse von Nutzenfunktionen, $\succeq_{\mathcal{U}}$ eine Mengen-Dominanzrelation und P_1 und P_2 Lotterien. Dann gilt*

$$P_1 \succeq_{\mathcal{U}} P_2 \iff q \circ P_1 \succeq_{\mathcal{U}} q \circ P_2 \quad \forall q \in Q_1$$

Beweis: Sei $\mathcal{T}$ das die Mengen-Dominanzrelation $\succeq_{\mathcal{U}}$ definierende Mengensystem. Dann gilt

$$P_1 \succeq_{\mathcal{U}} P_2 \iff P_1(T) \geq P_2(T) \ \forall T \in \mathcal{T}$$

und

$$q \circ P_1 \succeq_{\mathcal{U}} q \circ P_2 \iff q \circ P_1(T) \geq q \circ P_2(T) \ \forall T \in \mathcal{T}.$$

Für jede streng monoton wachsende Verzerrungsfunktion q und für jedes $T \in \mathcal{T}$ gilt aber

$$P_1(T) \geq P_2(T) \iff q \circ P_1(T) \geq q \circ P_2(T).$$

Hieraus folgt direkt die Behauptung. ■

Satz 4.5 kann wie folgt interpretiert werden. Zieht jeder EU-Maximierer mit Nutzenfunktion u in $\mathcal{U}$ die Lotterie P_1 gegenüber P_2 vor, so tut dies auch jeder AU-Maximierer mit Nutzenfunktion u in $\mathcal{U}$ und Verzerrungsfunktion q in Q_1. Die Umkehrung gilt natürlich sowieso. Bemerkenswert daran ist, daß die Klasse Q_1 eine sehr allgemeine Klasse von Verzerrungsfunktionen ist. Ist die stochastische Dominanzrelation $\succeq_{\mathcal{U}}$ also eine Mengen-Dominanzrelation, so sind „fast“ keine weiteren Kenntnisse über die Verzerrungsfunktion nötig.

Die Situation wird komplizierter, wenn die Dominanzrelation $\succeq_{\mathcal{U}}$ keine Mengen-Dominanzrelation ist. Man erhält ein zu Satz 4.5 analoges Resultat, wenn die Dominanzrelation stochastische Dominanz zweiten Grades und die Klasse der Verzerrungsfunktionen eine Teilmenge der konvexen Verzerrungsfunktionen ist. Konvexe Verzerrungsfunktion und konkave Nutzenfunktion sind äquivalent zu starker Risikoaversion, siehe Chew, Karni und Safra (1987).

Satz 4.6 *Sei $\mathcal{U}_2 = \{u : \mathbb{R} \to \mathbb{R}_+ \,|\, u$ zweimal differenzierbar, $u' \geq 0,\ u'' \leq 0\}$, und sei $Q_2 = \{q \in Q_1 \,|\, q$ zweimal differenzierbar, $q' \geq 0,\ q'' \geq 0\}$. Seien weiter P_1 und P_2 Lotterien. Dann gilt*

$$q \circ P_1 \succeq_{\mathcal{U}_2} q \circ P_2 \quad \forall q \in Q_2 \iff P_1 \succeq_{\mathcal{U}_2} P_2$$

Beweis: Sei $\overline{F}(t) = P_1([t,\infty))$, $t \in \mathbb{R}$ und $\overline{G}(t) = P_2([t,\infty))$, $t \in \mathbb{R}$. Wir definieren die Quasi-Inverse h^{-1} einer monoton fallenden Funktion h wie üblich durch $h^{-1}(t) = \inf\{\, x \mid h(x) \leq t \,\}$. Man kann zeigen (siehe Röell 1987), daß

$$\int_0^x [\overline{F}(t) - \overline{G}(t)]dt \geq 0 \;\forall x \in \mathbb{R} \iff \int_z^1 [\overline{F}^{-1}(s) - \overline{G}^{-1}(s)]ds \geq 0 \;\forall z \in [0,1] \;.$$

Aus Beispiel 4.9 ist bekannt, daß

$$q \circ P_1 \succeq_{\mathcal{U}_2} q \circ P_2 \iff \int_0^x [q \circ \overline{F}(t) - q \circ \overline{G}(t)]dt \geq 0 \quad \forall x \in \mathbb{R}$$

Da $(q \circ \overline{F})^{-1} = \overline{F}^{-1} \circ q^{-1}$ ist, erhalten wir

$$\begin{aligned} & \int_0^x [q \circ \overline{F}(t) - q \circ \overline{G}(t)]dt \geq 0 \quad \forall x \in \mathbb{R} \\ \iff & \\ & \int_z^1 [\overline{F}^{-1}(q^{-1}(s)) - \overline{G}^{-1}(q^{-1}(s))]ds \geq 0 \quad \forall z \in [0,1] \\ \iff & \\ & \int_x^1 [\overline{F}^{-1}(u) - \overline{G}^{-1}(u)]q'(u)du \geq 0 \quad \forall x \in [0,1] \end{aligned}$$

Partielle Integration liefert dann

$$\begin{aligned} & \int_x^1 [\overline{F}^{-1}(u) - \overline{G}^{-1}(u)]q'(u)du \\ & \quad = \int_x^1 [\overline{F}^{-1}(s) - \overline{G}^{-1}(s)]ds \; q'(x) + \int_x^1 \int_t^1 [\overline{F}^{-1}(s) - \overline{G}^{-1}(s)]ds \; q''(t)dt \end{aligned}$$

Gilt $P_1 \succeq_{\mathcal{U}_2} P_2$, so ist

$$\int_x^1 [\overline{F}^{-1}(s) - \overline{G}^{-1}(s)]ds \;\geq 0 \quad \forall x \in [0,1]$$

und da $q'(x) \geq 0$ und $q''(x) \geq 0$ sind, erhalten wir schließlich $q \circ P_1 \succeq_{\mathcal{U}_2} q \circ P_2$, wie behauptet. Die Umkehrung ist klar, da die identische Verzerrungsfunktion in Q_2 liegt. ∎

Kapitel 5

Risikoaversion und Unsicherheitsaversion

In diesem Kapitel wollen wir Risikoaversion und Unsicherheitsaversion in AU- und CEU-Modellen darstellen. Dazu muß zunächst einmal geklärt werden, was unter Risikoaversion bzw. Unsicherheitsaversion zu verstehen ist. Der Begriff Risikoaversion bezieht sich dabei auf Entscheidungsprobleme unter Risiko, der Begriff der Unsicherheitsaversion dagegen auf Entscheidungsprobleme unter Unsicherheit. Dementsprechend werden wir zunächst Risikoaversion in AU charakterisieren, in einem weiteren Abschnitt dann Unsicherheitsaversion in CEU.

5.1 Risikoaversion in AU

Voraussetzung: Es sei $((\mathbb{R}, \mathcal{B}), \mathcal{P}, \preceq)$ ein Entscheidungsproblem unter Risiko. Wir betrachten die W-Maße in $\mathcal{P}$ als Verteilungen von Zufallsvariablen auf einem gemeinsamen W-Raum $(S, \mathcal{A}, P)$.

In EU gibt es verschiedene Definitionen für den Begriff der Risikoaversion. Die Theorie der Risikoaversion in EU wurde von Pratt (1964) und Arrow (1965, 1971) entwickelt. Sie nennen einen ET risikoavers, wenn er den Erwartungswert einer Lotterie der Lotterie selbst vorzieht. Diesen Begriff der Risikoaversion wollen wir im folgenden als *schwache Risikoaversion* bezeichnen.

Definition 5.1 (Schwache Risikoaversion) *Ein ET heißt* schwach risikoavers*, wenn $E(X) \succeq X$ für alle integrierbaren Zufallsvariablen X mit $P^X \in \mathcal{P}$ gilt.*

Eine andere Definition des Begriffes Risikoaversion wurde von Rothschild und Stiglitz (1970) gegeben. Sie nennen eine Verteilung P^X riskanter als eine andere Verteilung P^Y, wenn die Verteilung P^X mehr Gewicht in den Enden hat. Um dies zu präzisieren, benötigen wir die folgende Definition.

Definition 5.2 (Erwartungswerterhaltende Streckung, Rothschild und Stiglitz 1970) *Seien X und Y zwei Zufallsvariablen mit Verteilungsfunktionen F_X und F_Y. Y unterscheidet sich von X durch eine* erwartungswerterhaltende Streckung *(engl.* mean preserving spread*), wenn die folgenden Bedingungen erfüllt sind:*

$$\int_{-\infty}^{x} [F_Y(t) - F_X(t)]\, dt \geq 0 \qquad \text{für alle } x \in \mathbb{R},$$
$$\int_{-\infty}^{\infty} [F_Y(t) - F_X(t)]\, dt = 0.$$

Wir sagen auch „Y unterscheidet sich von X durch einen MPS" und schreiben dafür $X \stackrel{MPS}{\longrightarrow} Y$.

Starke Risikoaversion wird dann wie folgt definiert.

Definition 5.3 (Starke Risikoaversion) *Ein ET heißt* stark risikoavers*, wenn für alle X, Y gilt:*

$$X \stackrel{MPS}{\longrightarrow} Y \Rightarrow X \succeq Y.$$

Da sich jede Zufallsvariable X von $E(X)$ durch einen MPS unterscheidet ist ein stark risikoaverser ET auch schwach risikoavers. In EU gilt sogar, daß ein schwach risikoaverser ET auch stark risikoavers ist. Die beiden Begriffe von Risikoaversion sind also in EU äquivalent und gleichwertig zur Konkavität der Nutzenfunktion u. In AU fallen schwache und starke Risikoaversion nicht länger zusammen. Wir zitieren zunächst die Charakterisierung der starken Risikoaversion, die von Chew, Karni und Safra (1987) stammt.

Satz 5.1 (Chew, Karni und Safra (1987)) *Ein AU-ET ist genau dann stark risikoavers, wenn u konkav und q konvex ist.*

Bevor wir zur schwachen Risikoaversion kommen, erklären wir, was man in AU unter einem Pessimisten bzw. einem Optimisten versteht. Dies wird durch den folgenden Satz vorbereitet.

Satz 5.2 *Es gelten die folgenden Aussagen:*

(i) Ist $q(p) \leq p$ für alle $p \in [0,1]$, so gilt $E_{q \circ P}(X) \leq E_P(X)$.

(ii) Ist $q(p) \geq p$ für alle $p \in [0,1]$, so gilt $E_{q \circ P}(X) \geq E_P(X)$.

Beweis: Wir beweisen nur (i). (ii) zeigt man analog. Es ist

$$\begin{aligned} E_{q \circ P}(X) &= \int_0^\infty q \circ P(X \geq t)\, dt + \int_{-\infty}^0 [q \circ P(X \geq t) - 1]\, dt \\ &\leq \int_0^\infty P(X \geq t)\, dt + \int_{-\infty}^0 [P(X \geq t) - 1]\, dt \\ &= E_P(X). \end{aligned}$$

■

Daher nennen wir einen AU-ET, dessen Verzerrungsfunktion $q(p) \leq p$ für alle $p \in [0,1]$ erfüllt, einen *Pessimisten.* Entprechend nennen wir einen AU-ET mit einer Verzerrungsfunktion q, für die $q(p) \geq p$ für alle $p \in [0,1]$ ist, einen *Optimisten.*

Die Charakterisierung von schwacher Risikaversion in AU ist schwieriger. Die folgenden Sätze aus Chateauneuf und Cohen (1990) geben notwendige und hinreichende Bedingung für schwache Risikoaversion in AU.

Satz 5.3 (Chateauneuf und Cohen 1990) *In einem AU-Modell seien u und q differenzierbar.*

(a) Ist der AU-ET schwach risikoavers, so gelten die folgenden Aussagen.

(i) $q(p) \leq p$ für alle $p \in [0,1]$.

(ii) Ist u zweimal differenzierbar und gibt es $p_0 \in [0,1]$ mit $q(p_0) = p_0$, so ist u konkav.

(iii) Es gibt ein $k \geq 1$, so daß $u'(x) \leq k \dfrac{u(x) - u(y)}{x - y}$ für alle $y < x$.

Ist (iii) erfüllt und ist $q(p) \leq p^k$ für alle $p \in [0,1]$, so ist der AU-ET schwach risikoavers.

(b) Ist der AU-ET schwach risikofreudig, so gelten die folgenden Aussagen.

(i) $q(p) \geq p$ für alle $p \in [0,1]$.

(ii) Ist u zweimal differenzierbar und gibt es $p_0 \in [0,1]$ mit $q(p_0) = p_0$, so ist u konvex.

(iii) Es gibt ein $h \geq 1$, so daß $u'(y) \leq h \dfrac{u(x) - u(y)}{x - y}$ für alle $y < x$.

Ist (iii) erfüllt und ist $q(p) \geq 1 - (1-p)^h$ für alle $p \in [0,1]$, so ist der AU-ET schwach risikofreudig.

Bedingung (a)(i) besagt also, daß ein schwach risikoaverser AU-ET ein Pessimist sein muß. Insbesondere folgt daraus, daß ein AU-ET mit S-förmiger Verzerrungsfunktion nicht schwach risikoavers sein kann. Die Bedingung (a)(ii) wird in Chateauneuf und Cohen (1990) wie folgt interpretiert: Ist der AU-ET risikoavers und lokal (nämlich für $p = p_0$) ein EU-Maximierer, so muß u konkav sein. Bedingung (a)(iii) schließlich fordert, daß u nicht „zu konvex“ ist. Ist $k = 1$ so ist u offensichtlich konkav.

Ein ET kann also auch dann schwach risikoavers sein, wenn u nicht konkav oder sogar konvex ist; er muß lediglich pessimistisch genug sein. Darüberhinaus kann eine Nutzenfunktion gleichzeitig (a)(iii) und (b)(iii) erfüllen. In diesem Fall ist der ET schwach risikoavers, wenn $q(p) \leq p^k$, $p \in [0,1]$, ist, und schwach risikofreudig, wenn $q(p) \geq 1 - (1-p)^h$, $p \in [0,1]$, ist.

5.2 Unsicherheitsaversion in CEU

Die erste Definition von Unsicherheitsaversion stammt von Schmeidler (1989, erste Version 1982) im Rahmen des Anscombe-Aumann Ansatzes (vgl. Anhang B.1). Wir erinnern noch einmal daran, daß in diesem Modell der Konsequenzenraum $\mathcal{C}$ die Menge aller Lotterien mit endlichem Träger auf einem Raum Γ ist und die Nutzenfunktion U auf $\mathcal{C}$ durch den Erwartungsnutzen der Lotterie in bezug auf eine Nutzenfunktion u auf Γ gegeben ist. Insbesondere ist also U affin-linear auf $\mathcal{C}$.

Definition 5.4 (Unsicherheitsaversion nach Schmeidler 1989) *In einem CEU-Modell nach Schmeidler (1989) heißt der ET* unsicherheitsavers*, wenn für alle $f, g, h \in \mathcal{F}$ und $\alpha \in (0,1)$ gilt*

$$f \succeq h,\ g \succeq h \Rightarrow \alpha f + (1-\alpha) g \succeq h.$$

Von Wakker (1990a) stammt eine weitere Definition von Unsicherheitsaversion (von Wakker als Pessimismus bezeichnet) in Schmeidlers Modell. Er definiert Unsicherheitsaversion, indem er das komonotone Unabhängigkeitsaxiom (vgl. Anhang B.1) wie folgt abschwächt.

Definition 5.5 (Unsicherheitsaversion nach Wakker 1990a) *In einem CEU-Modell nach Schmeidler (1989) heißt der ET* unsicherheitsavers*, wenn für alle $f, g, h \in \mathcal{F}$ und $\alpha \in (0,1)$ gilt*

$$f \succ g,\ g \text{ und } h \text{ komonoton} \Rightarrow \alpha f + (1-\alpha) h \succ \alpha g + (1-\alpha) h.$$

Sind die Aktionen f und h nicht komonoton, so kommt es zu einer Verringerung der Unsicherheit durch „hedging“. Bei den komonotonen Aktionen g und h ist dies jedoch nicht der Fall. Ein unsicherheitsaverser ET wird daher die α-Mischung aus f und h (mit möglicherweise verringerter Unsicherheit) der α-Mischung aus g und h vorziehen.

Es zeigt sich jedoch, daß beide Definitionen von Unsicherheitsaversion äquivalent sind.

Satz 5.4 (Schmeidler 1989, Wakker 1990a) *In einem CEU-Modell nach Schmeidler (1989) sind die folgenden Aussagen äquivalent:*

(i) Der ET ist unsicherheitsavers im Sinne von Schmeidler.

(ii) Der ET ist unsicherheitsavers im Sinne von Wakker.

(iii) μ ist stark superadditiv.

Weitere Definitionen von Unsicherheitsaversion finden sich in Chateauneuf (1991, 1993). Chateauneuf betrachtet ein CEU-Modell $((S, \mathcal{A}), (\mathbb{R}, \mathcal{B}), \mathcal{F}, \preceq)$, in dem die Nutzenfunktion u linear und nicht konstant ist. In Chateauneuf (1991) ist $\mathcal{F}$ die Menge aller beschränkten, nichtnegativen Zufallsvariablen, in Chateauneuf (1993) ist $\mathcal{F} = \mathcal{F}^b$ die Menge aller beschränkten reellen Zufallsvariablen. Seine Definition von Unsicherheitsaversion entspricht im wesentlichen der Wakkers:

Definition 5.6 (Unsicherheitsaversion nach Chateauneuf 1993) *In einem CEU-Modell mit $\mathcal{C} = (\mathbb{R}, \mathcal{B})$ und linearer und nicht konstanter Nutzenfunktion heißt ein ET* unsicherheitsavers*, wenn für alle $f, g, h \in \mathcal{F}^b$ gilt*

$$f \sim g, \;\; g \text{ und } h \text{ komonoton} \Rightarrow f + h \succeq g + h.$$

Auch diese Definition von Unsicherheitsaversion ist äquivalent zur starken Superadditivität von μ.

Satz 5.5 (Chateauneuf 1993) *In einem CEU-Modell mit $\mathcal{C} = (\mathbb{R}, \mathcal{B})$ und linearer und nicht konstanter Nutzenfunktion sind die folgenden Aussagen äquivalent.*

(i) Der ET ist unsicherheitsavers im Sinne von Chateauneuf.

(ii) μ ist stark superadditiv.

Kapitel 6

Dekomposition multivariater Nutzenfunktionen in CEU und AU

In praktischen Anwendungen können die Konsequenzen der verschiedenen Entscheidungsalternativen oft nicht durch ein einzelnes Attribut beschrieben werden, sondern nur durch eine Menge von Attributen. Der Konsequenzenraum ist somit ein kartesisches Produkt (oder eine Teilmenge eines solchen) von meheren Attributräumen. Die Bestimmung von multiattributiven Nutzenfunktionen ist bereits in EU schwierig (siehe z.B. Keeney/Raiffa 1976). Es besteht daher die Notwendigkeit, die Anzahl der Fragen an den ET möglichst gering zu halten. Die Bestimmung der Nutzenfunktion läßt sich wesentlich vereinfachen, wenn die multivariate Nutzenfunktion in Nutzenfunktionen möglichst geringer Dimension zerlegbar ist. In EU sind Zerlegungen gebräuchlich, bei denen die Nutzenfunktion eine Summe von Produkten, ein Produkt oder eine Summe von eindimensionalen Nutzenfunktionen ist. Diese Nutzenfunktionen sind als *multilineare, multiplikative* und *additive Nutzenfunktionen* bekannt. Ist die Nutzenfunktion auf eine dieser Arten zerlegbar, so reduziert sich die Bestimmung der multiattributiven Nutzenfunktion auf die Bestimmung der einattributiven Nutzenfunktionen und einer Anzahl von Skalierungskonstanten. In EU sind viele notwendige oder hinreichenden Bedingungen für derartige Zerlegungen bekannt, siehe z.B. Keeney/Raiffa (1976), Fishburn (1982), von Stengel (1993) sowie für allgemeinere Modelle Miyamoto (1988). Es handelt sich dabei um Bedingungen an die Präferenzen des ETs, die in Anwendungssituationen überprüft werden können.

Da in den hier untersuchten Modellen der Entscheidungstheorie nicht nur die Nutzenfunktion des ETs zu bestimmen ist, sondern außerdem die Verzerrungsfunktion (in AU), bzw. eine Kapazität (in CEU), ist die Notwendigkeit von Bedingungen, die diese Bestimmungen erleichtern, um so größer.

Weiter geben wir Charakterisierungen bivariater Risikohaltungen in AU. In EU kann ein ET mit einer multiplikativen Nutzenfunktion lediglich drei verschiedene Typen von bivariater Risikohaltung zeigen, nämlich bivariate Risikoaversion, bivariate Risikofreude und bivariate Risikoneutralität, je nachdem, ob der Parameter β der multiplikativen Nutzenfunktion kleiner, größer oder gleich Null ist. Dagegen kann in AU der ET für jeden Wert des Parameters β bivariat risikoavers sein. Weiter muß ein AU-Maximierer mit einer multiplikativen Nutzenfunktion keine der drei oben erwähnten bivariaten Risikohaltungen zeigen. Wir verallgemeinern diese Ergebnisse auch auf den Fall von mehr als zwei Attributen.

In einem weiteren Abschnitt geben wir eine Charakterisierung von Korrelationsaversion in AU. Korrelationsaversion kann auch als Verschärfung von bivariater Risikoaversion gedeutet werden, so daß man Korrelationsaversion auch als *starke bivariate Risikaversion* bezeichnen könnte. Wir werden zeigen, daß Korrelationsaversion dadurch charakterisiert werden kann, daß die zweite gemischte Ableitung der Nutzenfunktion kleiner als Null und die Verzerrungsfunktion konvex ist. Dies steht in Analogie zur Charakterisierung von starker Risikoaversion gemäß Chew, Karni und Safra (1987) (vergleiche Satz 5.1).

Dieses Kapitel stützt sich mit Ausnahme von Abschnitt 6.5 auf Dyckerhoff (1994).

6.1 Grundlagen

In diesem Kapitel gehen wir davon aus, daß der Konsequenzenraum ein kartesisches Produkt $\mathcal{C} = \times_{i=1}^{n} \mathcal{C}_i$ ist. Wir nehmen an, daß die $\mathcal{C}_i$, $i = 1, \ldots, n$, zusammenhängende, separable topologische[1] Räume sind. Der Raum $\mathcal{C}$ sei mit der Produkttopologie versehen. Dann ist $\mathcal{C}$ ebenfalls ein zusammenhängender, separabler topologischer Raum. Wir nennen i ein *Attribut* und den Raum $\mathcal{C}_i$ einen *Attributraum*. Wir bezeichnen die Menge $\{1, \ldots, n\}$ aller Attribute mit N und das Komplement einer Teilmenge I von N mit I^c. Für eine gegebene Teilmenge I von Attributen schreiben wir $\mathcal{C}_I$ für den Raum $\times_{i \in I} \mathcal{C}_i$ und $\mathcal{C}_{-I}$ für den Raum $\times_{i \notin I} \mathcal{C}_i$. Analog schreiben wir x_I für die Elemente von $\mathcal{C}_I$ und x_{-I} für die Elemente von $\mathcal{C}_{-I}$. Ein Element $x \in \mathcal{C}$ schreiben wir auch als $x = (x_I, x_{-I})$. Ist $I = \{i\}$, so schreiben wir x_i statt $x_{\{i\}}$ und x_{-i} statt $x_{-\{i\}}$.

Ein Attribut i heißt *relevant*, wenn es ein $x_{-i} \in \mathcal{C}_{-i}$ und $y_i, z_i \in \mathcal{C}_i$ gibt, so daß $(x_{-i}, y_i) \prec (x_{-i}, z_i)$. Wir setzen voraus, daß alle Attribute relevant sind.

Eine Nutzenfunktion u heißt *zerlegbar*, wenn sie dargestellt werden kann als

$$u(x_1, \ldots, x_n) = F(u_1(x_1), \ldots, u_n(x_n)),$$

wobei die u_i, $i = 1, \ldots, n$, Nutzenfunktionen auf den Attributräumen sind und $F : \mathbb{R}^n \to \mathbb{R}$ eine Funktion ist.

[1] Diese und weitere in diesem Kapitel verwendeten topologischen Begriffe und Ergebnisse finden sich z.B. in Willard (1970).

Wir nehmen in diesem Kapitel generell an, daß die Nutzenfunktion u beschränkt ist und daß minimale und maximale Konsequenzen x^- und x^+ existieren. In einem nichtdegeneriertem CEU- oder AU-Modell ist die Nutzenfunktion eindeutig bis auf positiv affine Transformationen (Satz 3.1). Somit kann die Nutzenfunktion so normiert werden, daß $u(x^-) = 0$ und $u(x^+) = 1$ ist. Ist $x = (x_I, x_{-I}^-)$, so schreiben wir kürzer $u(x_I)$ statt $u(x_I, x_{-I}^-)$. Die Funktion $u_I : \mathcal{C}_I \to [0,1]$ ist definiert durch $u_I(x_I) = u(x_I)/\alpha_I$, wobei $\alpha_I = \sup\{u(x_I)|x_I \in \mathcal{C}_I\}$ ist. Dann ist u_I eine Nutzenfunktion auf $\mathcal{C}_I$ und so normiert, daß $u_I(x_I^-) = 0$ und $\sup\{u_I(x_I)|x_I \in \mathcal{C}_I\} = 1$ ist.

In einigen Sätzen setzen wir voraus, daß u stetig ist. In diesem Fall sind insbesondere auch alle Funktionen u_I, $\emptyset \neq I \subset N$, stetig.

Wir formulieren jetzt die Bedingungen, die in EU bei der Zerlegung von Nutzenfunktionen eine Rolle spielen.

Definition 6.1 (Präferenzunabhängigkeit) *Eine Menge I von Attributen heißt* präferenzunabhängig*, wenn Präferenzen über sichere Konsequenzen mit gemeinsamen, festem Niveau $x_{-I} \in \mathcal{C}_{-I}$ nicht von diesem festen Niveau abhängen.*

Dies bedeutet, daß alle bedingten Präferenzen über sichere Konsequenzen auf $\mathcal{C}_I$ identisch sind. Definiert man durch $x_I \preceq_I y_I \iff [(x_I, z_{-I}) \preceq (y_I, z_{-I})$ für ein $z_{-I} \in \mathcal{C}_{-I}]$ eine Relation $\preceq_I$ auf $\mathcal{C}_I$, so ist $\preceq_I$ wohldefiniert. Ist I präferenzunabhängig, so ist $u_I(x_I) = u(x_I)/u(x_I^+)$. Somit ist in diesem Fall $u_I(x_I^-) = 0$ und $u_I(x_I^+) = 1$, und die Skalierungskonstante α_I ist durch $u(x_I^+)$ gegeben.

Bemerkung: Ist die Menge I von Attributen präferenzunabhängig, so sind die Funktionen $u(\cdot, x_{-I})$, $x_{-I} \in \mathcal{C}_{-I}$, paarweise komonoton.

Da $\mathcal{C}_I$ zusammenhängend ist, ist – sofern u_I stetig ist – das Bild von $\mathcal{C}_I$ unter u_I gleich $[0,1]$.

Definition 6.2 (Nutzenunabhängigkeit) *Eine Menge I von Attributen heißt* nutzenunabhängig*, wenn Präferenzen über Lotterien oder Aktionen mit gemeinsamen, festen und sicherem Niveau $x_{-I} \in \mathcal{C}_{-I}$ nicht von diesem festen Niveau abhängen.*

Ist die Menge I von Attributen nutzenunabhängig, so ist I auch präferenzunabhängig. Dies ist klar, da jede sichere Konsequenz als eine Aktion aufgefasst werden kann, die für jeden Umweltzustand diese Konsequenz liefert.

Definition 6.3 (Randverteilungsbedingung) *Sei N eine disjunkte Vereinigung von nichtleeren Mengen $I_1, \ldots, I_k$. Die Mengen von Attributen $I_1, \ldots, I_k$ erfüllen die* Randverteilungsbedingung*, wenn Präferenzen über Lotterien nur von den Randverteilungen der Lotterien bezüglich der Attribute $\mathcal{C}_{I_1}, \ldots \mathcal{C}_{I_k}$ abhängen.*

Eine weitere Klasse von Bedingungen bezeichnet man als *bivariate Risikohaltung*. Diese Bedingungen wurden zuerst von Richard (1975) betrachtet.

Definition 6.4 (Bivariate Risikohaltung) *Ein ET heißt* bivariat risikoavers *in bezug auf I und I^c ($I(BRA)I^c$), wenn I und I^c präferenzunabhängig sind und für alle $x_I, y_I \in \mathcal{C}_I$ und $x_{-I}, y_{-I} \in \mathcal{C}_{-I}$ mit $x_I \prec_I y_I$ und $x_{-I} \prec_{-I} y_{-I}$*

$$L_1 = \begin{bmatrix} (x_I, x_{-I}) & (y_I, y_{-I}) \\ 0.5 & 0.5 \end{bmatrix} \preceq \begin{bmatrix} (x_I, y_{-I}) & (y_I, x_{-I}) \\ 0.5 & 0.5 \end{bmatrix} = L_2$$

gilt, d.h., falls er die 50-50-Lotterie, die zumindest in einem der beiden Attribute eine gute Konsequenz liefert, vorzieht.

Ein ET heißt bivariat risikofreudig *in Bezug auf I und I^c ($I(BRS)I^c$), wenn I und I^c präferenzunabhängig sind und $L_1 \succeq L_2$ für alle $x_I, y_I \in \mathcal{C}_I$ und $x_{-I}, y_{-I} \in \mathcal{C}_{-I}$ mit $x_I \prec_I y_I$ und $x_{-I} \prec_{-I} y_{-I}$ gilt.*

Ein ET heißt bivariat risikoneutral *in Bezug auf I und I^c ($I(BRN)I^c$), wenn I und I^c präferenzunabhängig sind und $L_1 \sim L_2$ für alle $x_I, y_I \in \mathcal{C}_I$ und $x_{-I}, y_{-I} \in \mathcal{C}_{-I}$ mit $x_I \prec_I y_I$ und $x_{-I} \prec_{-I} y_{-I}$ gilt.*

Wir fassen nun noch einmal die Voraussetzungen dieses Kapitels zusammen:

Voraussetzung:

CEU-Modelle: $((S, \mathcal{A}), (\mathcal{C}, \mathcal{D}), \mathcal{F}, \preceq)$ sei ein nichtdegeneriertes CEU-Modell.

AU-Modelle: $((\mathcal{C}, \mathcal{D}), \mathcal{P}, \preceq)$ sei ein nichtdegeneriertes AU-Modell. Die Verzerrungsfunktion q sei stetig und streng monoton wachsend.

CEU- und AU-Modelle: Die Attributräume $\mathcal{C}_i$, $i \in N$, sind zusammenhängende, separable topologische Räume. Der Konsequenzenraum $\times_{i=1}^n \mathcal{C}_i$ ist mit der Produkttopologie versehen. Es gibt minimale und maximale Konsequenzen x^- und x^+. Jedes Attribut ist relevant. Die Nutzenfunktion u ist stetig.

6.2 Multiplikative und multilineare Nutzenfunktionen

In diesem Abschnitt untersuchen wir unter welchen Bedingungen eine Nutzenfunktion in einem AU- oder CEU-Modell multilinear oder multiplikativ ist.

Eine Nutzenfunktion heißt *multiplikativ*, falls sie darstellbar ist als

$$u(x_1, \ldots, x_n) = \sum_{\emptyset \neq I \subset N} \beta^{|I|-1} \prod_{i \in I} \alpha_i u_i(x_i) \ , \tag{6.1}$$

wobei die Nutzenfunktionen u_i, $i = 1, \ldots n$, so normiert sind, daß $u_i(x_i^-) = 0$, $u_i(x_i^+) = 1$ für jedes $i = 1, \ldots, n$ erfüllt ist und $\alpha_1, \ldots, \alpha_n, \beta \in \mathbb{R}$ sind. Man bezeichnet diese Nutzenfunktion als multiplikativ, da für $\beta \neq 0$ Gleichung (6.1) äquivalent zu

$$1 + \beta u(x_1, \ldots, x_n) = \prod_{i=1}^{n} (1 + \beta \alpha_i u_i(x_i)) \tag{6.2}$$

ist. Für $\beta = 0$ vereinfacht sich (6.1) zu

$$u(x_1, \ldots, x_n) = \sum_{i=1}^{n} \alpha_i u_i(x_i). \tag{6.3}$$

Eine Nutzenfunktion dieser Art heißt *additiv*. Mit additiven Nutzenfunktionen werden wir uns in Abschnitt 6.4 beschäftigen.

Eine Nutzenfunktion heißt *multilinear*, falls sie darstellbar ist als

$$u(x_1, \ldots, x_n) = \sum_{\emptyset \neq I \subset N} \gamma_I \prod_{i \in I} u_i(x_i) \ , \tag{6.4}$$

wobei die Nutzenfunktionen u_i, $i = 1, \ldots n$, so normiert sind, daß $u_i(x_i^-) = 0$, $u_i(x_i^+) = 1$ für jedes $i = 1, \ldots, n$ erfüllt ist und $\gamma_I \in \mathbb{R}$, $\emptyset \neq I \subset N$. Setzt man $\gamma_I = \beta^{|I|-1} \prod_{i \in I} \alpha_i$, so sieht man, daß (6.1) ein Spezialfall von (6.4) ist.

In EU ist bekannt (siehe z.B. Keeney/Raiffa 1976), daß eine Nutzenfunktion multiplikativ ist, falls jede Teilmenge I von Attributen nutzenunabhängig ist. Ein weiteres Ergebnis besagt, daß eine Nutzenfunktion multilinear ist, falls jedes einzelne Attribut i nutzenunabhängig ist. Diese Resultate behalten ihre Gültigkeit auch für den Fall, daß der ET ein CEU- oder AU-Maximierer ist. Dies läßt sich ähnlich wie in Keeney/Raiffa (1976) beweisen. Die Beweise dieses und der folgenden Sätze finden sich in Abschnitt 6.6.

Satz 6.1 *Es liege ein nichtdegeneriertes AU- oder CEU-Modell vor. Dann gelten die folgenden zwei Aussagen:*

(i) Die Nutzenfunktion u ist multiplikativ, falls jede Teilmenge von Attributen nutzenunabhängig ist.

(ii) Falls u multiplikativ ist und falls jede Teilmenge I von Attributen präferenzunabhägig ist, so ist jede Teilmenge I von Attributen auch nutzenunabhängig.

Satz 6.2 *Es liege ein nichtdegeneriertes AU- oder CEU-Modell vor. Dann gelten die folgenden zwei Aussagen:*

(i) Die Nutzenfunktion u ist multilinear, falls jedes einzelne Attribut i nutzenunabhängig ist.

(ii) Falls u multilinear ist und falls jedes einzelne Attribut i präferenzunabhägig ist, so ist jedes einzelne Attribut i auch nutzenunabhängig.

Bevor wir uns in Abschnitt 6.4 der Untersuchung additiver Nutzenfunktionen zuwenden, geben wir im nächsten Abschnitt noch eine Interpretation des Parameters β der multiplikativen Nutzenfunktion.

6.3 Bivariate Risikohaltung in AU

In EU kann der Parameter β der multiplikativen Nutzenfunktion im Rahmen der bivariaten Risikohaltung wie folgt interpretiert werden (vgl. Richard 1975). Ist die Nutzenfunktion des ET in EU multiplikativ, so ist der ET genau dann bivariat risikoavers (risikofreudig, risikoneutral), wenn $\beta \leq 0$ ($\beta \geq 0$, $\beta = 0$) ist. In AU dagegen ist dieser Zusammenhang nicht gültig. Der Grund dafür ist, daß in AU die Präferenzen des ET nicht nur durch die Bewertung der Konsequenzen sondern auch durch die Bewertung der Wahrscheinlichkeiten bestimmt werden. Die nächsten beiden Sätze zeigen, daß in AU ein ET mit multiplikativer Nutzenfunktion für jeden Wert von β bivariat risikoavers sein kann, vorausgesetzt er ist „pessimistisch genug“. Zur Präzisierung dieses „pessimistisch genug“ betrachten wir im folgenden den Bruch $(1 - q(0.5))/q(0.5)$. Wir bezeichnen diesen Quotienten als *Pessimismus-Parameter* und kürzen ihn mit c ab. Zur Motivation dieses Namens betrachten wir eine einfache Lotterie

$$L = \begin{bmatrix} x & y \\ 0.5 & 0.5 \end{bmatrix} ,$$

wobei $x \prec y$ gelten soll. Der antizipierte Nutzen dieser Lotterie berechnet sich zu

$$\begin{aligned} AU(L) &= u(x) + (u(y) - u(x))q(0.5) \\ &= u(x)[1 - q(0.5)] + u(y)q(0.5) \ . \end{aligned}$$

Ist $c > 1$ so ist der ET pessimistisch, da die Wahrscheinlichkeit 0.5 der schlechten Konsequenz x überbewertet wird, während die Wahrscheinlichkeit 0.5 der guten Konsequenz unterbewertet wird. Ist $c < 1$ so ist der ET optimistisch.

Satz 6.3 *In einem AU-Modell mit bivariatem Konsequenzenraum $\mathcal{C} = \mathcal{C}_1 \times \mathcal{C}_2$ sei die Nutzenfunktion stetig und multiplikativ gemäß (6.1). Dann gelten die folgenden Aussagen.*

(i) Der ET ist genau dann bivariat risikoavers, wenn eine der folgenden Bedingungen erfüllt ist:

  (a) $\beta \leq 0$, $c \geq 1$,
  (b) $\beta > 0$, $c \geq 1 + \beta \max\{\alpha_1, \alpha_2\}$.

(ii) Der ET ist genau dann bivariat risikofreudig, wenn eine der folgenden Bedingungen erfüllt ist:

  (a) $\beta \geq 0$, $c \leq 1$,
  (b) $\beta < 0$, $c \leq 1 + \beta \max\{\alpha_1, \alpha_2\}$.

(iii) Der ET ist genau dann bivariat risikoneutral, wenn $\beta = 0$ und $c = 1$ ist.

Wir verallgemeinern den vorhergehenden Satz auf den Fall von mehr als zwei Attributen.

Satz 6.4 *In einem AU-Modell mit Konsequenzenraum* $\mathcal{C} = \times_{i=1}^{n} \mathcal{C}_i$ *sei die Nutzenfunktion stetig und multiplikativ gemäß (6.1). Dann gelten die folgenden Aussagen.*

(i) $I(BRA)I^c$ *gilt genau dann für jede nichtleere Teilmenge* $I \subset N$*, wenn eine der folgenden Bedingungen erfüllt ist:*

  (a) $\beta \leq 0$, $c \geq 1$,

  (b) $\beta > 0$, $c \geq \max_{j \in N} \dfrac{1+\beta}{1+\beta\alpha_j}$.

(ii) $I(BRS)I^c$ *gilt genau dann für jede nichtleere Teilmenge* $I \subset N$*, wenn eine der folgenden Bedingungen erfüllt ist:*

  (a) $\beta \geq 0$, $c \leq 1$,

  (b) $\beta < 0$, $c \leq \min_{j \in N} \dfrac{1+\beta}{1+\beta\alpha_j}$.

(iii) $I(BRN)I^c$ *gilt genau dann für jede nichtleere Teilmenge* $I \subset N$*, wenn* $\beta = 0$ *und* $c = 1$ *ist.*

Bemerkung 1: Teil (*iii*) des Satzes gibt eine hinreichende Bedingung für eine additive Nutzenfunktion. Dabei wird vorausgesetzt, daß die Nutzenfunktion multiplikativ ist. Wir werden in Abschnitt 6.4 zeigen, daß auch ohne diese Voraussetzung bivariate Risikoneutralität eine additive Nutzenfunktiom impliziert.

Bemerkung 2: Teil (*iii*) besagt ebenfalls, daß für die Verzerrungsfunktion des ET $q(0.5) = 0.5$ gilt. Diese Einschränkung wurde in Quiggin (1982) allgemein als Eigenschaft einer Verzerrungsfunktion gefordert. Seitdem wurde diese Eigenschaft viel diskutiert, siehe dazu Segal (1987), Quiggin (1987).

In EU kann ein ET mit einer multiplikativen Nutzenfunktion lediglich drei verschiedene Arten von bivariater Risikohaltung zeigen. Ist $\beta < 0$, so gilt $I(BRA)I^c$ für jede nichttriviale Teilmenge $I \subset N$. Entsprechend gilt $I(BRS)I^c$ für jede nichttriviale Teilmenge $I \subset N$, wenn $\beta > 0$ ist und $I(BRN)I^c$ für jede nichttriviale Teilmenge $I \subset N$, wenn $\beta = 0$ ist. Die obigen Sätze zeigen, daß in AU eine größere Zahl von multivariaten Risikohaltungen möglich ist. Wir unterscheiden vier Fälle.

(i) Ist $\beta \leq 0$ und $c \geq 1$, so gilt $I(BRA)I^c$ für *jede* nichttriviale Teilmenge $I \subset N$.

(ii) Ist $\beta > 0$ und $c > 1$, so kann $I(BRA)I^c$ für gewisse Teilmengen $I \subset N$ erfüllt sein. Für die übrigen Teilmengen J gilt aber weder $J(BRA)J^c$, noch $J(BRS)J^c$, noch $J(BRN)J^c$.

(iii) Ist $\beta \geq 0$ und $c \leq 1$, so gilt $I(BRS)I^c$ für *jede* nichttriviale Teilmenge $I \subset N$.

(iv) Ist $\beta < 0$ und $c < 1$, so kann $I(BRS)I^c$ für gewisse Teilmengen $I \subset N$ erfüllt sein. Für die übrigen Teilmengen J gilt aber weder $J(BRA)J^c$, noch $J(BRS)J^c$, noch $J(BRN)J^c$.

Zur Illustration der vorhergehenden Sätze geben wir noch zwei Beispiele. Im ersten Beispiel präsentieren wir ein AU-Modell mit bivariat risikoaversem ET und multiplikativer Nutzenfunktion mit $\beta > 0$. Im zweiten Beispiel ist die Nutzenfunktion multiplikativ. Jedoch gilt weder bivariate Risikoaversion, noch Risikofreudigkeit, noch Risikoneutralität.

Beispiel 6.1 (Bivariate Risikoaversion mit $\beta > 0$) *Seien $\mathcal{C} = [0,1]^2$, q eine Verzerrungsfunktion mit $q(0.5) = 0.25$, und*

$$u(x_1, x_2) = 0.25x_1 + 0.25x_2 + 0.5x_1x_2 \,.$$

Dann ist $\alpha_1 = \alpha_2 = 0.25$ und $\beta = 8 > 0$. Man sieht jedoch leicht, daß der ET bivariat risikoavers ist.

Beispiel 6.2 (u multiplikativ, keine globale bivariate Risikohaltung) *Seien $\mathcal{C}$ und u wie in Beispiel 6.1, und sei q eine Verzerrungsfunktion mit $q(0.5) =$ 0.35. Dann hat der ET die folgenden Präferenzen.*

$$L_1 = \begin{bmatrix} (0,0) & (0.5,0.5) \\ 0.5 & 0.5 \end{bmatrix} \succ \begin{bmatrix} (0.5,0) & (0,0.5) \\ 0.5 & 0.5 \end{bmatrix} = L_2$$

and

$$L_3 = \begin{bmatrix} (0.5,0.5) & (1,1) \\ 0.5 & 0.5 \end{bmatrix} \prec \begin{bmatrix} (0.5,1) & (1,0.5) \\ 0.5 & 0.5 \end{bmatrix} = L_4$$

Wegen $L_1 \succ L_2$ ist der ET nicht bivariat risikoavers. Wegen $L_3 \prec L_4$ ist der ET nicht bivariat risikofreudig.

6.4 Additive Nutzenfunktionen in AU

In diesem Abschnitt untersuchen wir, unter welchen Bedingungen eine Nutzenfunktion in AU additiv ist. Eine Nutzenfunktion u heißt *additiv*, wenn sie darstellbar ist als

$$u(x_1, \ldots, x_n) = \sum_{i=1}^{n} \alpha_i u_i(x_i) \;, \tag{6.5}$$

wobei die Nutzenfunktionen u_i, $i = 1, \ldots n$, so normiert sind, daß $u_i(x_i^-) = 0$, $u_i(x_i^+) = 1$ für jedes $i = 1, \ldots, n$ erfüllt ist und $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, n$.

Wie wir bereits in Abschnitt 6.2 dargelegt haben, ist eine additive Nutzenfunktion ein Spezialfall einer multiplikativen Nutzenfunktion.

In EU gibt es verschiedene Bedingungen an die Präferenzen des ET, die zu einer additiven Nutzenfunktion des ET gleichwertig sind. Ein wichtiges Ergebnis besagt, daß die Nutzenfunktion additiv ist, wenn die Attribute die Randverteilungsbedingung erfüllen (vgl. Keeney/Raiffa 1976, Fishburn 1982). In AU erweist sich diese Bedingung als besonders stark. Wir werden nämlich zeigen, daß die Randverteilungsbedingung, auf ein AU-Modell angewendet, impliziert, daß die Nutzenfunktion additiv *und* die Verzerrungsfunktion die identische Abbildung ist. Das heißt aber nichts anderes, als daß aus dem AU-Modell ein einfaches EU-Modell wird. Der folgende Satz besagt somit: Die Randverteilungsbedingung impliziert EU.

Satz 6.5 *In einem AU-Modell sei der Konsequenzenraum $\mathcal{C}$ ein kartesisches Produkt der Attributräume $\mathcal{C}_1$ and $\mathcal{C}_2$. $\mathcal{C}_1$ und $\mathcal{C}_2$ erfüllen genau dann die Randverteilungsbedingung, wenn die folgenden beiden Bedingungen erfüllt sind:*

(i) $q = id|_{[0,1]}$.

(ii) u *ist additiv.*

Bemerkung: Satz 6.5 besagt umgekehrt, daß in einem AU-Modell mit nichttrivialer Verzerrungsfunktion (d.h. es ist $q(p) \neq p$ für mindestens ein $p \in [0,1]$) die Präferenzen des ET über multiattributive Lotterien immer auch von der Abhängigkeitsstruktur der Lotterien abhängen. In einem AU-Modell mit nichttrivialer Verzerrung beeinflussen also selbst bei additiver Nutzenfunktion die Abhängigkeiten der Verteilung die Präferenzen des ET.

In Abschnitt 6.3 haben wir unter der Voraussetung einer multiplikativen Nutzenfunktion gezeigt, daß bivariate Risikoneutralität eine additive Nutzenfunktion und $q(0.5) = (0.5)$ impliziert. Der folgende Satz zeigt, daß diese Aussage auch ohne die Voraussetzung einer multiplikativen Nutzenfunktion gültig bleibt.

Satz 6.6 *Der ET ist genau dann bivariat risikoneutral, wenn die folgenden beiden Bedingungen erfüllt sind.*

(i) $q(0.5) = 0.5$

(ii) u *ist additiv.*

Satz 6.5 und Satz 6.6 können leicht auf den Fall von mehr als zwei Attributen verallgemeinert werden. Ist z.B. die Randverteilungsbedingung bezüglich einer Teilmenge I von Attributen und deren Komplement I^c erfüllt, so kann man Satz 6.5 auf den „bivariaten" Konsequenzenraum $\mathcal{C} = \mathcal{C}_I \times \mathcal{C}_{-I}$ anwenden und daraus schließen, daß die Verzerrungsfunktion die Identität ist. Somit liegt wieder ein einfaches EU-Modell vor, und die bekannten Resultate über Zerlegungen von Nutzenfunktionen in EU können angewendet werden.

6.5 Korrelationsaversion in AU

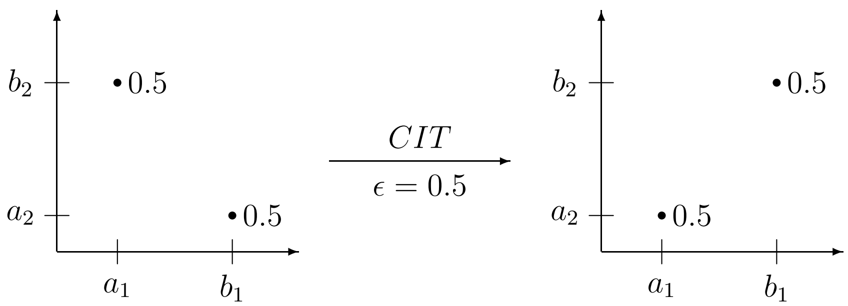
Der Begriff der bivariaten Risikohaltung kann auch wie folgt interpretiert werden. Ein bivariat risikoaverser ET präferiert die negativ korrelierte 50-50-Lotterie in Definition 6.4 gegenüber der positiv korrelierten 50-50-Lotterie. Verglichen werden dabei jedoch lediglich 50-50-Lotterien. Fordert man allgemein, daß der ET von zwei Lotterien, die sich lediglich darin unterscheiden, daß eine der Lotterien stärker korreliert ist als die andere, immer die weniger korrelierte bevorzugt, so gelangt man zum Begriff der *Korrelationsaversion.* Analog lassen sich auch *Korrelationsfreudigkeit* und *Korrelationsneutralität* definieren. Wir wollen in diesem Abschnitt untersuchen, unter welchen Bedingungen ein AU-ET korrelationsavers ist. Dazu muß allerdings zunächst definiert werden, was unter der Aussage „Die Zufallsvariable X ist korrelierter als Y“ zu verstehen ist. Diese Frage wurde von Epstein und Tanny (1980) diskutiert. Sie definieren eine *korrelationserhöhende Transformation (CIT)* auf die folgende Weise:

Definition 6.5 (Korrelationserhöhende Transformation) *Seien L und M Lotterien auf $\mathbb{R}^2$ mit endlichem Träger. L unterscheidet sich von M durch eine* korrelationserhöhende Transformation *(CIT), wenn es $a, b \in \mathbb{R}^2$ mit $a_i < b_i$, $i = 1, 2$, und $\epsilon > 0$ gibt, so daß*

$$M(\{x\}) - L(\{x\}) = \begin{cases} -\epsilon, & \textit{falls } x = (a_1, a_2) \textit{ oder } x = (b_1, b_2), \\ \epsilon, & \textit{falls } x = (a_1, b_2) \textit{ oder } x = (b_1, a_2), \\ 0, & \textit{sonst.} \end{cases}$$

Wir schreiben dann auch kürzer $M \overset{CIT}{\longrightarrow} L$.

Die folgende Abbildung verdeutlicht die Wirkung einer korrelationserhöhenden Transformation für $\epsilon = 0.5$:



Definition 6.6 (Stärkere Korrelation) *Seien P und Q W-Maße auf $(\mathbb{R}^2, \mathcal{B}^2)$, F und G die zugehörigen Verteilungsfunktionen und $F^{(i)}, G^{(i)}$, $i = 1, 2$, die Verteilungsfunktionen der Randverteilungen. Wir definieren eine Relation $\leq_D$ wie folgt. $P \leq_D Q$ genau dann, wenn $F(x) \leq G(x)$ für alle $x \in \mathbb{R}^2$ und $F^{(i)} = G^{(i)}$, $i = 1, 2$. Gilt $P \leq_D Q$, so sagen wir auch „Q ist stärker korreliert als P“.*

Man zeigt leicht, daß $\leq_D$ eine partielle Ordnung ist. Daß die Interpretation „Q ist stärker korreliert als P“ für $P \leq_D Q$ sinnvoll ist, zeigen die folgenden Sätze, die in Epstein and Tanny (1980) bewiesen werden.

Satz 6.7 (Epstein und Tanny 1980) *Seien L und M Lotterien auf $\mathbb{R}^2$ mit endlichem Träger. $L \leq_D M$ gilt genau dann, wenn eine Folge von Lotterien mit endlichem Träger $L = L_0, L_1, \ldots, L_n = M$, existiert, so daß $L_{i-1} \xrightarrow{CIT} L_i$ für alle $i = 1, \ldots, n$ gilt.*

Satz 6.8 (Epstein und Tanny 1980) *Seien P und Q W-Maße auf $(\mathbb{R}^2, \mathcal{B}^2)$. Gilt $P \leq_D Q$, so existieren Folgen $(P_n)_{n\in\mathbb{N}}$ und $(Q_n)_{n\in\mathbb{N}}$ von Lotterien mit endlichem Träger, so daß P_n schwach[2] gegen P und Q_n schwach gegen Q konvergiert und $P_n \leq_D Q_n$ für alle $n \in \mathbb{N}$ gilt.*

Wir kommen nun zu der schon angekündigten Definition der Korrelationsaversion.

Definition 6.7 (Korrelationsaversion) *In einem Entscheidungsproblem unter Risiko sei der Konsequenzenraum $(\mathcal{C}, \mathcal{D}) = (\mathbb{R}^2, \mathcal{B}^2)$. Ein ET heißt* korrelationsavers*, wenn für alle $P, Q \in \mathcal{P}$ gilt*

$$P \leq_D Q \Rightarrow P \succeq Q.$$

Er heißt korrelationsfreudig*, falls $P \leq_D Q \Rightarrow P \preceq Q$ gilt, und* korrelationsneutral*, falls $P \leq_D Q \Rightarrow P \sim Q$ gilt.*

Der folgende Satz liefert eine Charakterisierung von Korrelationsaversion im AU-Modell. Dabei bezeichnen wir mit u_{x_1,x_2} die partielle Ableitung nach x_1 und x_2. Eine Charakterisierung von Korrelationsaversion im EU-Modell findet sich in Epstein und Tanny (1980).

Satz 6.9 *In einem AU-Modell mit Konsequenzenraum $(\mathbb{R}^2, \mathcal{B}^2)$ seien Nutzenfunktion u und Verzerrungsfunktion q zweimal differenzierbar. Ferner sei die Nutzenfunktion u monoton wachsend. Ist der ET korrelationsavers, so ist $u_{x_1,x_2} \leq 0$ und $q'' \geq 0$. Ist u beschränkt, so gilt hiervon auch die Umkehrung.*

Den Beweis von Satz 6.9 findet man im folgenden Abschnitt.

[2] Eine Folge $(P_n)_{n\in\mathbb{N}}$ von W-Maßen auf $(\mathbb{R}^k, \mathcal{B}^k)$ heißt schwach konvergent gegen ein W-Maß P, wenn für die zugehörigen Verteilungsfunktionen $F_n, n \in \mathbb{N}$, und F gilt: Für alle Stetigkeitsstellen x von F ist $\lim_{n\to\infty} F_n(x) = F(x)$.

6.6 Beweise

Bevor wir mit den Beweisen beginnen können, benötigen wir noch ein Lemma.

Lemma 6.1 *Es liege ein nichtdegeneriertes AU- oder CEU-Modell vor. Die Menge I von Attributen sei nutzenunabhängig. Dann gibt es eine Funktion $\gamma : \mathcal{C}_{-I} \to \mathbb{R}_+ \setminus \{0\}$, so daß für jedes $x \in \mathcal{C}$*

$$u(x) = \gamma(x_{-I})u(x_I) + u(x_{-I})$$

erfüllt ist.

Beweis: Wir führen den Beweis für ein AU-Modell. Zum Beweis der Aussage für ein CEU-Modell hat man im folgenden lediglich „Lotterien“ durch „Aktionen“ zu ersetzen.

Da I nutzenunabhängig ist, hängen Präferenzen über Lotterien mit gemeinsamer und sicherer Konsequenz $x_{-I} \in \mathcal{C}_{-I}$ nicht von x_{-I} ab. Dies bedeutet, daß für jedes $x_{-I} \in \mathcal{C}_{-I}$ die Funktion $u(\cdot, x_{-I}) : \mathcal{C}_I \to \mathbb{R}$ eine Nutzenfunktion für Lotterien über $\mathcal{C}_I$ ist und daß alle diese Nutzenfunktionen dieselbe Präferenz über Lotterien darstellen. Da unter der Voraussetung der Nichtdegeneriertheit des Modells die Nutzenfunktion eindeutig bis auf positive affine Transformationen ist, existieren Funktionen $\gamma : \mathcal{C}_{-I} \to \mathbb{R}_+ \setminus \{0\}$ und $\beta : \mathcal{C}_{-I} \to \mathbb{R}$, so daß

$$u(x) = \gamma(x_{-I})u(x_I, x_{-I}^-) + \beta(x_{-I}) \quad \forall x \in \mathcal{C}\,.$$

Setzt man in diese Formel $x = (x_I^-, x_{-I})$ ein, so ergibt sich $\beta(x_{-I}) = u(x_I^-, x_{-I})$ und somit

$$u(x) = \gamma(x_{-I})u(x_I, x_{-I}^-) + u(x_I^-, x_{-I}) \quad \forall x \in \mathcal{C}\,,$$

wie behauptet. ∎

Beweis von Satz 6.1

Zum Beweis von (ii) beachten wir zunächst, daß für jede nichtleere Teilmenge $I \subset N$ von Attributen

$$\alpha_I = u(x_I^+) = \sum_{\emptyset \neq J \subset I} \beta^{|J|-1} \prod_{j \in J} \alpha_j \tag{6.6}$$

und

$$u_I(x_I) = \frac{u(x_I)}{u(x_I^+)} = \alpha_I^{-1} \sum_{\emptyset \neq J \subset I} \beta^{|J|-1} \prod_{j \in J} \alpha_j u_j(x_j)$$

gilt. Damit erhält man

$$\begin{aligned} u(x) &= \alpha_I u_I(x_I) + \alpha_{-I} u_{-I}(x_{-I}) + \beta \alpha_I u_I(x_I) \alpha_{-I} u_{-I}(x_{-I}) \\ &= \alpha_I u_I(x_I) \left[1 + \beta \alpha_{-I} u_{-I}(x_{-I})\right] + \alpha_{-I} u_{-I}(x_{-I}) \end{aligned}$$

für alle $x \in \mathcal{C}$. Dies ist aber äquivalent zu

$$[1 + \beta \alpha_{-I} u_{-I}(x_{-I})] = \frac{u(x_I, x_{-I}) - u(x_I^-, x_{-I})}{u(x_I, x_{-I}^-) - u(x_I^-, x_{-I}^-)}. \quad \text{für alle } x_I \in \mathcal{C}_I,\ x_{-I} \in \mathcal{C}_{-I}\,.$$

Da I präferenzunabhängig ist, haben Zähler und Nenner das gleiche Vorzeichen. Daher ist der Ausdruck in eckigen Klammern größer als 0. Somit sind alle Funktionen $u(\cdot, x_{-I}) : \mathcal{C}_I \to \mathbb{R}$, $x_{-I} \in \mathcal{C}_{-I}$, positive affine Transformationen von $u(\cdot, x_{-I}^-)$ und repräsentieren somit dieselbe Präferenzordnung auf den Lotterien (bzw. Aktionen) mit gemeinsamen sicherem Niveau $x_{-I} \in \mathcal{C}_{-I}$. Dies bedeutet aber, daß I nutzenunabhängig ist. Da I beliebig gewählt war, ist (ii) gezeigt.

Zum Beweis von (i) ist es hinreichend anzunehmen, daß für jedes $i \in N$ die Menge $-\{i\}$ von Attributen nutzenunabhängig ist. Nach Lemma 6.1 erhalten wir für jedes $i \in N$ die folgende Gleichung:

$$u(x) = \gamma_i(x_i) u(x_{-i}) + u(x_i)\ . \tag{6.7}$$

Für $i \neq j$ ergibt sich

$$\gamma_i(x_i) u(x_{-i}) + u(x_i) = \gamma_j(x_j) u(x_{-j}) + u(x_j)\ .$$

Setzt man $x = (x_{\{i,j\}}, x^-_{-\{i,j\}}$, so erhält man

$$\gamma_i(x_i) u(x_j) + u(x_i) = \gamma_j(x_j) u(x_i) + u(x_j)$$

und somit

$$\frac{\gamma_j(x_j) - 1}{u(x_j)} = \frac{\gamma_i(x_i) - 1}{u(x_i)}\ .$$

Da die linke Seite eine Funktion von x_j allein ist und die rechte Seite nur von x_i abhängt, müssen beide Quotienten konstant sein. Wir bezeichnen diesen konstanten Wert mit β. Dann ist für jedes $i \in N$

$$\gamma_i(x_i) = 1 + \beta u(x_i)\ .$$

Wir bemerken für später, daß insbesondere $1 + \beta u(x_i) > 0$ für alle $x_i \in \mathcal{C}_i$, $i \in N$, gilt.

Ersetzt man in Gleichung (6.7) $\gamma_i(x_i)$ durch $1 + \beta u(x_i)$, erhält man

$$u(x) = u(x_i) + [1 + \beta u(x_i)] u(x_{-\{i\}})\ .$$

Einsetzen von $x = (x^-_{\{1,\dots,i-1\}}, x_{\{i,\dots,n\}})$ führt zu

$$u(x_{\{i,\dots,n\}}) = u(x_i) + [1 + \beta u(x_i)]u(x_{\{i+1,\dots,n\}})\,. \tag{6.8}$$

Für $\beta \neq 0$ ist dies zu

$$1 + \beta u(x_{\{i,\dots,n\}}) = [1 + \beta u(x_i)][1 + \beta u(x_{\{i+1,\dots,n\}})]\ .$$

äquivalent. Mit vollständiger Induktion folgt, daß

$$\begin{aligned} 1 + \beta u(x) &= \prod_{i=1}^{n} [1 + \beta u(x_i)] \\ &= \sum_{I \subset N} \beta^{|I|} \prod_{i \in I} u(x_i) \end{aligned}$$

Hieraus folgt nun aber sofort

$$u(x) = \sum_{\emptyset \neq I \subset N} \beta^{|I|-1} \prod_{i \in I} u(x_i)$$

und da $u(x_i) = \alpha_i u_i(x_i)$ erhält man schließlich das gewünschte Ergebnis.
Ist $\beta = 0$, so ist Gleichung (6.8) äquivalent zu

$$u(x_{\{i,\dots,n\}}) = u(x_i) + u(x_{\{i+1,\dots,n\}})\,.$$

Wieder mit vollständiger Induktion erhält man

$$u(x) = \sum_{i=1}^{n} u(x_i) = \sum_{i=1}^{n} \alpha_i u_i(x_i)\ .$$

Damit ist nun auch (i) vollständig bewiesen. ∎

Beweis von Satz 6.2

Wir zeigen zuerst (ii). Ist u multilinear, so kann für jedes $i \in N$ die Nutzenfunktion u als

$$\begin{aligned} u(x) &= u_i(x_i)\Big(\gamma_i + \sum_{\emptyset \neq I \subset N \setminus \{i\}} \gamma_{I \cup \{i\}} \prod_{j \in I} u_j(x_j)\Big) + \sum_{\emptyset \neq I \subset N \setminus \{i\}} \gamma_I \prod_{j \in I} u_j(x_j) \\ &= u_i(x_i)\Big(\gamma_i + \sum_{\emptyset \neq I \subset N \setminus \{i\}} \gamma_{I \cup \{i\}} \prod_{j \in I} u_j(x_j)\Big) + u(x_{-i}) \end{aligned}$$

geschrieben werden. Wie im Beweis von Theorem 6.1 zeigt man, daß der Ausdruck in eckigen Klammern größer als 0 ist. Dieselben Argumente wie dort zeigen dann, daß (ii) gilt.

Zum Beweis von (i) führen wir die folgende Bezeichnung ein. Für jede Teilmenge $I \subset N$ sei die Differenz $\Delta_I u(x)$ rekursiv definiert durch $\Delta_\emptyset u(x) = u(x)$ und $\Delta_{I\cup\{i\}} u(x) = \Delta_I u(x_i^+, x_{-i}) - \Delta_I u(x_i^-, x_{-i})$, falls $i \notin I$. Wir benutzen wieder Lemma 6.1. Da i nutzenunabhängig ist erhält man für jedes $i \in N$

$$u(x) = \gamma_{-i}(x_{-i})u(x_i) + u(x_{-i})$$

Einsetzen von $x = (x_i^+, x_{-i})$ und auflösen nach $\gamma_{-i}(x_{-i})$ liefert

$$\begin{aligned} \gamma_{-i}(x_{-i}) &= \frac{u(x_i^+, x_{-i}) - u(x_i^-, x_{-i})}{u(x_i^+)} \\ &= u(x_i^+)^{-1}\Delta_{\{i\}}u(x_{-i}). \end{aligned}$$

Somit hat man für jedes $i \in N$

$$u(x) = \frac{u(x_i)}{u(x_i^+)}\Delta_i u(x_{-i}) + \Delta_\emptyset u(x_{-i})$$

und da $u_i(x_i) = [u(x_i^+)]^{-1}u(x_i)$ erhalten wir

$$u(x) = u_i(x_i)\Delta_i u(x_{-i}) + \Delta_\emptyset u(x_{-i}) \ . \tag{6.9}$$

Setzt man in Gleichung (6.9) $x = (x^-_{\{1,\ldots,m-1\}}, x_{\{m,\ldots,n\}})$ und $i = m$ ein, so folgt zunächst

$$u(x_{\{m,\ldots,n\}}) = u_m(x_m)\Delta_m u(x_{\{m+1,\ldots,n\}}) + \Delta_\emptyset u(x_{\{m+1,\ldots,n\}})$$

und damit

$$\begin{aligned} \Delta_I u(x_{\{m,\ldots,n\}}) &= \Delta_I\Big(u_m(x_m)\Delta_m u(x_{\{m+1,\ldots,n\}}) + \Delta_\emptyset u(x_{\{m+1,\ldots,n\}})\Big) \\ &= u_m(x_m)\Delta_{I\cup\{m\}}u(x_{\{m+1,\ldots,n\}}) + \Delta_I u(x_{\{m+1,\ldots,n\}}). \end{aligned} \tag{6.10}$$

Wir zeigen mit Induktion, daß

$$u(x) = \sum_{I\subset\{1,\ldots m\}} \Delta_I u(x_{\{m+1,\ldots n\}}) \prod_{i\in I} u_i(x_i) \tag{6.11}$$

für $m = 1, \ldots n$ gilt. Setzt man in Gleichung (6.9) $i = 1$ liefert dies gerade die Behauptung für $m = 1$. Sei nun die Behauptung für $m-1$ bewiesen. Dann erhält man mit (6.10)

$$\begin{aligned} u(x) &= \sum_{I\subset\{1,\ldots m-1\}} \Delta_I u(x_{\{m,\ldots n\}}) \prod_{i\in I} u_i(x_i) \\ &= \sum_{I\subset\{1,\ldots m-1\}} \Big(\Delta_{I\cup\{m\}}u(x_{\{m+1,\ldots n\}})u_m(x_m) + \Delta_I u(x_{\{m+1,\ldots n\}}\Big) \prod_{i\in I} u_i(x_i) \\ &= \sum_{I\subset\{1,\ldots m\}} \Delta_I u(x_{\{m+1,\ldots n\}}) \prod_{i\in I} u_i(x_i). \end{aligned}$$

Also gilt (6.11) für alle $m = 1, \dots n$. Für $m = n$ erhält man aber

$$u(x) = \sum_{I \subset N} \Delta_I u(x^-) \prod_{i \in I} u_i(x_i) \, .$$

Da $\Delta_\emptyset u(x^-) = 0$ ist, vereinfacht sich obige Gleichung zu

$$u(x) = \sum_{\emptyset \neq I \subset N} \Delta_I u(x^-) \prod_{i \in I} u_i(x_i).$$

Mit $\gamma_I = \Delta_I u(x^-)$ ergibt sich dann die Behauptung. ■

Beweis von Satz 6.3

Wir beginnen mit dem Beweis von Teil (i). Sei $x_1 \prec_1 y_1$. Wir wählen dann $x_2, y_2 \in \mathcal{C}_2$ so, daß $x_2 \prec_2 y_2$ und $(x_1, y_2) \prec (y_1, x_2)$ gelten. Dies ist möglich, da die Nutzenfunktion u stetig und die Attributräume $\mathcal{C}_1$ und $\mathcal{C}_2$ zusammenhängend sind. Da der ET bivariat risikoavers ist, wird die Lotterie

$$L_1 = \begin{bmatrix} (x_1, y_2) & (y_1, x_2) \\ 0.5 & 0.5 \end{bmatrix}$$

gegenüber der Lotterie

$$L_2 = \begin{bmatrix} (x_1, x_2) & (y_1, y_2) \\ 0.5 & 0.5 \end{bmatrix}$$

präferiert. Berechnung des antizipierten Nutzens zeigt, daß dies gleichwertig ist zu

$$[1 - q(0.5)]u(x_1, y_2) + q(0.5)u(y_1, x_2) \geq [1 - q(0.5)]u(x_1, x_2) + q(0.5)u(y_1, y_2).$$

Division der Ungleichung durch $q(0.5)$, ersetzen von $(1 - q(0.5))/q(0.5)$ durch c und Umformung der Terme ergibt

$$c[u(x_1, y_2) - u(x_1, x_2)] \geq u(y_1, y_2) - u(y_1, x_2) \; . \tag{6.12}$$

Da $(x_1, y_2) \succ (x_1, x_2)$ ist, ist die Differenz $u(x_1, y_2) - u(x_1, x_2)$ größer als 0. Ungleichung (6.12) ist damit äquivalent zu

$$c \geq \frac{u(y_1, y_2) - u(y_1, x_2)}{u(x_1, y_2) - u(x_1, x_2)} \; . \tag{6.13}$$

Da die Nutzenfunktion u nach Voraussetzung multiplikativ ist, können wir (6.13) auch als

$$c \geq \frac{\alpha_2[u_2(y_2) - u_2(x_2)] + \beta\alpha_1 u_1(y_1)\alpha_2[u_2(y_2) - u_2(x_2)]}{\alpha_2[u_2(y_2) - u_2(x_2)] + \beta\alpha_1 u_1(x_1)\alpha_2[u_2(y_2) - u_2(x_2)]} \tag{6.14}$$

schreiben. Kürzen des Quotienten ergibt dann

$$c \geq \frac{1+\beta\alpha_1 u_1(y_1)}{1+\beta\alpha_1 u_1(x_1)} \,. \tag{6.15}$$

Nun muß diese Ungleichung aber für alle $x_1, y_1 \in \mathcal{C}_1$ mit $x_1 \prec_1 y_1$ gültig sein. Dies bedeutet, daß

$$c \geq \sup\left\{\frac{1+\beta\alpha_1 u_1(y_1)}{1+\beta\alpha_1 u_1(x_1)} \;\Big|\; x_1, y_1 \in \mathcal{C}_1,\ x_1 \prec_1 y_1\right\} \,. \tag{6.16}$$

Ist $\beta \leq 0$, so ist die rechte Seite der Ungleichung (6.16) gleich 1. Dies folgt aus der Tatsache, daß u stetig und die Attributräume zusammenhängend sind. Daher existiert eine Folge $(y_1^n)_{n\in\mathbb{N}}$ in $\mathcal{C}_1$, so daß $u_1(y_1^n)$ gegen $u_1(x_1)$ konvergiert.

Ist $\beta > 0$, so ist das Supremum durch $1+\beta\alpha_1$ gegeben. Dies sieht man durch Einsetzen von $x_1 = x_1^-$ und $y_1 = x_1^+$. Aus (6.16) folgt somit

$$\beta \leq 0,\ c \geq 1 \quad \text{oder} \quad \beta > 0,\ c \geq 1+\beta\alpha_1 \,. \tag{6.17}$$

Vertauscht man die Rollen von $\mathcal{C}_1$ und $\mathcal{C}_2$ so führt dieselbe Argumentation zu

$$\beta \leq 0,\ c \geq 1 \quad \text{oder} \quad \beta > 0,\ c \geq 1+\beta\alpha_2 \,. \tag{6.18}$$

Kombiniert man (6.17) und (6.18), so erhält man schließlich

$$\beta \leq 0,\ c \geq 1 \quad \text{oder} \quad \beta > 0,\ c \geq 1+\beta\max\{\alpha_1,\alpha_2\} \,.$$

Damit haben wir nun Teil (i) gezeigt.

Der Beweis von Teil (ii) verläuft völlig analog. Zum Beweis von Teil (iii) bemerken wir, daß ein ET bivariat risikoneutral ist, wenn er bivariat risikoavers *und* bivariat risikofreudig ist. Teil (iii) folgt dann direkt aus (i) und (ii). ∎

Beweis von Satz 6.4

Wir zeigen nur Teil (i). Teil (ii) folgt wieder analog zu (i), und Teil (iii) ist eine einfache Folgerung aus (i) und (ii).

Da

$$\alpha_I = \sum_{\emptyset\neq J\subset I} \beta^{|J|-1} \prod_{j\in J} \alpha_j \tag{6.19}$$

und

$$u_I(x_I) = \alpha_I^{-1} \sum_{\emptyset\neq J\subset I} \beta^{|J|-1} \prod_{j\in J} \alpha_j u_j(x_j) \,,$$

sieht man, daß

$$u(x_I, x_{-I}) = \alpha_I u_I(x_I) + \alpha_{-I} u_{-I}(x_{-I}) + \beta \alpha_I u_I(x_I) \alpha_{-I} u_{-I}(x_{-I}) \ .$$

Wegen Satz 6.3 ist der ET genau dann bivariat risikavers bezüglich $\mathcal{C}_I$ und $\mathcal{C}_{-I}$, wenn

$$\beta \leq 0, \qquad c \geq 1 \qquad \text{oder} \tag{6.20}$$

$$\beta > 0, \qquad c \geq 1 + \beta \max\{\alpha_I, \alpha_{-I}\} \tag{6.21}$$

erfüllt ist. Da entweder $\beta \leq 0$ oder $\beta > 0$ gilt, schließen wir, daß entweder $c \geq 1$ gelten muß oder daß $c \geq 1 + \beta\alpha_I$ für alle nichttrivialen Teilmengen $I \subset N$ gelten muß. Es bleibt somit nur noch zu zeigen, daß $c \geq 1 + \beta\alpha_I$ für alle nichttrivialen Teilmengen $I \subset N$ zu der im Satz aufgestellten Bedingung äquivalent ist. Mit (6.19) ergibt sich, daß $c \geq 1 + \beta\alpha_I$ äquivalent zu

$$c \geq \prod_{i \in I} (1 + \beta\alpha_i) \tag{6.22}$$

ist. Ist also $\beta > 0$, so muß Bedingung (6.22) für alle nichttrivialen Teilmengen I von N gelten. Wegen $\beta > 0$ ist der Ausdruck in Klammern größer 1 und somit

$$c \geq \max_{\emptyset \neq I \subset N,\ I \neq N} \prod_{i \in I} (1 + \beta\alpha_i) = \max_{j \in N} \prod_{\substack{i \in N \\ i \neq j}} (1 + \beta\alpha_i) \,. \tag{6.23}$$

Wegen

$$\prod_{i \in N} (1 + \beta\alpha_i) = 1 + \beta$$

kann man dann (6.23) auch als

$$c \geq \max_{j \in N} \frac{1 + \beta}{1 + \beta\alpha_j}$$

schreiben. ∎

Beweis von Satz 6.5

Der Beweis, daß aus der Randverteilungsbedingung eine additive Nutzenfunktion und $q = id|_{[0,1]}$ folgt, besteht aus drei Teilen. Im ersten Teil zeigen wir, daß die Randverteilungsbedingung Präferenzunabhängigkeit der Attribute impliziert. Dann zeigen wir, daß q die Identität ist. Im dritten Schritt schließlich wird gezeigt, daß u additiv ist.

Schritt 1: Die Attribute sind präferenzunabhängig.
Wir nehmen an, daß Attribut 1 nicht präferenzunabhängig ist. Dann gibt es

$x_i, y_i \in \mathcal{C}_i$, $i = 1, 2$, so daß $u(x_1, x_2) \leq u(y_1, x_2)$ ist aber $u(x_1, y_2) > u(y_1, y_2)$. Die Lotterien

$$\begin{bmatrix} (x_1, x_2) & (y_1, y_2) \\ 0.5 & 0.5 \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} (y_1, x_2) & (x_1, y_2) \\ 0.5 & 0.5 \end{bmatrix}$$

haben dieselben Randverteilungen. Wegen der Randverteilungsbedingung ist der ET also indifferent zwischen beiden Lotterien. Berechnet man aber den antizipierten Nutzen der beiden Lotterien, so sieht man, daß der Nutzen der rechten Lotterie echt größer ist, als der der linken Lotterie. Die Annahme, daß Attribut i nicht präferenzunabhängig ist, führt also zu einem Widerspruch. Attribut 1 muß also präferenzunabhängig sein. Entsprechend sieht man, daß auch Attribut 2 präferenzunabhängig ist.

Schritt 2: q ist die identische Abbildung.
Dazu wählen wir $x_i, y_i \in \mathcal{C}_i$ so, daß $x_i \prec_i y_i$, $i = 1, 2$. Wir betrachten nun Lotterien der Form:

$$L = \begin{bmatrix} (x_1, x_2) & (x_1, y_2) & (y_1, x_2) & (y_1, y_2) \\ 1 - p_0 - p & p & p & p_0 - p \end{bmatrix},$$

wobei $p_0 \in [0, 1]$ und $0 \leq p \leq \min\{p_0, 1 - p_0\}$. Dann sind die Randverteilungen L_1 und L_2 gegeben durch

$$L_1 = \begin{bmatrix} x_1 & y_1 \\ 1 - p_0 & p_0 \end{bmatrix} \quad \text{und} \quad L_2 = \begin{bmatrix} x_2 & y_2 \\ 1 - p_0 & p_0 \end{bmatrix}.$$

Somit haben für festes p_0 alle diese Lotterien dieselben Randverteilungen. Wir nehmen o.B.d.A. an, daß $(x_1, y_2) \preceq (y_1, x_2)$ ist. Da die Attribute präferenzunabhängig sind, gilt

$$(x_1, x_2) \prec (x_1, y_2) \preceq (y_1, x_2) \prec (y_1, y_2) \ .$$

Wir berechnen den antizipierten Nutzen der Lotterien L zu

$$\begin{aligned} \int_{\mathcal{C}} u \; d(q \circ L) \;=\; u(x_1, x_2) \;&+\; [u(x_1, y_2) - u(x_1, x_2)] \; q(p_0 + p) \\ &+\; [u(y_1, x_2) - u(x_1, y_2)] \; q(p_0) \\ &+\; [u(y_1, y_2) - u(y_1, x_2)] \; q(p_0 - p) \ . \end{aligned}$$

Da die Randverteilungsbedingung erfüllt ist, ist obiger Ausdruck für festes p_0 und $0 \leq p \leq \min\{p_0, 1 - p_0\}$ konstant in p, Dann gilt aber für alle $p \in [0, \min\{p_0, 1 - p_0\}]$

$$[u(x_1, y_2) - u(x_1, x_2)] \; q(p_0 + p) + [u(y_1, y_2) - u(y_1, x_2)] \; q(p_0 - p) = c(p_0) \ . \quad (6.24)$$

Setzt man

$$d = \frac{u(y_1, y_2) - u(y_1, x_2)}{u(x_1, y_2) - u(x_1, x_2)}\,, \quad (6.25)$$

so schreibt sich diese Gleichung als

$$\Big(u(x_1, y_2) - u(x_1, x_2)\Big)\Big(q(p_0 + p) + dq(p_0 - p)\Big) = c(p_0). \qquad (6.26)$$

Setzt man nun $p = 0$ ein, erhält man

$$\Big(u(x_1, y_2) - u(x_1, x_2)\Big)(1 + d)q(p_0) = c(p_0)\,.$$

Einsetzen in (6.26) ergibt dann

$$\Big(u(x_1, y_2)-u(x_1, x_2)\Big)\Big(q(p_0+p)+dq(p_0-p)\Big) = \Big(u(x_1, y_2)-u(x_1, x_2)\Big)(1+d)q(p_0)\,. \qquad (6.27)$$

Wegen $(x_1, y_2) \succ (x_1, x_2)$ ist die Differenz $u(x_1, y_2) - u(x_1, x_2)$ nicht Null, und wir können (6.27) durch $u(x_1, y_2) - u(x_1, x_2)$ teilen. Somit erhalten wir für alle $p \in [0, \min\{p_0, 1 - p_0\}]$ die folgende Funktionalgleichung

$$q(p_0 + p) + dq(p_0 - p) = (1 + d)q(p_0). \qquad (6.28)$$

Wir zeigen nun, daß die Konstante d gleich 1 ist. Dazu setzen wir spezielle Werte von p_0 und p ein und berechnen so den Wert von q an bestimmten Punkten. Für $p_0 = \frac{1}{2}$ und $p = \frac{1}{2}$ erhält man

$$q(\tfrac{1}{2}) = \frac{1}{1 + d}\,.$$

Mit $p_0 = \frac{1}{4}$ und $p = \frac{1}{4}$ erhält man

$$q(\tfrac{1}{4}) = \frac{1}{(1 + d)^2}\,,$$

mit $p_0 = \frac{1}{2}$ und $p = \frac{1}{4}$

$$q(\tfrac{3}{4}) = 1 - \frac{d}{(1 + d)^2},$$

und mit $p_0 = \frac{3}{4}$ und $p = \frac{1}{4}$ schließlich

$$q(\tfrac{3}{4}) = \frac{1 + 2d}{(1 + d)^2}.$$

Durch Gleichsetzen der letzten Gleichungen ergibt sich, daß d nur die Werte $d = 0$ oder $d = 1$ annehmen kann. Da aber $(y_1, y_2) \succ (y_1, x_2)$ ist, folgt, daß $d \neq 0$ ist. Also muß $d = 1$ sein. Gleichung (6.28) vereinfacht sich damit zu

$$q(p_0+p)+q(p_0-p) = 2q(p_0) \qquad \text{falls } 0 \le p_0 \le 1,\ 0 \le p \le \min\{p_0, 1-p_0\}\,. \qquad (6.29)$$

Wir zeigen nun, daß diese Funktionalgleichung nur für $q(x) = x$ für alle $x \in [0, 1]$ erfüllt ist. Mit $p_0 = p$ vereinfacht sich (6.29) zu

$$q(2p) = 2q(p) \qquad \text{falls } 0 \le p \le 0.5\,.$$

Mit vollständiger Induktion folgt dann, daß

$$q(2^{-n}) = 2^{-n} \qquad \text{für alle } n \in \mathbb{N}. \tag{6.30}$$

Setzt man $p_0 = m \cdot 2^{-n}$ mit $1 \leq m \leq 2^n - 1$ und $p = 2^{-n}$, so erhält man aus (6.6) und (6.30)

$$q((m+1) \cdot 2^{-n}) = 2q(m \cdot 2^{-n}) - q((m-1) \cdot 2^{-n}).$$

Mit vollständiger Induktion folgt hieraus, daß

$$q(m \cdot 2^{-n}) = m \cdot 2^{-n} \qquad \text{für alle } m, n \in \mathbb{N},\ 0 \leq m \leq 2^n. \tag{6.31}$$

Gleichung (6.31) besagt nun aber, daß $q(x) = x$ für eine dichte Teilmenge von $[0,1]$ erfüllt ist. Da q monoton wachsend ist, muß dann aber $q(x) = x$ für alle $x \in [0,1]$ gelten.

Schritt 3: u ist additiv.
Da q die Identität ist, vereinfacht sich das AU-Funktional zu einem EU-Funktional. Damit können wir nun aber die bekannten Sätze über Zerlegungen von Nutzenfunktionen in EU anwenden. Nach Theorem 6.4 in Keeney/Raiffa (1976) folgt dann, daß u additiv ist. Also ist gezeigt, daß aus der Randverteilungsbedingung eine additive Nutzenfunktion und die identische Verzerrungsfunktion folgt.

Für die andere Richtung des Beweises haben wir nur noch zu zeigen, daß aus $q = id|_{[0,1]}$ und einer additiven Nutzenfunktion u folgt, daß die Randverteilungsbedingung erfüllt ist. Dies ist aber klar, da wegen $q = id|_{[0,1]}$ wieder ein einfaches EU-Modell vorliegt. Mit Theorem 6.4 in Keeney/Raiffa (1976) erhält man, daß $\mathcal{C}_1$ und $\mathcal{C}_2$ die Randverteilungsbedingung erfüllen. ∎

Beweis von Satz 6.6

Sei $x \in \mathcal{C}$. Da der ET bivariat risikoneutral ist, ist er indifferent zwischen den Lotterien

$$L_1 = \begin{bmatrix} (x_1^-, x_2^-) & (x_1, x_2) \\ 0.5 & 0.5 \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} (x_1^-, x_2) & (x_1, x_2^-) \\ 0.5 & 0.5 \end{bmatrix} = L_2 \ . \tag{6.32}$$

Um den antizipierten Nutzen der beiden Lotterien zu bestimmen, müssen wir zwei Fälle betrachten, nämlich $(x_1^-, x_2) \preceq (x_1, x_2^-)$ und $(x_1, x_2^-) \prec (x_1^-, x_2)$. Im ersten Fall ist Gleichung (6.32) äquivalent zu

$$u(x_1^-, x_2^-) + [u(x_1, x_2) - u(x_1^-, x_2^-)]\, q(0.5) = u(x_1^-, x_2) + [u(x_1, x_2^-) - u(x_1^-, x_2)]\, q(0.5)$$

und da $u(x^-) = 0$ ist, auch zu

$$u(x_1, x_2) \ = \ u(x_1, x_2^-) + u(x_1^-, x_2)\, \frac{1 - q(0.5)}{q(0.5)}.$$

Wir ersetzen $(1-q(0.5))/q(0.5)$ durch c. Da $u(x_1, x_2^-) = \alpha_1 u_1(x_1)$ und $u(x_1^-, x_2) = \alpha_2 u_2(x_2)$ ist, können wir obige Gleichung als

$$u(x_1, x_2) \;=\; \alpha_1 u_1(x_1) + c\alpha_2 u_2(x_2)$$

schreiben. Im Fall $(x_1, x_2^-) \prec (x_1^-, x_2)$ schließen wir analog auf

$$u(x_1, x_2) \;=\; c\alpha_1 u_1(x_1) + \alpha_2 u_2(x_2)\ .$$

Zusammenfassend erhalten wir also

$$u(x_1, x_2) \;=\; \begin{cases} \alpha_1 u_1(x_1) + c\alpha_2 u_2(x_2), & \text{falls } (x_1^-, x_2) \preceq (x_1, x_2^-), \\ c\alpha_1 u_1(x_1) + \alpha_2 u_2(x_2), & \text{falls } (x_1^-, x_2) \succ (x_1, x_2^-). \end{cases}$$

Wir müssen zeigen, daß $c = 1$ ist. Dazu wählen wir $y_i, z_i \in \mathcal{C}_i$ so, daß $x_i^- \prec_i y_i \prec_i z_i$, $i = 1, 2$, ist und nehmen o.B.d.A. an, daß $(x_1^-, y_2) \preceq (y_1, x_2^-)$ ist. Damit erhalten wir

$$u(y_1, y_2) = \alpha_1 u_1(y_1) + c\alpha_2 u_2(y_2)\ .$$

Da die Attribute präferenzunabhängig sind, ist $(x_1^-, y_2) \prec (z_1, x_2^-)$ und es folgt, daß

$$u(z_1, y_2) = \alpha_1 u_1(z_1) + c\alpha_2 u_2(y_2)\ .$$

Da der ET bivariat risikoneutral ist, gilt die folgende Indifferenz

$$\begin{bmatrix} (y_1, x_2^-) & (z_1, y_2) \\ 0.5 & 0.5 \end{bmatrix} \sim \begin{bmatrix} (y_1, y_2) & (z_1, x_2^-) \\ 0.5 & 0.5 \end{bmatrix} \tag{6.33}$$

Um den antizipierten Nutzen der rechten Lotterie zu bestimmen, unterscheiden wir wieder zwei Fälle. Im Fall $(y_1, y_2) \preceq (z_1, x_2^-)$ besagt Gleichung (6.33), daß

$$\begin{aligned} &\alpha_1 u_1(y_1) + [\alpha_1 u_1(z_1) + c\alpha_2 u_2(y_2) - \alpha_1 u_1(y_1)]q(0.5) \\ &\quad = \alpha_1 u_1(y_1) + c\alpha_2 u_2(y_2) + [\alpha_1 u_1(z_1) - \alpha_1 u_1(y_1) - c\alpha_2 u_2(y_2)]q(0.5) \end{aligned}\ .$$

Dies formen wir um zu

$$c\alpha_2 u_2(y_2) q(0.5) = c\alpha_2 u_2(y_2) - c\alpha_2 u_2(y_2) q(0.5)\,.$$

Teilt man diese Gleichung durch $c\alpha_2 u_2(y_2)$, so erhält man

$$q(0.5) = 1 - q(0.5)\,,$$

was zu $c = 1$ äquivalent ist.

Ist $(z_1, x_2^-) \prec (y_1, y_2)$, so erhält man durch eine analoge Argumentation ebenfalls $c = 1$.

Da $c = 1$ äquivalent zu $q(0.5) = 0.5$ ist, haben wir nun gezeigt, daß (i) erfüllt ist. Wir haben ferner gezeigt, daß

$$u(x_1, x_2) \;=\; \alpha_1 u_1(x_1) + \alpha_2 u_2(x_2) \qquad \text{für alle } (x_1, x_2) \in \mathcal{C}_1 \times \mathcal{C}_2\,.$$

Dies ist aber genau die additive Zerlegung der Nutzenfunktion. Aus der bivariaten Risikoneutralität folgt also eine additive Nutzenfunktion und $q(0.5) = 0.5$.

Wir nehmen nun umgekehrt an, daß u additiv ist und daß $q(0.5) = 0.5$ gilt. Dann ist $c = 1$ und wir erhalten aus Satz 6.3, Teil (*iii*), daß der ET bivariat risikoneutral ist. ■

Beweis von Satz 6.9

Teil 1: Korrelationsaversion impliziert $u_{x_1,x_2} \leq 0$ und $q'' \geq 0$.
Wir betrachten Lotterien L und M, die durch

$$L = \begin{bmatrix} (a_1, a_2) & (a_1, a_2 + \delta) & (b_1, a_2) & (b_1, a_2 + \delta) \\ 1 - p_0 - p & p & p & p_0 - p \end{bmatrix},$$

und

$$M = \begin{bmatrix} (a_1, a_2) & (a_1, a_2 + \delta) & (b_1, a_2) & (b_1, a_2 + \delta) \\ 1 - p_0 - p + \epsilon & p - \epsilon & p - \epsilon & p_0 - p + \epsilon \end{bmatrix},$$

gegeben sind, wobei $a_1 < b_1$, $\delta > 0$, $p_0 \in [0, 1]$ und $0 \leq p \leq \min\{p_0, 1 - p_0\}$ ist. Es gilt also $L \xrightarrow{CIT} M$.

Da u stetig ist, ist für hinreichend kleines δ $u(b_1, a_2) > u(a_1, a_2 + \delta)$. Da u außerdem monoton wachsend ist, berechnet sich der antizipierte Nutzen der Lotterien zu

$$\begin{aligned} AU(L) = u(a_1, a_2) &+ [u(a_1, a_2 + \delta) - u(a_1, a_2)]\, q(p_0 + p) \\ &+ [u(b_1, a_2) - u(a_1, a_2 + \delta)]\, q(p_0) \\ &+ [u(b_1, a_2 + \delta) - u(b_1, a_2)]\, q(p_0 - p) \end{aligned}$$

und

$$\begin{aligned} AU(M) = u(a_1, a_2) &+ [u(a_1, a_2 + \delta) - u(a_1, a_2)]\, q(p_0 + p - \epsilon) \\ &+ [u(b_1, a_2) - u(a_1, a_2 + \delta)]\, q(p_0) \qquad . \\ &+ [u(b_1, a_2 + \delta) - u(b_1, a_2)]\, q(p_0 - p + \epsilon)\ , \end{aligned}$$

Wegen $L \leq_D M$ muß dann $AU(L) \geq AU(M)$ gelten. Dies ist gleichwertig mit

$$\begin{aligned} &[u(a_1, a_2 + \delta) - u(a_1, a_2)][q(p_0 + p) - q(p_0 + p - \epsilon)] \\ &\qquad \geq [u(b_1, a_2 + \delta) - u(b_1, a_2)][q(p_0 - p + \epsilon) - q(p_0 - p)]\,. \end{aligned}$$

Dividiert man diese Ungleichung durch $[u(b_1, a_2 + \delta) - u(b_1, a_2)]$ und $[q(p_0 + p) - q(p_0 + p - \epsilon)]$, so erhält man

$$\frac{u(a_1, a_2 + \delta) - u(a_1, a_2)}{u(b_1, a_2 + \delta) - u(b_1, a_2)} \geq \frac{q(p_0 - p) - q(p_0 - p + \epsilon)}{q(p_0 + p) - q(p_0 + p - \epsilon)}\,.$$

Grenzübergang für $\delta \to 0$ und $\epsilon \to 0$ ergibt

$$\frac{u_{x_2}(a_1,a_2)}{u_{x_2}(b_1,a_2)} \geq \frac{q'(p_0-p)}{q'(p_0+p)}\,.$$

Da u zweimal differenzierbar ist, konvergiert die linke Seite für b_1 gegen a_1 gegen 1. Genauso konvergiert die rechte Seite für p gegen 0 gegen 1. Daher ist $u_{x_2}(b_1,a_2) \leq u_{x_2}(a_1,a_2)$ und $q'(p_0+p) \geq q'(p_0-p)$. Wegen der Differenzierbarkeitsvoraussetzungen ist dann aber auch $u_{x_1,x_2}(a_1,a_2) \leq 0$ and $q'' \geq 0$.

Teil 2: $u_{x_1,x_2} \leq 0$ und $q'' \geq 0$ impliziert Korrelationsaversion.
Wir zerlegen den Beweis in mehrere Schritte. Wir zeigen zuerst, daß die Behauptung richtig ist, falls P und Q Lotterien mit endlichem Träger sind, die sich nur um eine korrelationserhöhende Transformation unterscheiden. Im zweiten Schritt nehmen wir nur noch an, daß P und Q Lotterien mit endlichem Träger sind. Danach zeigen wir, daß die Behauptung für beliebige W-Maße auf $\mathbb{R}^2$ gilt.

Schritt 1: Seien P und Q Lotterien mit endlichem Träger, und gelte $P \xrightarrow{CIT} Q$. Sei $\{x^{(1)},\dots,x^{(m)}\}$ der gemeinsame Träger von P und Q. Wir schreiben kurz u_i für $u(x^{(i)})$, $i=1,\dots,m$, und nehmen an, daß $u_1 < \dots < u_m$. Sei weiter $p'_i = P(\{x^{(i)}\})$ und $p_i = Q(\{x^{(i)}\})$, $i=1,\dots,m$. Der antizipierte Nutzen von P ist genau dann größer als der von Q, wenn

$$\sum_{i=1}^{m}(u_i-u_{i-1})\Big[q(\sum_{j=i}^{m}p'_i) - q(\sum_{j=i}^{m}p_i)\Big] \geq 0$$

ist. Da sich P von Q nur durch eine korrelationserhöhende Transformation unterscheidet, gibt es Indizes $i_1 < \dots i_4$ mit $x^{(i_1)} = (a_1,a_2)$, $x^{(i_2)} = (a_1,b_2)$, $x^{(i_3)} = (b_1,a_2)$ und $x^{(i_4)} = (b_1,b_2)$, so daß

$$p'_i - p_i = \begin{cases} -\epsilon, & \text{falls } i=i_1 \text{ oder } i=i_4,\\ \epsilon, & \text{falls } i=i_2 \text{ oder } i=i_3,\\ 0, & \text{sonst.}\end{cases}$$

Die Differenz der verzerrten Wahrscheinlichkeiten ist dann durch

$$\Big[q(\sum_{j=i}^{n}p'_i) - q(\sum_{j=i}^{n}p_i)\Big] = \begin{cases} q(\sum_{j=i}^{m}p_i+\epsilon) - q(\sum_{j=i}^{m}p_i), & \text{falls } i_1 < i \leq i_2,\\ q(\sum_{j=i}^{m}p_i-\epsilon) - q(\sum_{j=i}^{m}p_i), & \text{falls } i_3 < i \leq i_4,\\ 0, & \text{sonst,}\end{cases}$$

gegeben. Somit gilt $AU(P) \geq AU(Q)$ genau dann, wenn

$$\begin{aligned}\sum_{i=i_1+1}^{i_2}(u_i-u_{i-1})\Big[q(\sum_{j=i}^{n}p_i+\epsilon) - q(\sum_{j=i}^{n}p_i)\Big]&\\ \geq \sum_{i=i_3+1}^{i_4}(u_i-u_{i-1})\Big[q(\sum_{j=i}^{n}p_i) - q(\sum_{j=i}^{n}p_i-\epsilon)\Big]&\end{aligned}$$

ist. Die eckigen Klammern auf der linken Seite sind nach unten durch $\epsilon q'(\sum_{i=i_2}^{n} p_i)$ beschränkt, die eckigen Klammern auf der rechten Seite nach oben durch $\epsilon q'(\sum_{i=i_3+1}^{n} p_i)$. Dies folgt mit dem Mittelwertsatz der Differentialrechnung und da q' monoton wachsend ist. Damit haben wir nur noch zu zeigen , daß

$$\epsilon q'(\sum_{i=i_2}^{n} p_i) \sum_{i=i_1+1}^{i_2} (u_i - u_{i-1}) \geq \epsilon q'(\sum_{i=i_3+1}^{n} p_i) \sum_{i=i_3+1}^{i_4} (u_i - u_{i-1})$$

gilt. Dies ist äquivalent zu

$$\epsilon q'(\sum_{i=i_2}^{n} p_i)[u(a_1, b_2) - u(a_1, a_2)] \geq \epsilon q'(\sum_{i=i_3+1}^{n} p_i)[u(b_1, b_2) - u(b_1, a_2)].$$

Wegen $u_{x_1,x_2} \leq 0$ ist klar, daß $[u(a_1, b_2) - u(a_1, a_2)] \geq [u(b_1, b_2) - u(b_1, a_2)]$ ist. Wegen $\sum_{i=i_2}^{n} p_i > \sum_{i=i_3+1}^{n} p_i$ und der Konvexität von q ist $q'(\sum_{i=i_2}^{n} p_i) > q'(\sum_{i=i_3+1}^{n} p_i)$. Damit ist der Ausdruck auf der linken Seite größer als der Ausdruck auf der rechten Seite. Wir haben also gezeigt, daß $AU(P) \geq AU(Q)$ gilt, falls P und Q Lotterien mit $P \stackrel{CIT}{\longrightarrow} Q$ sind.

Schritt 2: Seien P und Q Lotterien mit endlichem Träger, und gelte $P \leq_D Q$. Nach Satz 6.7 existiert eine Folge von Lotterien mit endlichem Träger $P = P_0$, $P_1, \ldots, P_n = Q$, so daß sich für $i = 1, \ldots n$ P_i von P_{i-1} durch eine korrelationserhöhende Transformation unterscheidet. Nach Schritt 1 gilt dann $AU(P) = AU(P_0) \geq AU(P_1) \geq \ldots \geq AU(P_n) = AU(Q)$.

Schritt 3: Seien P und Q W-Maße mit $P \leq_D Q$.
Nach Satz 6.8 existieren Folgen $(P_n)_{n\in\mathbb{N}}$ und $(Q_n)_{n\in\mathbb{N}}$ von Lotterien mit endlichem Träger, die schwach gegen P und Q konvergieren und für die $P_n \leq_D Q_n$ für alle $n \in \mathbb{N}$ gilt. Wir haben bereits gezeigt, daß dann $AU(P_n) \leq AU(Q_n)$ für alle $n \in \mathbb{N}$ gilt. Es bleibt zu zeigen, daß $\lim_{n\to\infty} AU(P_n) = AU(P)$ und $\lim_{n\to\infty} AU(Q_n) = AU(Q)$ gelten.

Da die Folge P_n schwach gegen P konvergiert und da u stetig ist, konvergieren auch die Bildmaße P_n^u schwach gegen P^u, siehe z.B. Billingsley (1968, S. 30). Daher konvergieren die zugehörigen dekumulativen Verteilungsfunktionen $F_n(t) = P_n(\{x : u(x) \geq t\})$ fast überall gegen die dekumulative Verteilungsfunktion $F(t) = P(\{x : u(x) \geq t\})$. Da q stetig ist, konvergiert dann auch die Folge $(q \circ F_n)_{n\in\mathbb{N}}$ fast überall gegen $q \circ F$. Da u beschränkt ist, können wir o.B.d.A. annehmen, daß $u(x) \in [0, 1]$ für alle $x \in \mathbb{R}^2$ ist. Dann ist $|q \circ F_n| \leq 1$ und 1 ist auf $[0, 1]$ integrierbar. Nach dem Satz von der majorisierten Konvergenz erhalten wir

$$AU(P_n) = \int_0^1 (q \circ F_n)(t)\, dt \stackrel{n\to\infty}{\longrightarrow} \int_0^1 (q \circ F)(t)\, dt = AU(P)\,.$$

Genauso erhält man, daß $\lim_{n\to\infty} AU(Q_n) = AU(Q)$ ist. Damit ist der noch fehlende Teil des Satzes bewiesen. ∎

Kapitel 7

Bestimmung der Modellparameter in AU und CEU

In diesem Kapitel wollen wir uns mit der Frage beschäftigen, wie ein konkretes Entscheidungsproblem zu modellieren ist. Insbesondere wollen wir die Frage untersuchen, wie Nutzen- und Verzerrungsfunktion in AU und Nutzenfunktion und Choquet-Kapazität in CEU bestimmt werden können.

Ein erster Schritt besteht in der Festlegung des zu verwendenden Modells. Es ist zunächst zu klären, ob ein Entscheidungsproblem unter Unsicherheit oder unter Risiko vorliegt. Bei Entscheidungen unter Unsicherheit ist zu entscheiden, ob man das Anscombe-Aumann-Modell zugrundelegt oder ein Modell, das ohne Verwendung objektiver Wahrscheinlichkeiten auskommt. Außerdem müssen Zustandsraum, Konsequenzenraum und die zur Auswahl stehenden Entscheidungsalternativen definiert werden.

Nachdem das benutzte Modell vollständig spezifiziert ist, kann mit der Elizitierung begonnen werden. Darunter versteht man die Bestimmung der Modellparameter, indem man den ET über seine Präferenzen auf geeignet gewählten Aktionen befragt. Die Modellparameter sind im AU-Modell Nutzenfunktion und Verzerrungsfunktion, im CEU-Modell Nutzenfunktion und Choquet-Kapazität.

In EU existieren eine Reihe von Standardverfahren zur Elizitierung der Nutzenfunktion. Zu nennen sind hier vor allem die Wahrscheinlichkeitsäquivalent-Methode und die Sicherheitsäquivalent-Methode. In diesen beiden Verfahren werden jeweils eine Lotterie mit zwei möglichen Konsequenzen mit einer sicheren Konsequenz verglichen. Für weitere Methoden der Elizitierung in EU und zu den dabei auftretenden Problemen vergleiche Farquhar (1984), McCord und de Neufville (1986), von Nitzsch und Weber (1986,1988) und Holz und Mosler (1992).

Wir werden in den folgenden Abschnitten zunächst Schmeidlers Axiomatisierung eines CEU-Modells im Rahmen des Anscombe-Aumann Ansatz betrachten. Danach werden wir CEU-Modelle ohne objektive Wahrscheinlichkeiten untersuchen. Wir werden dabei sowohl auf den Fall, daß der Zustandsraum reichhaltig ist (Savage-Ansatz) als auch auf den Fall, daß der Konsequenzenraum reichhaltig ist (topologischer Ansatz, algebraischer Ansatz), eingehen. Ferner werden wir Methoden zur Elizitierung von Nutzen- und Verzerrungsfunktion in AU-Modellen vorstellen. Für das AU-Modell wird des weiteren ein parametrischer Ansatz zur Bestimmung der Verzerrungsfunktion vorgeschlagen.

Besondere Anforderungen gegenüber der Elizitierung in EU liegen vor allem in der Tatsache begründet, daß im Gegensatz zum EU-Modell zusätzlich zur Nutzenfunktion noch die Verzerrungsfunktion bzw. statt des subjektiven W-Maßes die Choquet-Kapazität zu bestimmen ist.

7.1 Die Bestimmung von Nutzenfunktion und Choquet-Kapazität in Schmeidlers Axiomatisierung

In diesem Modell gestaltet sich die Bestimmung von Nutzenfunktion und Kapazität relativ einfach. Der Grund dafür ist die Reichhaltigkeit des Konsequenzenraums (alle Lotterien auf einer Menge Γ) sowie die Annahme, daß beim Vorliegen objektiver Wahrscheinlichkeiten Lotterien gemäß EU beurteilt werden. Daher kann der Prozeß der Bestimmung von Nutzenfunktion und Kapazität in zwei Phasen aufgespalten werden. In der ersten Phase werden dem ET zunächst Referenzlotterien, d.h. Lotterien mit objektiven Wahrscheinlichkeiten, zur Entscheidung vorgelegt. Damit kann gemäß den Verfahren zur Bestimmung der Nutzenfunktion in EU die Nutzenfunktion des ET erfragt werden. Sobald die Nutzenfunktion bekannt ist, kann dann in der zweiten Phase die Choquet-Kapazität bestimmt werden. Da jetzt nur noch die Choquet-Kapazität unbekannt ist, gestaltet sich die Elizitierung hier relativ einfach. Wir werden dieses Vorgehen im folgenden präzisieren.

Phase I: Bestimmung der Nutzenfunktion.
Da die Nutzenfunktion eindeutig bis auf positiv-affine Transformationen ist, kann man zunächst zwei Konsequenzen $x^-, x^+ \in \Gamma$ mit $x^- \prec x^+$ auswählen und $u(x^-) = 0$ und $u(x^+) = 1$ setzen.

- **Wahrscheinlichkeitsäquivalent-Methode:** Dem ET wird die Aufgabe gestellt zu einem gegebenen Ergebnis $x \in \Gamma$ einen Wert p_x so zu bestimmen, daß

$$x \sim \begin{bmatrix} x_1 & x_2 \\ 1 - p_x & p_x \end{bmatrix}$$

gilt. Sind $u(x_1)$ und $u(x_2)$ bekannt, so läßt sich der Wert von $u(x)$ zu $u(x) = (1-p_x)u(x_1)+p_x u(x_2)$ berechnen. Speziell für $x_1 = x^-$ und $x_2 = x^+$ erhält man $u(x) = p_x$. Auf diese Weise läßt sich also für beliebige Ergebnisse die Nutzenfunktion ermitteln.

- **Sicherheitsäquivalent-Methode:** Dem ET wird eine Lotterie

$$L = \begin{bmatrix} x_1 & x_2 \\ 1-p & p \end{bmatrix}$$

vorgelegt. Durch Befragen wird dann ein Ergebnis x_p bestimmt, so daß der ET indifferent zwischen L und x_p ist. Dann gilt

$$u(x_p) = u(x_1)(1-p) + u(x_2)p\,.$$

Sind also $u(x_1)$ und $u(x_2)$ bekannt, so ist auch der Nutzen von x_p bekannt.

Es ist jedoch zu beachten, daß ein Sicherheitsäquivalent i.a. nicht exisitieren muß, sofern der Raum Γ nicht reichhaltig genug ist. Ist Γ jedoch z.B. ein zusammenhängender, topologischer Raum und die Nutzenfunktion u stetig, so existiert immer ein Sicherheitsäquivalent.

Ein häufig benutztes Vorgehen erhält man mit $p = 0.5$. Man beginnt bei diesem Verfahren mit den zuvor festgelegten Ergebnissen x^- und x^+ und erhält dann im ersten Schritt ein Ergebnis $x_{0.5}$ mit $u(x_{0.5}) = 0.5$. Der Wert $x_{0.5}$ ist also bezüglich des Nutzens Mittelpunkt des Intervalls von x^- bis x^+. Man kann nun fortfahren, indem man die Intervalle $[x^-, x_{0.5}]$ und $[x_{0.5}, x^+]$ wieder halbiert. Setzt man dieses Verfahren weiter fort, kann man das Intervall $[x^-, x^+]$ beliebig fein aufteilen.

Phase II: Bestimmung der Choquet-Kapazität.
Wir gehen bei Phase II davon aus, daß die Nutzenfunktion u bekannt ist.

- **Lotterieäquivalent-Methode:** Dem ET wird eine Aktion

$$f = \begin{bmatrix} x_1 & x_2 \\ A^c & A \end{bmatrix}$$

präsentiert, wobei $x_1 \prec x_2$ ist. Der ET muß nun angeben, bei welcher Wahrscheinlichkeit p_A für ihn die Indifferenz

$$\begin{bmatrix} x_1 & x_2 \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} x_1 & x_2 \\ 1-p_A & p_A \end{bmatrix} \tag{7.1}$$

gilt. In diesem Fall ist

$$u(x_1)[1-\mu(A)] + u(x_2)\mu(A) = u(x_1)[1-p_A] + u(x_2)p_A$$

und somit $\mu(A) = p_A$. Wesentlich ist dabei, daß $x_1 \prec x_2$ ist, da der Choquet-Erwartungsnutzen einer Aktion ja von der Rangfolge der möglichen Ausgänge abhängt. Gilt (7.1) mit $x_1 \succ x_2$, so ergibt sich $\mu(A^c) = 1 - p_A$. Mit der Lotterieäquivalent-Methode kann also für beliebige Ereignisse A der Wert der Choquet-Kapazität $\mu(A)$ ermittelt werden.

- **Sicherheitsäquivalent-Methode:** Dem ET wird eine Aktion

$$f = \begin{bmatrix} x_1 & x_2 \\ A^c & A \end{bmatrix}$$

präsentiert, wobei $x_1 \prec x_2$ gilt. Sodann wird das Sicherheitsäquivalent x_A der Aktion f ermittelt, d.h. x_A wird so ermittelt, daß $x_A \sim f$ gilt. Dann ist

$$u(x_A) = u(x_1) + [u(x_2) - u(x_1)]\mu(A)$$

bzw.

$$\mu(A) = \frac{u(x_A) - u(x_1)}{u(x_2) - u(x_1)}\,.$$

Wie schon bei der Sicherheitsäquivalent-Methode in Phase I bemerkt, muß jedoch ein Sicherheitsäquivalent i.a. nicht existieren.

- **Aktionsäquivalent-Methode:** Sind $x_1 \prec x_2$ und $x_3 \prec x_4$ und ist die Indifferenz

$$\begin{bmatrix} x_1 & x_2 \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} x_3 & x_4 \\ A^c & A \end{bmatrix}$$

erfüllt, so gilt

$$u(x_1) + [u(x_2) - u(x_1)]\mu(A) = u(x_3) + [u(x_4) - u(x_3)]\mu(A)$$

und somit

$$\mu(A) = \frac{u(x_1) - u(x_3)}{u(x_4) - u(x_3) - u(x_2) + u(x_1)}\,.$$

Bei bekannter Nutzenfunktion erhält man also wieder den Wert von $\mu(A)$.

- **Ratio-Schätzungen aus dem Vergleich zweier Aktionen:** Zu gegebenen Ereignissen A, B und x_1, x_2 mit $x_1 \prec x_2$ ist $x_3 \succ x_1$ so zu bestimmen, daß

$$\begin{bmatrix} x_1 & x_2 \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} x_1 & x_3 \\ B^c & B \end{bmatrix}$$

gilt. Dann ist

$$u(x_1) + [u(x_2) - u(x_1)]\mu(A) = u(x_1) + [u(x_3) - u(x_1)]\mu(B)$$

bzw.

$$\frac{\mu(A)}{\mu(B)} = \frac{u(x_3) - u(x_1)}{u(x_2) - u(x_1)}\,.$$

Sind also die Nutzenfunktion u und z.B. $\mu(B)$ bekannt (etwa durch Anwendung der Sicherheitsäquivalent-Methode), so kann auf diese Art $\mu(A)$ ermittelt werden.

Die oben beschriebenen Methoden erlauben die Bestimmung der Nutzenfunktion und der Choquet-Kapazität. Wesentlich bei diesem Vorgehen ist jedoch die Voraussetzung, daß Lotterien mit objektiven Wahrscheinlichkeiten gegeben sind und gemäß EU beurteilt werden. Diese Voraussetzung erlaubt es, die Bestimmung von Nutzenfunktion und Choquet-Kapazität zu trennen.

7.2 Die Bestimmung von Nutzenfunktion und Choquet-Kapazität in CEU ohne objektive Wahrscheinlichkeiten

In den CEU-Modellen von Gilboa (1987), Wakker (1989a, 1989b, 1991), Nakamura (1990) stehen objektive Wahrscheinlichkeiten nicht zur Verfügung. Im Rahmen dieser Ansätze können dem ET keine Lotterien mit objektiven Wahrscheinlichkeiten zur Entscheidung vorgelegt werden. Eine separate Bestimmung der Nutzenfunktion ist also auf diese Weise nicht möglich. Wir unterscheiden nun zwei Fälle. Zunächst untersuchen wir Modelle, in denen der Konsequenzenraum reichhaltig ist, anschließend Modelle, in denen der Zustandsraum reichhaltig ist.

7.2.1 Bestimmung von Nutzenfunktion und Choquet-Kapazität in CEU bei reichhaltigem Konsequenzenraum

Wir wollen in diesem Abschnitt zeigen, wie in Modellen mit reichhaltigem Konsequenzenraum Nutzenfunktion und Choquet-Kapazität durch Befragung des ET ermittelt werden können. Wir wählen speziell den topologischen Ansatz, bemerken aber, daß die vorgestellten Verfahren auch im algebraischen Ansatz durchführbar sind.

Voraussetzung (für diesen Abschnitt): Wir betrachten ein Entscheidungsproblem unter Unsicherheit $((S,\mathcal{A}),(\mathcal{C},\mathcal{D}),\mathcal{F},\preceq)$. Der Konsequenzenraum $\mathcal{C}$ sei ein zusammenhängender topologischer Raum. Es existiere ein Ereignis $A \in \mathcal{A}$, das weder null noch universell ist. Es gelte CEU mit stetiger Nutzenfunktion u und Choquet-Kapazität μ.

Über den Zustandsraum machen wir keine weiteren Voraussetzungen. Er kann endlich oder auch unendlich sein. Demzufolge sind z.B. Fragen nach Sicherheitsäquivalenten erlaubt, da wir deren Existenz garantieren können. Da im schlimmsten Fall der Zustandsraum nur aus zwei Elementen bestehen kann, können wir jedoch keine Fragen nach Ereignisäquivalenten stellen, da es in dem Modell kein Ereignis geben muß, für das Indifferenz gilt.

Wir beschreiben nun eine Methode, mit deren Hilfe ohne die Verwendung objektiver Wahrscheinlichkeiten die Nutzenfunktion in einem CEU-Modell bestimmt werden kann.

Definition 7.1 (Standardfolge in CEU) *Sei* $N^+ = \{0, \ldots, n\}$, $n \in \mathbb{N}$, *oder* $N^+ = \mathbb{N}_0$. *Sei* A *nicht null oder universell, und seien* $a, b \in \mathcal{C}$. *Eine Menge von Konsequenzen* $\{z_i \,|\, i \in N^+\}$ *heißt* aufsteigende Standardfolge *bezüglich* (A, a, b), *wenn* $a \prec b \preceq z_i$ *für alle* $i \in N^+$ *und*

$$\begin{bmatrix} b & z_i \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} a & z_{i+1} \\ A^c & A \end{bmatrix}$$

für alle $i, i+1 \in N^+$ *gilt.*

Sei $N^- = \{0, \ldots, -n\}$, $n \in \mathbb{N}$, *oder* $N^- = \mathbb{Z}_0^-$. *Sei* A *nicht null oder universell, und seien* $a, b \in \mathcal{C}$. *Eine Menge von Konsequenzen* $\{z_i \,|\, i \in N^-\}$ *heißt* absteigende Standardfolge *bezüglich* (A, a, b), *wenn* $z_i \preceq a \prec b$ *für alle* $i \in N^-$ *und*

$$\begin{bmatrix} z_i & a \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} z_{i-1} & b \\ A^c & A \end{bmatrix}$$

für alle $i, i-1 \in N^-$ *gilt.*

Die für uns wesentliche Eigenschaft einer Standardfolge ist, daß die Nutzendifferenz zweier aufeinanderfolgender Elemente einer Standardfolge konstant ist. Insbesondere kann die Nutzenfunktion so gewählt werden, daß $u(z_i) = i$ für alle Elemente der Standardfolge gilt.

Satz 7.1 *Es sei* $\{z_i\}$ *eine (aufsteigende oder absteigende) Standardfolge in einem CEU-Modell. Dann kann die Nutzenfunktion* u *so gewählt werden, daß* $u(z_i) = i$ *für alle* $i \in N$ *ist.*

Beweis: Wir zeigen die Behauptung nur für aufsteigende Standardfolgen. Der Beweis für den Fall einer absteigenden Standardfolge verläuft analog. Wegen

$$\begin{bmatrix} b & z_0 \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} a & z_1 \\ A^c & A \end{bmatrix}$$

gilt

$$u(b)[1-\mu(A)] + u(z_0)\mu(A) \;=\; u(a)[1-\mu(A)] + u(z_1)\mu(A)\,.$$

Eine einfache Umformung ergibt

$$u(z_1) - u(z_0) \;=\; [u(b) - u(a)]\frac{1-\mu(A)}{\mu(A)} > 0\,.$$

Daher kann die Nutzenfunktion so gewählt werden, daß $u(z_0) = 0$ und $u(z_1) = 1$ ist. Für diese Wahl von u gilt dann

$$u(b) - u(a) \;=\; \frac{\mu(A)}{1-\mu(A)}\,. \tag{7.2}$$

Für $i = 0$ und $i = 1$ ist die Aussage des Satzes also erfüllt. Wir zeigen nun, daß aus $u(z_i) = i$ folgt, daß auch $u(z_{i+1}) = i + 1$ ist, sofern $i + 1 \in N$ ist. Die Behauptung folgt dann durch vollständige Induktion.

Nach Definition einer Standardfolge gilt

$$\begin{bmatrix} b & z_i \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} a & z_{i+1} \\ A^c & A \end{bmatrix},$$

was gleichbedeutend ist mit

$$u(b)[1-\mu(A)] + u(z_i)\mu(A) \;=\; u(a)[1-\mu(A)] + u(z_{i+1})\mu(A)\,.$$

Umformung der Gleichung ergibt

$$u(z_{i+1}) - u(z_i) \;=\; [u(b)-u(a)]\frac{1-\mu(A)}{\mu(A)}\,.$$

Unter Beachtung von (7.2) und mit der Voraussetzung $u(z_i) = i$ erhält man schließlich $u(z_{i+1}) = i + 1$, wie behauptet. ■

Gelingt es uns also, eine Standardfolge zu bestimmen, so haben wir auch für die Elemente der Standardfolge eine Nutzenfunktion bestimmt. Die Konstruktion einer aufsteigenden Standardfolge liefert jedoch nur Nutzenwerte für Konsequenzen oberhalb von z_0. Um Nutzenwerte für weniger präferierte Konsequenzen zu bestimmen, gibt es nun zwei Möglichkeiten. Wir können zum einen z_0 möglichst klein wählen, so daß nur „wenige“ Konsequenzen schlechter als z_0 beurteilt werden. Die zweite Möglichkeit besteht darin, ausgehend von z_0 eine absteigende Standardfolge zu bestimmen, deren Nutzendifferenz mit der Nutzendifferenz der aufsteigenden Standardfolge identisch ist. Eine Antwort darauf, wie dies zu bewerkstelligen ist, gibt der folgende Satz:

Satz 7.2 *Seien $\{z_i \,|\, i \in N^+\}$ eine aufsteigende Standardfolge bezüglich (A, a, z_0) und $\{z_i \,|\, i \in N^-\}$ eine absteigende Standardfolge bezüglich (A, z_1, b). Gilt*

$$z_1 \sim \begin{bmatrix} z_0 & b \\ A^c & A \end{bmatrix}, \tag{7.3}$$

so kann u so gewählt werden, daß $u(z_i) = i$ für alle $i \in N^+ \cup N^-$ gilt.

Beweis: Nach dem vorangehendem Satz kann u so gewählt werden, daß für die Elemente der aufsteigenden Standardfolge $u(z_i) = i$ gilt.

Ist (7.3) erfüllt, so muß wegen $z_0 \prec z_1$ auch $z_1 \prec b$ gelten. Somit ist (7.3) äquivalent zu $1 = \mu(A) \cdot u(b)$.

Sind $i, i-1 \in \mathbb{N}^-$, so gilt

$$\begin{bmatrix} z_i & z_1 \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} z_{i-1} & b \\ A^c & A \end{bmatrix},$$

was wiederum zu

$$u(z_i)[1-\mu(A)] + \mu(A) = u(z_{i-1})[1-\mu(A)] + u(b)\mu(A)$$

äquivalent ist. Wegen $\mu(A) \cdot u(b) = 1$ ist dann

$$u(z_i)[1-\mu(A)] + \mu(A) = u(z_{i-1})[1-\mu(A)] + 1$$

und somit $u(z_{i-1}) = u(z_i) - 1$. Die Elemente der absteigenden Standardfolge haben also ebenfalls die Nutzendifferenz 1. Mit Induktion folgt dann, daß $u(z_i) = i$ für alle $i \in N^+ \cup N^-$ gilt. ∎

Wir beschreiben nun ein Verfahren zur Bestimmung der Nutzenfunktion mittels Verwendung von Standardfolgen.

Schritt 1: Bestimmung der aufsteigenden Standardfolge.
Zunächst werden ein Ereignis A und zwei Konsequenzen $a, b \in \mathcal{C}$ mit $a \prec b$ zur Konstruktion der Standardfolge beliebig gewählt. Allerdings sollten a und b nicht zu „weit auseinander“ liegen. Je näher a und b beisammen liegen, desto feiner wird der Konsequenzenraum unterteilt. Als nächstes muß das erste Element z_0 der Standardfolge gewählt werden. Da $b \preceq z_0$ gelten muß, können wir $z_0 = b$ wählen. Die Standardfolge wird also bezüglich (A, a, z_0) gebildet. Man erhält dann z_1 aus der folgenden Aufgabe.

Bestimme z_1 so, daß $z_0 \sim \begin{bmatrix} a & z_1 \\ A^c & A \end{bmatrix}$ gilt.

Ist ein Wert z_i der Standardfolge bestimmt, so erhält man einen weiteren Wert z_{i+1} aus der folgenden Aufgabe.

Bestimme z_{i+1} so, daß $\begin{bmatrix} z_0 & z_i \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} a & z_{i+1} \\ A^c & A \end{bmatrix}$ gilt.

Dieses Verfahren wird so lange fortgesetzt, bis entweder genug Information über die Nutzenfunktion beschafft ist oder bis kein weiteres z_{i+1} mehr bestimmt werden kann.

Schritt 2: Abstimmung der aufsteigenden mit einer absteigenden Standardfolge.
Dem ET wird die folgende Aufgabe gestellt:

$$\text{Bestimme } b \text{ so, daß } z_1 \sim \begin{bmatrix} z_0 & b \\ A^c & A \end{bmatrix} \text{ gilt.}$$

Gemäß Satz 7.2 hat jede bezüglich (A, z_1, b) gebildete absteigende Standardfolge die gleiche Nutzendifferenz wie die in Schritt 1 konstruierte aufsteigende Standardfolge.

Schritt 3: Bestimmung der absteigenden Standardfolge.
Die absteigende Standardfolge wird also bezüglich (A, z_1, b) gebildet, wobei b die in Schritt 2 gewonnene Konsequenz ist. Ist z_i bereits konstruiert, so gewinnt man z_{i-1} aus der Aufgabe:

$$\text{Bestimme } z_{i-1} \text{ so, daß } \begin{bmatrix} z_i & z_1 \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} z_{i-1} & b \\ A^c & A \end{bmatrix} \text{ gilt.}$$

Dieses Verfahren wird wieder so lange fortgesetzt, bis entweder genug Information über die Nutzenfunktion beschafft ist oder bis kein weiteres z_{i-1} mehr bestimmt werden kann.

Verfeinerung der Standardfolge.
Ist der Konsequenzenraum dann noch nicht fein genug unterteilt, kann man eine zweite Standardfolge $\{v_i\}$ konstruieren, für die $u(v_{2i}) = i$, $u(v_{2i+1}) = i + \frac{1}{2}$ ist, d.h. eine Standardfolge, deren Nutzendifferenz gerade halb so groß ist wie die der ersten Standardfolge. Dies wird im folgenden beschrieben.

Zur Verfeinerung der Standardfolge müssen wir b' so klein machen, daß für die bezüglich (A, a, b') mit $v_0 = z_0$ konstruierte Standardfolge gerade $v_2 = z_1$ gilt. Man wählt also zunächst ein b' mit $a \prec b' \prec b$. v_1 wird dann gemäß folgender Aufgabe bestimmt.

$$\text{Bestimme } v_1 \text{ so, daß } \begin{bmatrix} b' & z_0 \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} a & v_1 \\ A^c & A \end{bmatrix} \text{ gilt.}$$

Wir unterscheiden sodann drei Fälle:

(i) Ist $\begin{bmatrix} b' & v_1 \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} a & z_1 \\ A^c & A \end{bmatrix}$, so ist b' richtig gewählt.

(ii) Ist $\begin{bmatrix} b' & v_1 \\ A^c & A \end{bmatrix} \prec \begin{bmatrix} a & z_1 \\ A^c & A \end{bmatrix}$, so ist b' zu klein gewählt. Das Verfahren muß dann mit einem größeren b' wiederholt werden.

(iii) Ist $\begin{bmatrix} b' & v_1 \\ A^c & A \end{bmatrix} \succ \begin{bmatrix} a & z_1 \\ A^c & A \end{bmatrix}$, so ist b' zu groß gewählt. Das Verfahren muß dann mit einem kleineren b' wiederholt werden.

Ist mit diesem iterativen Verfahren ein geeignetes b' gefunden, so ergeben sich die Konsequenzen z_{2i+1} aus folgender Aufgabe (die v_{2i} sind sowieso mit z_i identisch).

Bestimme v_{2i+1} so, daß $\begin{bmatrix} b' & z_i \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} a & v_{2i+1} \\ A^c & A \end{bmatrix}$ gilt.

Auf diese Art kann der Konsequenzenraum beliebig fein eingeteilt werden. Ist die Unterteilung hinreichend fein, können im nächsten Schritt für beliebige Ereignisse B obere und untere Schranken für $\mu(B)$ bestimmt werden.

Bestimmung von $\mu(B)$

Hat man eine auf- und eine absteigende Standardfolge $\{z_i \,|\, i \in N^+ \cup N^-\}$ bestimmt, so erhält man Abschätzungen für $\mu(B)$ wie folgt. Der ET wird aufgefordert zu gegebenen $i, j \in N^+ \cup N^-$, $i < j$, den Wert $k \in N^+ \cup N^-$ so zu bestimmen, daß

$$z_k \preceq \begin{bmatrix} z_i & z_j \\ B^c & B \end{bmatrix} \prec z_{k+1}$$

gilt. Dann ist offensichtlich $k \leq i + (j - i)\mu(B) < k + 1$ und somit

$$\frac{k - i}{j - i} \leq \mu(B) < \frac{k - i + 1}{j - i}.$$

Man erhält als untere und obere Schranken für $\mu(B)$. Diese Schranken sind umso enger je größer die Differenz $j - i$ ist. Durch Verfeinerung der Standardfolge wie oben beschrieben, kann also auch $\mu(B)$ beliebig gut abgeschätzt werden.

7.2.2 Bestimmung von Nutzenfunktion und Choquet-Kapazität in CEU bei reichhaltigem Zustandsraum

In diesem Abschnitt untersuchen wir Modelle, in denen der Zustandsraum in einem gewissen Sinne reichhaltig ist, und zeigen, wie dort Nutzenfunktion und Choquet-Kapazität bestimmt werden können. Wir werden speziell die Axiomatisierung von Gilboa (1987) betrachten und machen daher die folgende

Voraussetzung (für diesen Abschnitt): Wir betrachten ein Entscheidungsproblem unter Unsicherheit $((S, \mathcal{A}), (\mathcal{C}, \mathcal{D}), \mathcal{F}, \preceq)$. Es gelte CEU mit Nutzenfunktion u und Choquet-Kapazität μ. Für alle Ereignisse A, B mit $B \subset A$ und alle $\alpha \in [0, 1]$ gibt es ein Ereignis C mit $B \subset C \subset A$, so daß

$$\mu(C) = \alpha\mu(B) + (1 - \alpha)\mu(A)$$

ist. Der Konsequenzenraum enthält mindestens drei Konsequenzen x_*, x, x^*, so daß $x_* \prec x \prec x^*$ ist.

Über den Konsequenzenraum setzen wir also lediglich voraus, daß es mindestens drei unterschiedlich präferierte Konsequenzen gibt. Wir können daher keine Fragen nach Sicherheitsäquivalenten stellen, da diese unter den gemachten Voraussetzungen nicht existieren müssen. Stattdessen können wir Fragen nach Ereignisäquivalenten stellen.

Wir beschreiben nun eine Methode, die die Bestimmung der Nutzenfunktion und Choquet-Kapazität in einem CEU-Modell mit reichen Zustandsraum ermöglicht. In Analogie zum vorigen Abschnitt besteht die Grundidee des Verfahrens darin, Folgen von Ereignissen zu konstruieren. Wir gehen dabei wie folgt vor:

Wir setzen zunächst $u(x_*) = 0$ und $u(x^*) = 1$. Dann ist $u(x) = c \in (0,1)$.

Schritt 1: Bestimmung des Ereignisäquivalents zu x.
Bestimme eine Ereignis A_0 so, daß

$$x \sim \begin{bmatrix} x_* & x^* \\ A_0^c & A_0 \end{bmatrix} \tag{7.4}$$

gilt.

Schritt 2: Konstruktion einer absteigenden Folge von Ereignissen.
Wir nehmen an, daß Ereignis A_i, $i \in \mathbb{Z}_-$, bereits konstruiert ist. Dann gewinnt man A_{i-1} aus der folgenden Aufgabe:

Bestimme ein Ereignis A_{i-1} so, daß $A_{i-1} \subset A_i$ ist und

$$\begin{bmatrix} x_* & x \\ A_i^c & A_i \end{bmatrix} \sim \begin{bmatrix} x_* & x^* \\ A_{i-1}^c & A_{i-1} \end{bmatrix} \tag{7.5}$$

gilt.

Dieser Schritt kann theoretisch beliebig oft wiederholt werden. In der Praxis wird man diesen Schritt solange wiederholen, bis eine genügend große Anzahl von Ereignissen bestimmt ist.

Schritt 3: Konstruktion einer aufsteigenden Folge von Ereignissen.
Wir nehmen an, daß Ereignis A_i, $i \in \mathbb{N}$, bereits konstruiert ist. Dann gewinnt man A_{i+1} aus der folgenden Aufgabe:

Bestimme ein Ereignis A_{i+1} so, daß $A_i \subset A_{i+1}$ ist und daß

$$\begin{bmatrix} x & x^* \\ A_i^c & A_i \end{bmatrix} \sim \begin{bmatrix} x_* & x^* \\ A_{i+1}^c & A_{i+1} \end{bmatrix} \tag{7.6}$$

gilt.

Wie Schritt 2 wird man auch diesen Schritt so oft wiederholen, bis eine geeignete Anzahl von Mengen A_i bestimmt ist.

Wir erhalten durch diese drei Schritte eine bezüglich Teilmengenbeziehung vollständig geordnete Kette $A_{-m} \subset \ldots A_n$, $m, n \in \mathbb{N}$, von Ereignissen. Wir können für die Elemente dieser Kette die Werte der Kapazität in Abhängigkeit von $c = u(x)$ bestimmen.

Satz 7.3 *Für die gemäß Schritt 1 bis 3 konstruierte Folge von Ereignissen $A_{-m} \subset \ldots \subset A_n$, $m, n \in \mathbb{N}$, gilt*

$$\mu(A_i) = \begin{cases} 1 - (1-c)^{i+1}, & \textit{falls } i \geq 0, \\ c^{1-i}, & \textit{falls } i < 0. \end{cases}$$

Beweis: Aus der Bestimmung von A_0 errechnet man sofort, daß $c = \mu(A_0)$ ist.

Wir zeigen zunächst mit vollständiger Induktion, daß die Behauptung für $i \leq 0$ richtig ist. Für $i = 0$ haben wir bereits gezeigt, daß $\mu(A_0) = c$ ist. Sei also $\mu(A_i) = c^{1-i}$ bereits gezeigt. Wir zeigen, daß die Behauptung dann auch für $i-1$ gilt. Nach Konstruktion ist $c\mu(A_i) = \mu(A_{i-1})$. Somit ist $\mu(A_{i-1}) = c^{1-(i-1)}$, wie behauptet.

Wir zeigen ebenfalls mit vollständiger Induktion, daß die Behauptung auch für $i \geq 0$ richtig ist. Für $i = 0$ ist $\mu(A_0) = 1-(1-c)$ und die Behauptung somit richtig. Wir nehmen nun an, daß $\mu(A_i) = 1-(1-c)^{i+1}$ für ein $i \geq 0$ bewiesen ist und zeigen, daß die Behauptung dann auch für $i+1$ gültig ist. Nach Konstruktion von A_{i+1} ist

$$c[1-\mu(A_i)] + \mu(A_i) = \mu(A_{i+1}).$$

Wegen der Induktionsvoraussetzung ist dann

$$\mu(A_{i+1}) = c(1-c)^{i+1} + 1 - (1-c)^{i+1} = 1-(1-c)^{i+2},$$

wie behauptet. ∎

Das Problem ist natürlich, daß der Wert von c nicht bekannt ist. Wir können jedoch durch einfache Aktionenvergleiche obere und untere Schranken für c gewinnen. Dem ET werden für $i, j > 0$ die beiden Aktionen

$$L_{i,j} = \begin{bmatrix} x_* & x & x^* \\ A_i^c & A_i \setminus A_{-j} & A_{-j} \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} x_* & x^* \\ A_0^c & A_0 \end{bmatrix} = L_0$$

präsentiert. Der ET muß nun angeben, ob $L_{i,j} \prec L_0$, $L_{i,j} \sim L_0$ oder $L_{i,j} \succ L_0$ gilt.

Die Differenz der Choquet-Erwartungsnutzen der beiden Aktionen ergibt sich zu

$$\begin{aligned} CEU(L_{i,j}) - CEU(L_0) &= c[1-(1-c)^{i+1}] + (1-c)c^{j+1} - c \\ &= c(1-c)[c^j - (1-c)^i]. \end{aligned}$$

Da $c \in (0,1)$ ist, ist $c(1-c) > 0$ und es ist $CEU(L_{i,j}) \geq CEU(L_0)$ genau dann, wenn $c^j - (1-c)^i \geq 0$ ist. Entsprechend ist $CEU(L_{i,j}) \leq CEU(L_0)$ genau dann, wenn $c^j - (1-c)^i \leq 0$ ist. Da $\lim_{c\to 0} c^j - (1-c)^i = -1$, $\lim_{c \to 1} c^j - (1-c)^i = 1$ und $c^j - (1-c)^i$ streng monoton wachsend in c ist, hat $c^j - (1-c)^i$ genau eine Nullstelle in $(0,1)$. Diese bezeichnen wir mit $c_{i,j}$. Wir können somit festhalten:

(i) Gilt $L_{i,j} \prec L_0$, so ist $c < c_{i,j}$.

(ii) Gilt $L_{i,j} \sim L_0$, so ist $c = c_{i,j}$.

(iii) Gilt $L_{i,j} \succ L_0$, so ist $c > c_{i,j}$.

Indem wir den ET für verschiedene Paare i, j nach seinen Präferenzen fragen, können wir also obere und untere Schranken für c bestimmen. Dazu muß lediglich jeweils $c_{i,j}$ als eindeutige Lösung der Gleichung $c^j = (1-c)^i$ in $(0,1)$ bestimmt werden. Dies ist numerisch jedoch leicht zu bewerkstelligen, z.B. durch Anwendung des Newton-Verfahrens oder des Bisektionsverfahrens. In der folgenden Tabelle sind für einige Paare von i und j die Werte von $c_{i,j}$ vertafelt.

i \ j	1	2	3	4	5	6	7	8
1	0.5000	0.6180	0.6823	0.7245	0.7549	0.7781	0.7965	0.8117
2	0.3820	0.5000	0.5698	0.6180	0.6540	0.6823	0.7053	0.7245
3	0.3177	0.4302	0.5000	0.5497	0.5877	0.6180	0.6431	0.6642
4	0.2755	0.3820	0.4503	0.5000	0.5386	0.5698	0.5959	0.6180
5	0.2451	0.3460	0.4123	0.4614	0.5000	0.5316	0.5581	0.5808
6	0.2219	0.3177	0.3820	0.4302	0.4684	0.5000	0.5267	0.5497
7	0.2035	0.2947	0.3569	0.4041	0.4419	0.4733	0.5000	0.5231
8	0.1883	0.2755	0.3358	0.3820	0.4192	0.4503	0.4769	0.5000

Hat man untere und obere Schranken für c, so hat man natürlich auch untere und obere Schranken für die Werte $\mu(A_i)$.

7.3 Die Bestimmung von Nutzen- und Verzerrungsfunktion in AU

Bei der Bestimmung von Nutzenfunktion und Verzerrungsfunktion in AU können geeignet modifizierte Varianten der in Abschnitt 7.2 vorgestellten Verfahren angewendet werden.

Voraussetzung (für diesen Abschnitt): Wir betrachten ein Entscheidungsproblem unter Risiko $((\mathcal{C},\mathcal{D}),\mathcal{P},\preceq)$. Es gelte AU mit Nutzenfunktion u und Verzerrungsfunktion q.

Reichhaltiger Konsequenzenraum

Ist der Konsequenzenraum reich, so kann die Nutzenfunktion wie in Abschnitt 7.2.1 mit Hilfe von Standardfolgen bestimmt werden.

Definition 7.2 (Standardfolge in AU) *Sei $N^+ = \{0,\dots,n\}$, $n \in \mathbb{N}$, oder $N^+ = \mathbb{N}_0$. Sei $p \in (0,1)$, und seien $a, b \in \mathcal{C}$. Eine Menge von Konsequenzen $\{z_i \,|\, i \in N^+\}$ heißt* aufsteigende Standardfolge *bezüglich (p,a,b), wenn $a \prec b \preceq z_i$ für alle $i \in N^+$ und*

$$\begin{bmatrix} b & z_i \\ 1-p & p \end{bmatrix} \sim \begin{bmatrix} a & z_{i+1} \\ 1-p & p \end{bmatrix}$$

für alle $i, i+1 \in N^+$ *gilt.*

Sei $N^- = \{0, \ldots, -n\}$, $n \in \mathbb{N}$, *oder* $N^- = \mathbb{Z}_0^-$. *Sei* $p \in (0,1)$, *und seien* $a, b \in \mathcal{C}$. *Eine Menge von Konsequenzen* $\{z_i \,|\, i \in N^-\}$ *heißt* absteigende Standardfolge *bezüglich* (p, a, b), *wenn* $z_i \preceq a \prec b$ *für alle* $i \in N^-$ *und*

$$\begin{bmatrix} z_i & a \\ 1-p & p \end{bmatrix} \sim \begin{bmatrix} z_{i-1} & b \\ 1-p & p \end{bmatrix}$$

für alle $i, i-1 \in N^-$ *gilt.*

Weiter gelten in Analogie zu den Sätzen 7.1 und 7.2 die folgenden Sätze:

Satz 7.4 *Es sei* $\{z_i\}$ *eine (aufsteigende oder absteigende) Standardfolge in einem AU-Modell. Dann kann die Nutzenfunktion* u *so gewählt werden, daß* $u(z_i) = i$ *für alle* $i \in N$ *ist.*

Satz 7.5 *Seien* $\{z_i \,|\, i \in N^+\}$ *eine aufsteigende Standardfolge bezüglich* (p, a, z_0) *und* $\{z_i \,|\, i \in N^-\}$ *eine absteigende Standardfolge bezüglich* (p, z_1, b). *Gilt*

$$z_1 \sim \begin{bmatrix} z_0 & b \\ 1-p & p \end{bmatrix}, \tag{7.7}$$

so kann u *so gewählt werden, daß* $u(z_i) = i$ *für alle* $i \in N^+ \cup N^-$ *gilt.*

Der Beweis der beiden Sätze verläuft völlig analog zum Beweis der Sätze 7.1 und 7.2. Die beiden Sätze zeigen, daß in AU mit reichem Konsequenzenraum eine Bestimmung der Nutzenfunktion über Standardfolgen wie in CEU möglich ist. Die in Abschnitt 7.2.1 beschrieben Verfahren zur Konstruktion von auf- und absteigenden Standardfolgen gelten mutatis mutandis auch in AU. Es muß jeweils „Ereignis A“ durch „Wahrscheinlichkeit p“ und „Ereignis A^c“ durch „Wahrscheinlichkeit $1-p$“ ersetzt werden.

Eine Abschätzung des Wertes der Verzerrungsfunktion an der Stelle p ist ebenso möglich wie die Abschätzung des Wertes von $\mu(A)$: Ist $i < j$ und

$$z_k \preceq \begin{bmatrix} z_i & z_j \\ 1-p & p \end{bmatrix} \preceq z_{k+1},$$

so ist

$$\frac{k-i}{j-i} \leq q(p) \leq \frac{k-i+1}{j-i}.$$

Konsequenzenraum nicht reichhaltig

Da wir davon ausgehen, daß $\mathcal{P}$ alle Lotterien mit endlichem Träger enthält (was der Reichhaltigkeit des Zustandsraumes in CEU entspricht), sind die Methoden aus Abschnitt 7.2.2 auf AU übertragbar und anwendbar. Wie setzen wie dort lediglich die Existenz dreier Konsequenzen x_*, x, x^* mit $x_* \prec x \prec x^*$ voraus. Wir beschreiben nun der Vollständigkeit halber kurz das Verfahren zur Bestimmung der Verzerrungsfunktion und der Nutzenfunktion.

Normierung: Wir setzen $u(x_*) = 0$ und $u(x^*) = 1$. Dann ist $u(x) = c \in (0,1)$.

Schritt 1: Bestimmung des Wahrscheinlichkeitsäquivalents zu x.
Bestimme eine Wahrscheinlichkeit p_0 so, daß

$$x \sim \begin{bmatrix} x_* & x^* \\ 1-p_0 & p_0 \end{bmatrix}$$

gilt.

Schritt 2: Konstruktion einer fallenden Folge von Wahrscheinlichkeiten.
Ist die Wahrscheinlichkeit p_i, $i \in \mathbb{Z}_-$, bereits konstruiert, so erhält man p_{i-1} aus der Aufgabe:

Bestimme eine Wahrscheinlichkeit p_{i-1} so, daß

$$\begin{bmatrix} x_* & x \\ 1-p_i & p_i \end{bmatrix} \sim \begin{bmatrix} x_* & x^* \\ 1-p_{i-1} & p_{i-1} \end{bmatrix}$$

gilt.

Wiederhole Schritt 2 bis eine geeignete Anzahl von Wahrscheinlichkeiten bestimmt ist.

Schritt 3: Konstruktion einer wachsenden Folge von Wahrscheinlichkeiten.
Ist die Wahrscheinlichkeit p_i, $i \in \mathbb{N}$, bereits konstruiert, so erhält man p_{i+1} aus der Aufgabe:

Bestimme eine Wahrscheinlichkeit p_{i+1} so, daß

$$\begin{bmatrix} x & x^* \\ 1-p_i & p_i \end{bmatrix} \sim \begin{bmatrix} x_* & x^* \\ 1-p_{i+1} & p_{i+1} \end{bmatrix}$$

gilt.

Wiederhole Schritt 3 bis eine geeignete Anzahl von Wahrscheinlichkeiten bestimmt ist.

Für die so gewonnen Werte $p_{-m}, \ldots, p_n$ ist $p_{-m} < \ldots < p_n$ und wir können die Werte der Verzerrungsfunktion q in Abhängigkeit von $c = u(x)$ bestimmen. Es gilt nämlich

$$q(p_i) = \begin{cases} 1-(1-c)^{i+1}, & \text{falls } i \geq 0, \\ c^{1-i}, & \text{falls } i < 0. \end{cases}$$

Bestimmung von Schranken für c: Dem ET werden die beiden Lotterien

$$L_{i,j} = \begin{bmatrix} x_* & x & x^* \\ 1-p_i & p_i - p_{-j} & p_{-j} \end{bmatrix} \quad \text{und} \quad \begin{bmatrix} x_* & x^* \\ 1-p_0 & p_0 \end{bmatrix} = L_0$$

zum Vergleich vorgelegt. Der ET muß angeben, ob $L_{i,j} \prec L_0$, $L_{i,j} \sim L_0$ oder $L_{i,j} \succ L_0$ gilt.

(i) Gilt $L_{i,j} \prec L_0$, so ist $c < c_{i,j}$.

(ii) Gilt $L_{i,j} \sim L_0$, so ist $c = c_{i,j}$.

(iii) Gilt $L_{i,j} \succ L_0$, so ist $c > c_{i,j}$.

Dabei ist $c_{i,j}$ wieder die eindeutige Lösung der Gleichung $c^j = (1-c)^i$ in $(0,1)$. Indem wir den ET für verschiedene Paare i, j nach seinen Präferenzen fragen, können wir also obere und untere Schranken für c bestimmen und damit auch ür die Werte $q(p_i)$.

7.4 Parametrische Ansätze zur Bestimmung der Verzerrungsfunktion

Aufgrund der oben genannten Schwierigkeiten bei der Bestimmung von Nutzen- und Verzerrungsfunktion in AU bieten sich auch parametrische Ansätze an. Wir schlagen vor, die Verzerrungsfunktion q aus einer parametrischen Klasse von Verzerrungsfunktionen zu bestimmen. Bei der Wahl einer geeigneten parametrischen Klasse sollten folgende Punkte beachtet werden.

- Die Klasse sollte durch wenige Parameter beschrieben werden, die einfach interpretierbar sind.
- Die Klasse sollte groß genug sein, um möglichst viele mögliche Präferenzrelationen modellieren zu können.
- Die Klasse sollte analytisch gut handhabbar sein.

Es macht also z.B. wenig Sinn eine Klasse zu wählen, die nur konvexe Verzerrungsfunktionen enthält, da mit ihnen (stark) risikoaverses Verhalten nicht beschreiben werden kann. (Das entsprechende gilt natürlich für eine Klasse, die nur konkave Funktionen enthält).

Verzerrungsfunktionen nach Tversky und Kahneman

Tversky und Kahneman (1992) benutzen Verzerrungsfunktionen der folgenden Gestalt:

$$q_\gamma(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}}.$$

Wir bezeichnen die Menge aller dieser Verzerrungsfunktionen mit Q_{KT}. Definiert sind diese Verzerrungsfunktionen für $\gamma > 0$. Ist jedoch γ kleiner als 0.2792, so ist q_γ nicht mehr monoton wachsend. Für $\gamma = 1$ erhält man die identische Verzerrungsfunktion. Für $\gamma < 2$, $\gamma \neq 1$ hat q_γ genau einen Schnittpunkt c mit der Winkelhalbierenden in $(0,1)$. Dieser Schnittpunkt ist monoton wachsend mit γ. Die Verzerrungsfunktion liegt für $\gamma < 1$ in $[0,c]$ oberhalb der Winkelhalbierenden und in $[c,1]$ unterhalb. Für $1 < \gamma < 2$ liegt q_γ in $[0,c]$ unterhalb der Winkelhalbierenden und in $[c,1]$ oberhalb. Für $\gamma \geq 2$ ist q_γ konvex.

Der erste und dritte Punkt unserer Forderungen sind zwar erfüllt, da die Klasse durch nur einen Parameter beschrieben wird und die Funktionen relativ einfach aufgebaut sind. Jedoch ist unserer Ansicht nach die Klasse nicht groß genug. Zwar sind S-förmige Verzerrungsfunktionen enthalten, doch können der Schnittpunkt mit der Winkelhalbierenden und die Krümmung der Verzerrungsfunktion nicht unabhängig voneinander variiert werden. Ebenso fehlen konvexe Funktionen, die zur Modellierung (stark) risikofreudigen Verhaltens notwendig sind. Außerdem ist die Klasse nicht abgeschlossen gegen Dualisierung, d.h. die Verzerrungsfunktion q_γ^D liegt i.a. nicht in Q_{KT}.

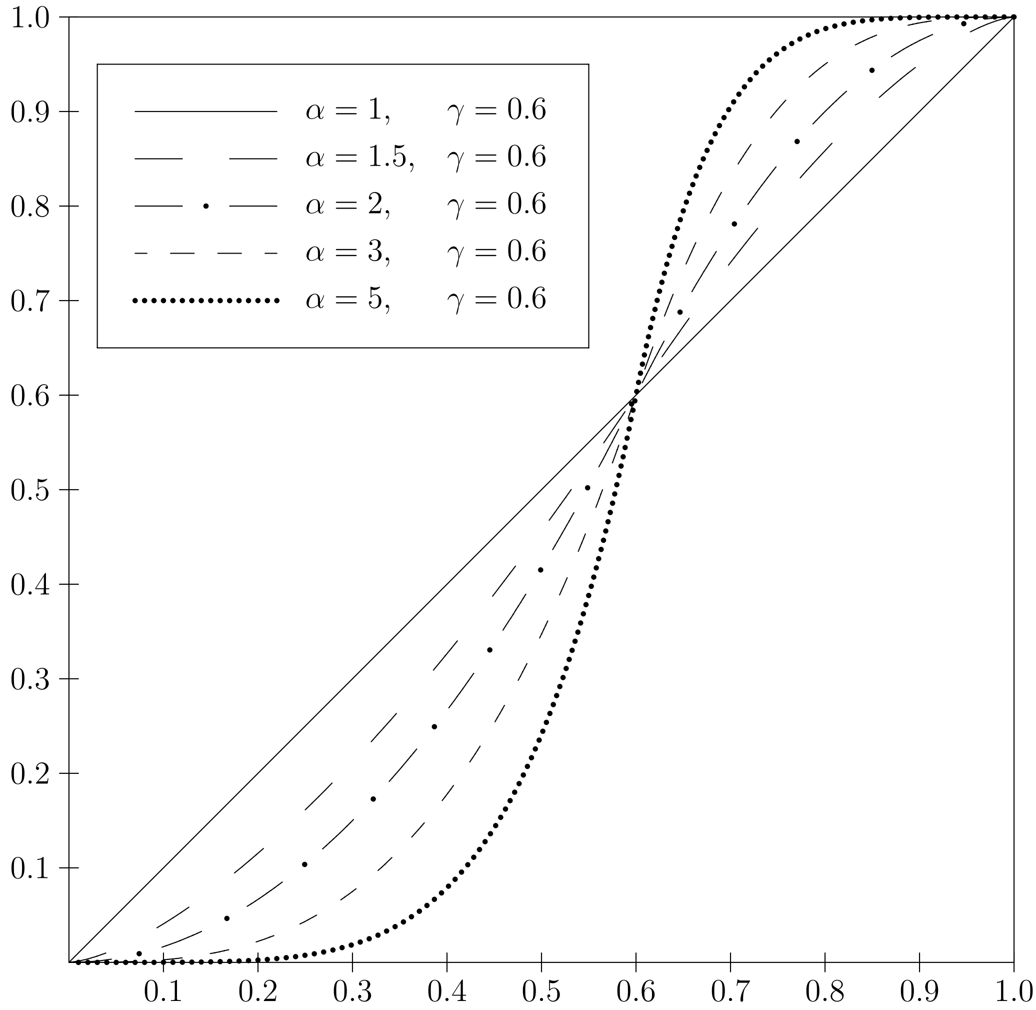
Eine Klasse Q_S von S-förmigen Verzerrungsfunktionen

Aufgrund der oben vorgebrachten Kritikpunkte an den von Tversky und Kahneman (1992) benutzten Verzerrungsfunktionen schlagen wir die folgende Klasse von Verzerrungsfunktionen vor. Für $\gamma \in [0,1]$ und $\alpha \in \mathbb{R}_+$ sei die Verzerrungsfunktion $q_{\alpha,\gamma}$ gegeben durch

$$q_{\alpha,\gamma}(p) = \begin{cases} \dfrac{p^\alpha}{\gamma^{\alpha-1}}, & \text{falls } p \leq \gamma, \\ 1 - \dfrac{(1-p)^\alpha}{(1-\gamma)^{\alpha-1}}, & \text{falls } p > \gamma \end{cases}.$$

Weiter sei $Q_S = \{q_{\alpha,\gamma} \,|\, \gamma \in [0,1],\, \alpha \in \mathbb{R}_+\}$, d.h. Q_S ist die Menge aller dieser Verzerrungsfunktionen. Für $\gamma \in (0,1)$ haben diese Verzerrungsfunktionen eine S-förmige Gestalt. Der Parameter γ bezeichnet den Schnittpunkt der Verzerrungsfunktion mit der Hauptdiagonalen, d.h. es ist $q_{\alpha,\gamma}(\gamma) = \gamma$. Für $\alpha \geq 1$ ist $q_{\alpha,\gamma}$ konvex in $[0,\gamma]$ und konkav in $[\gamma,1]$. Für $\alpha < 1$ ist $q_{\alpha,\gamma}$ konkav in $[0,\gamma]$ und konvex in $[\gamma,1]$. Der Parameter α bestimmt die Stärke der Krümmung der Verzerrungsfunktion.

Im Extremfall $\gamma = 1$ erhält man für $\alpha \geq 1$ konvexe und für $\alpha \leq 1$ konkave Verzerrungsfunktionen. In der folgenden Abbildung sind für einige Parameter die Verzerrungsfunktionen $q_{\alpha,\gamma}$ abgebildet.



Die Menge Q_S ist abgeschlossen gegenüber Dualisierung. Die zu $q_{\alpha,\gamma}$ duale Verzerrungsfunktion ist $q_{\alpha,1-\gamma}$.

Die Klasse Q_S erfüllt weitgehend die oben aufgestellten Forderungen. Die Klasse wird durch zwei Parameter beschrieben, die in einfacher Weise die Form der Verzerrungsfunktion beschreiben (γ den Schnittpunkt mit der Winkelhalbierenden, α die Stärke der Krümmung). Es sind sowohl konkave als auch konvexe Verzerrungsfunktionen enthalten, was die Beschreibung von sowohl risikoaversem als auch risikofreudigem Verhalten ermöglicht. Darüber hinaus sind aber auch S-förmige Funktionen enthalten, bei denen Schnittpunkt mit der Winkelhalbierenden und Krümmung der Verzerrungsfunktion unabhängig voneinander variiert werden können. Schließlich ist auch Abgeschlossenheit gegenüber Dualisierung erfüllt, wie oben dargelegt. Was die Handhabbarkeit dieser Verzerrungsfunktionen angeht, ist zwar nachteilig, daß die Funktionen nicht durch einen geschlossenen Ausdruck definiert sind. Beide Äste werden jedoch durch einfache Potenzfunktionen beschrieben.

Bestimmung einer Verzerrungsfunktion aus Q_S

Wir beschreiben, wie eine Verzerrungsfunktion in Q_S bestimmt werden kann. Dabei gehen wir wieder davon aus, daß ein AU-Modell vorliegt.

Wir wählen drei beliebige Konsequenzen x_*, x, x^* so, daß $x_* \prec x \prec x^*$ gilt und setzen $u(x_*) = 0$, $u(x^*) = 1$. Dann ist $u(x) = c \in (0, 1)$. Für $p_1, p_2 \in [0, 1]$ sei die Lotterie $L(p_1, p_2)$ definiert durch

$$L(p_1, p_2) = \begin{bmatrix} x_* & x & x^* \\ p_1 & 1 - p_1 - p_2 & p_2 \end{bmatrix}.$$

und es sei $\mathcal{L} = \{L(p_1, p_2) \,|\, p_1, p_2 \in [0, 1]\}$ die Menge aller dieser Lotterien. Für $L \in \mathcal{L}$ bezeichnen wir mit $AU_{\alpha,\gamma,c}(L)$ den antizipierten Nutzen der Lotterie L, wenn $q_{\alpha,\gamma}$ die Verzerrungsfunktion des ET und $u(x) = c$ ist.

Der ET muß eine gewisse Anzahl von Indifferenzen in der Klasse $\mathcal{L}$ angeben. Es liegen dann Paare $(L_{i,1}, L_{i,2})$, $i = 1, \ldots n$, von Lotterien in $\mathcal{L}$ vor, für die $L_{i,1} \sim L_{i,2}$, $i = 1 \ldots n$, gilt. Die Parameter α, γ und c werden dann nach der Methode der kleinsten Quadrate so bestimmt, daß die Summe der Fehlerquadrate

$$F(\alpha, \gamma, c) = \sum_{i=1}^{n} \left[AU_{\alpha,\gamma,c}(L_{i,1}) - AU_{\alpha,\gamma,c}(L_{i,2})\right]^2$$

minimal wird.

Der ET muß mindestens $n = 3$ Paare von Indifferenzen angeben, da die Parameter sonst nicht eindeutig bestimmt sind.

Indifferenzangaben zwischen Lotterien werden mit steigender Anzahl von möglichen Ergebnissen immer schwerer. Wir schlagen daher vor, nur mit Lotterien mit höchstens zwei möglichen Ergebnissen zu arbeiten. Es sind dann Indifferenzen der folgenden Formen möglich.

Typ 1: Bestimme p_0 so, daß $x \sim \begin{bmatrix} x_* & x^* \\ 1 - p_0 & p_0 \end{bmatrix}$ ist.

Typ 2: Wähle p_1 und p_2 so, daß $\begin{bmatrix} x_* & x \\ 1 - p_1 & p_1 \end{bmatrix} \sim \begin{bmatrix} x_* & x^* \\ 1 - p_2 & p_2 \end{bmatrix}$ gilt.

Typ 3: Wähle p_1 und p_2 so, daß $\begin{bmatrix} x & x^* \\ 1 - p_1 & p_1 \end{bmatrix} \sim \begin{bmatrix} x_* & x^* \\ 1 - p_2 & p_2 \end{bmatrix}$ gilt.

Während p_0 eindeutig bestimmt ist, sind beliebig viele Indifferenzangaben vom Typ 2 oder 3 möglich.

Das oben beschrieben Verfahren wurde auf einem PC implementiert. Der Benutzer hat die Möglichkeit beliebig viele Indifferenzen der Typen 1, 2 und 3 anzugeben. Das auftretende nichtlineare Minimierungsproblem wird mit dem BFGS-Algorithmus (siehe z.B. Stoer 1979, S. 278ff) gelöst. Der Algorithmus findet das Minimum i.a. in weniger als zehn Iterationen. Gute Ergebnisse erhält man i.a. bei 5 oder mehr Indifferenzangaben. Dabei ist jedoch darauf zu achten, daß die Anzahl der Indifferenzangaben von Typ 2 und Typ 3 etwa gleich groß ist.

Kapitel 8

AU und CEU in Anwendungen

In diesem Kapitel wollen wir an ausgewählten Anwendungssituationen zeigen, welche ökonomischen Implikationen sich aus der Verwendung von AU oder CEU ergeben. Insbesondere soll dargelegt werden, daß durch diese Modelle ein realistischeres Verhalten beschrieben werden kann als es durch EU möglich ist.

Im ersten Abschnitt, der Denneberg (1989) folgt, wird sich zeigen, daß das Choquet-Integral bezüglich eines verzerrten Wahrscheinlichkeitsmaßes zur Konstruktion von Prämienprinzipien für Versicherungen verwendet werden kann. Danach betrachten wir zwei klassische Grundprobleme der Portfeuilleauswahl, nämlich zum einen die Auswahl zwischen einer sicheren und einer risikobehafteten Anlageform, zum anderen das Problem der Aufteilung eines Betrages zwischen mehreren risikobehafteten Anlagen. Im letzten Abschnitt dieses Kapitels schließlich gehen wir noch auf eine Arbeit von Dow und Werlang (1992) ein, in der gezeigt wird, daß ein unsicherheitsaverser CEU-Maximierer genau dann mit eine Anlage handelt, wenn der Preis für diese Anlage außerhalb eines bestimmten Intervalls liegt. Liegt der Preis unterhalb des Intervalls, so kauft er Anteile dieser Anlage, liegt er oberhalb des Intervalls, so tätigt er Leerverkäufe. Die Größe dieses Intervalls hängt dabei vom Grad der Unsicherheitsaversion des ETs ab. Weitere Anwendungen von AU- und CEU-Modellen finden sich in Chateauneuf, Kast und Lapied (1992a,b) und Epstein und Wang (1992).

8.1 Prämienprinzipien für Versicherungen

Denneberg (1989) verwendet verzerrte Wahrscheinlichkeitsmaße zur Konstruktion von Prämienprinzipien für Versicherungen. Sei $\mathcal{X}$ eine Menge von Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$. Die Zufallsvariablen in der Menge $\mathcal{X}$ werden als die zu versichernden Risiken interpretiert. Ein *Prämienprinzip* ist eine Abbildung $H : \mathcal{X} \to \overline{\mathbb{R}}$. An ein Prämienprinzip werden i.a. die folgenden Anforderungen gestellt (vgl. hierzu Heilmann 1988, Reich 1985):

V1: $H(X) \geq E(X)$, falls $E(X)$ existiert,

V2: $H(X) \leq \sup_{\omega \in \Omega} X(\omega)$,

V3: $H(X + c) = H(X) + c$ für alle $c \in \mathbb{R}$,

V4: $H(cX) = cH(X)$ für alle $c \geq 0$.

Die Prämie soll also mindestens so hoch sein wie der erwartete Schaden (**V1**) jedoch nicht höher als der maximal mögliche Schaden (**V2**). **V3** besagt, daß durch eine sichere Erhöhung (Verringerung) des Schadens um c auch die Prämie um diesen Betrag erhöht (verringert) wird. Schließlich fordert **V4**, daß sich bei einer Multiplikation des Schadens mit einer positiven Konstante auch die Prämie entsprechend ändert.

Denneberg (1989) zeigt, daß das Choquet-Integral bezüglich eines verzerrten W-Maßes ein vernünftiges Prämienprinzip darstellt.

Satz 8.1 (Denneberg 1989) *Für die Verzerrungsfunktion q gelte $q(p) \geq p$, $0 \leq p \leq 1$. Dann ist*

$$H(X) = \int_{\Omega} X \, d(q \circ P)$$

ein Prämienprinzip, das die Eigenschaften **V1** - **V4** *erfüllt.*

Eine weitere sinnvolle Eigenschaft eines Prämienprinzips ist:

V5: $H(X + Y) \leq H(X) + H(Y)$.

Die Prämie für die Summe der Risiken $X+Y$ darf nicht größer sein als die Summe der Prämien der Einzelrisiken, da es sonst für den Versicherungsnehmer günstiger ist, die Risiken X und Y getrennt zu versichern.

Satz 8.2 (Denneberg 1989) *Sei q eine konkave Verzerrungsfunktion. Dann ist*

$$H(X) = \int_{\Omega} X \, d(q \circ P)$$

ein Prämienprinzip, das die Eigenschaften **V1**–**V5** *erfüllt.*

In diesem Kontext kann auch die komonotone Additivität des Choquet-Integrals schön interpretiert werden. Sind zwei Risiken X und Y komonoton, so bedeutet dies, daß kein Risikoausgleich zwischen X und Y stattfindet. Die Prämie ergibt sich dann als Summe der Prämien der Risiken X und Y. Sind jedoch die Risiken X und Y nicht komonoton, so findet ein Risikoausgleich statt. Die Prämie des Risikos $X + Y$ ist dann kleiner als die Summe der Einzelprämien.

8.2 Portfeuilleauswahl

8.2.1 Eine sichere und eine risikobehaftete Anlage

In diesem Abschnitt wollen wir AU auf ein einfaches Portfeuilleproblem anwenden, das in EU unter anderem von Arrow (1971, Essay 3) untersucht wurde. Der ET steht vor dem Problem, einen bestimmten Geldbetrag v zwischen einer sicheren und einer risikobehafteten Anlagemöglichkeit aufzuteilen. Der Ertrag der risikobehafteten Anlage pro investierter Geldeinheit werde durch eine Zufallsvariable $X \geq 0$ beschrieben, deren Verteilung bekannt sei. Der ET kann einen beliebigen Betrag a, $0 \leq a \leq v$ in die risikobehaftete Anlage investieren. Das Guthaben am Ende der Periode wird dann durch die Zufallsvariable

$$Y_a = v - a + aX$$

beschrieben. Zur Vereinfachung definiert man $R = X - 1$. Hierbei steht R für „rate of return". Damit ergibt sich

$$Y_a = v + aR.$$

Wir bezeichnen den optimalen Wert von a mit $a^\star$, d.h. es ist $Y_{a\star} \succeq Y_a$ für alle a mit $0 \leq a \leq v$. Agiert der ET gemäß EU, so gilt das folgende Ergebnis von Arrow (1971):

Sei u zweimal differenzierbar und konkav. Dann gelten die folgenden Aussagen:

(i) Ist $E(R) > 0$, so ist $a^\star > 0$.

(ii) Ist $E(R) \leq 0$, so ist $a^\star = 0$.

Mit anderen Worten: Ein risikoaverser EU-Maximierer investiert nichts in eine Anlage mit $E(R) \leq 0$. Ist jedoch $E(R) > 0$, so investiert der ET einen positiven Betrag in diese Anlage.

Bemerkenswert an obiger Aussage ist, daß ein risikoaverser EU-Maximierer bereits dann einen positiven Betrag in die Anlage investiert, wenn $E(R)$ nur geringfügig größer als Null ist. Im Gegensatz dazu kann in AU Verhalten modelliert werden, bei dem der ET erst dann in eine Anlage investiert, wenn deren erwartete Rückzahlungsrate $E(R)$ einen gewissen positiven Betrag erreicht.

Satz 8.3 *Für einen AU-Maximierer mit zweimal differenzierbarer, monoton wachsender und konkaver Nutzenfunktion u und stetiger Verzerrungsfunktion q ist $a^\star > 0$ genau dann, wenn*

$$\int R d(q \circ P) > 0$$

ist.

Beweis: Für $0 \leq a \leq v$ sind die Zufallsvariablen Y_a komonoton. Nach Satz 2.4 existiert ein W-Maß $\tilde{P}$, so daß

$$\int u(v + aR)d(q \circ P) = \int u(v + aR)d\tilde{P} \quad \text{für alle } a \in [0, v]$$

ist. Man kann also das Ergebnis von Arrow benutzen. Es ist $a^\star > 0$ genau dann, wenn $E_{\tilde{P}}(R) > 0$ ist. Wegen

$$E_{\tilde{P}}(R) = \int R d(q \circ P)$$

folgt daraus die Behauptung. ∎

Nach Satz 5.2 gelten die folgenden Aussagen über den Zusammenhang zwischen $E_{q \circ P}(R)$ und $E_P(R)$.

Ist $q(x) \leq x$ für alle $x \in [0, 1]$, so ist $E_{q \circ P}(R) \leq E_P(R)$ für alle Zufallsvariablen R.
Ist $q(x) \geq x$ für alle $x \in [0, 1]$, so ist $E_{q \circ P}(R) \geq E_P(R)$ für alle Zufallsvariablen R.

Bei konkaver Nutzenfunktion können also folgende Verhaltensweisen auftreten:

Ein pessimistischer ET (d.h. ein ET mit $q(x) \leq x$) investiert erst dann einen positiven Betrag a in die risikobehaftete Anlage, wenn der Erwartungswert $E(R)$ einen gewissen nichtnegativen Wert überschreitet.

Ein optimistischer ET (d.h. ein ET mit $q(x) \geq x$) investiert schon dann einen positiven Betrag a in die risikobehaftete Anlage, wenn der Erwartungswert $E(R)$ einen gewissen nichtpositiven Wert überschreitet.

8.2.2 Mehrere risikobehaftete Anlagen

Wir betrachten das folgende Portfeuilleproblem. Es stehen n risikobehaftete Anlagemöglichkeiten zur Auswahl, deren Erträge je investierter Geldeinheit durch die Zufallsvariablen $X_1, \ldots, X_n$ beschrieben werden. Ein Betrag k soll vollständig in diese Anlagen investiert werden. O.B.d.A. sei das zu investierende Kapital $k = 1$. Der Anteil des in die i-te Anlageform investierten Kapitals werde mit α_i bezeichnet. Es ist dann $\sum_{i=1}^n \alpha_i = 1$ und der Gesamtertrag des Portfeuilles ist $X = \sum_{i=1}^n \alpha_i X_i$. In Samuelson (1967) wurde gezeigt, daß im Rahmen von EU ein Investor sein Kapital zu gleichen Teilen auf die n Anlageformen aufteilt, sofern die Nutzenfunktion u konkav und die gemeinsame Verteilungsfunktion $F(x_1, \ldots, x_n) = P(X_1 \leq x_1, \ldots, X_n \leq x_n)$ symmetrisch in den Argumenten $x_1, \ldots x_n$ ist. Diese Ergebnis ist auch in AU gültig, wie wir zeigen werden.

Satz 8.4 *Seien $X_1, \ldots, X_n$ reelle nichtnegative Zufallsvariablen mit gemeinsamer Verteilungsfunktion F. F sei symmetrisch in den Argumenten. Dann gilt für* *jeden AU-Maximierer mit konkaver, zweimal differenzierbarer Nutzenfunktion u und konvexer, zweimal differenzierbarer Verzerrungsfunktion q:*

$$AU\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \geq AU\left(\sum_{i=1}^{n} \alpha_i X_i\right) \quad \textit{für alle } \alpha_1,\ldots,\alpha_n \geq 0 \textit{ mit } \textstyle\sum_{i=1}^n \alpha_i = 1.$$

Beweis: Für $\frac{1}{n}\sum_{i=1}^n X_i$ schreiben wir kürzer $\overline{X}$. Ist $\alpha = (\alpha_1, \ldots, \alpha_n)$ mit $\alpha_i \geq 0$, $i = 1, \ldots n$, und $\sum_{i=1}^n \alpha_i = 1$, so schreiben wir kürzer X_α anstelle von $\sum_{i=1}^n \alpha_i X_i$. Nach den klassischen Ergebnissen (siehe z.B. Rothschild und Stiglitz 1971) gilt für alle zweimal differenzierbaren, monoton wachsenden und konkaven Nutzenfunktionen u:

$$EU(\overline{X}) \geq EU(X_\alpha) \quad \text{für alle } \alpha_1, \ldots, \alpha_n \geq 0 \text{ mit } \textstyle\sum_{i=1}^n \alpha_i = 1.$$

Dies bedeutet aber, daß

$$P^{\overline{X}} \succeq_{\mathcal{U}_2} P^{X_\alpha}$$

gilt. Nach Satz 4.6 ist dann aber für alle zweimal differenzierbaren, konvexen Verzerrungsfunktionen q auch

$$q \circ P^{\overline{X}} \succeq_{\mathcal{U}_2} q \circ P^{X_\alpha},$$

was bedeutet, daß für alle solchen Verzerrungsfunktionen und alle zweimal differenzierbaren, konkaven Nutzenfunktionen

$$AU(\overline{X}) \geq AU(X_\alpha)$$

gilt. Das war aber zu beweisen. ∎

Ist die Nutzenfunktion u konkav und die Verzerrungsfunktion q konvex, so ist eine gleichmäßige Aufteilung des zu investierenden Betrages insbesondere dann optimal, wenn die n Anlagemöglichkeiten als unabhängige und identisch verteilte Zufallsvariablen angesehen werden können.

Ist q nicht mehr konvex, so kann selbst bei konkaver Nutzenfunktion und unabhängigen und identisch verteilten Anlagemöglichkeiten i.a. nicht mehr auf eine gleichmäßige Diversifikation geschlossen werden. Dies illustriert das folgende Beispiel.

Beispiel 8.1 (**u konkav, aber keine gleichmäßige Diversifikation**) *Wir betrachten zwei unabhängige und identisch verteilte Zufallsvariablen X_1 und X_2, die die Erträge zweier Anlagemöglichkeiten beschreiben. Die Verteilung von X_1 (und X_2) sei gegeben durch*

$$P^{X_1} = \begin{bmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Der ET habe die monoton wachsende und konkave Nutzenfunktion $u : [0,1] \to \mathbb{R}$, $u(x) = 1 - (1-x)^2$ *und die konkave Verzerrungsfunktion* $q = u$. *Die Verteilung von* $Z_\alpha = \alpha X_1 + (1-\alpha) X_2$ *ist gegeben durch*

$$P^{Z_\alpha} = \begin{bmatrix} 0 & \alpha & 1-\alpha & 1 \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

und der antizipierte Nutzen dieser Lotterie für $0 \le \alpha \le \frac{1}{2}$ *durch*

$$\begin{aligned} AU(Z_\alpha) = u(0) \; & + \; [u(\alpha) - u(0)] q(\tfrac{3}{4}) \\ & + \; [u(1-\alpha) - u(\alpha)] q(\tfrac{1}{2}) \\ & + \; [u(1) - u(1-\alpha)] q(\tfrac{1}{4})\,. \end{aligned}$$

bzw.

$$AU(Z_\alpha) = 12 + 6\alpha - 8\alpha^2$$

Man errechnet leicht, daß $AU(Z_\alpha)$ *sein Maximum für* $\alpha = \frac{3}{8}$ *annimmt. Also ist in der oben beschriebenen Situation eine Investition von* $3/8$ *des Betrages in Anlage* 1 *und* $5/8$ *des Betrages in Anlage* 2 *(oder auch umgekehrt) optimal.*

Wir zitieren ein weiteres Ergebnis aus EU (Samuelson 1967):

Ist $E(X_1) \ge E(X_j)$, $j = 1, \ldots, n$, und ist X_1 unabhängig von $(X_2, \ldots, X_n)$, so investiert jeder risikoaverse EU-Maximierer (d.h. jeder EU-Maximierer mit konkaver Nutzenfunktion) einen positiven Betrag in die Anlage X_1.

Diese Aussage ist in AU i.a. nicht mehr gültig, d.h. selbst wenn eine der zur Auswahl stehenden Anlagen unabhängig von den anderen Anlagen ist und einen Erwartungswert hat, der echt größer ist als der aller anderen Anlagen, ist nicht gewährleistet, daß ein stark risikoaverser AU-ET (d.h. ein AU-Maximierer mit konkaver Nutzen- und konvexer Verzerrungsfunktion) einen positiven Betrag in diese Anlage investiert. Dies wird deutlich wenn man den Extremfall betrachtet, in dem $q = 1_{\{1\}}$ ist. In diesem Fall agiert der AU-Maximierer nach dem Maximin-Prinzip. Stehen dann zwei stochastisch unabhängige Anlagen zur Verfügung, so investiert der ET den vollen Betrag in die Anlage, deren schlechtestes Ergebnis am günstigsten ist. Aber nicht nur in diesem Extremfall, sondern auch bei stetiger Verzerrungsfunktion ist das eben beschriebene Verhalten möglich.

Beispiel 8.2 (Keine Diversifikation) *Wir betrachten einen AU-ET mit Nutzenfunktion* $u(x) = -e^{-x}$ *und Verzerrungsfunktion* $q(x) = x^5$. *Betrachtet werden die beiden stochastisch unabhängigen Anlagen*

$$P^X = \begin{bmatrix} -2 & 2 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \quad \text{und} \quad P^Y = \begin{bmatrix} -1 & 1 \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Dann ist $E(X_1) = 0 = E(X_2)$. *Der ET ist risikoavers. Für* $\alpha \in [0,1]$ *ist die Verteilung von* $Z_\alpha = \alpha X + (1-\alpha) Y$ *gegeben durch*

$$P^{Z_\alpha} = \begin{bmatrix} -1-\alpha & -1+3\alpha & 1-3\alpha & 1+\alpha \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

Für $0 \leq \alpha \leq \frac{1}{3}$ *ist*

$$AU(Z_\alpha) = -e^{1+\alpha}\frac{781}{1024} - e^{1-3\alpha}\frac{211}{1024} - e^{-1+3\alpha}\frac{31}{1024} - e^{-1-\alpha}\frac{1}{1024}.$$

Wir betrachten die Funktion $f : \alpha \mapsto AU(Z_\alpha)$. *Man rechnet leicht nach, daß*

$$f'(\alpha) = -\frac{1}{1024}\left(781e^{1+\alpha} - 633e^{1-3\alpha} + 93e^{-1+3\alpha} - e^{-1-\alpha}\right)$$

und

$$f''(\alpha) = -\frac{1}{1024}\left(781e^{1+\alpha} + 1899e^{1-3\alpha} + 279e^{-1+3\alpha} + e^{-1-\alpha}\right) < 0$$

ist. Es ist

$$f'(0) = -\frac{1}{1024}\left(148e + 92e^{-1}\right) < 0.$$

Also ist f *in* $[0, \frac{1}{3}]$ *streng monoton fallend und konkav. Genauso zeigt man, daß* f *auch in* $[\frac{1}{3}, 1]$ *monoton fallend ist. Da* f *stetig ist, ist* f *somit in* $[0, 1]$ *monoton fallend. Das Maximum wird daher für* $\alpha = 0$ *angenommen. Das bedeutet aber, daß der Investor sein ganzes Kapital in Anlage* Y *investiert.*

8.3 Kauf- und Verkaufspreise

Dow und Werlang (1992) betrachten das Problem aus Abschnitt 8.2.1 unter den Annahmen, daß keine objektiven Wahrscheinlichkeiten gegeben sind und daß Leerverkäufe der Anlage möglich sind.

Das SEU-Modell schreibt unter diesen Anahmen für einen risikoaversen oder risikoneutralen ET (d.h. einen ET mit konkaver Nutzenfunktion) das folgende Verhalten vor.

- Ist $p < E(X)$, so investiert der ET einen positiven Betrag in die Anlage.
- Ist $p > E(X)$, so verkauft der ET einen positiven Betrag der Anlage leer.

Der ET wird also bis auf den Fall $p = E(X)$ in jedem Fall mit der Anlage handeln, sei es, daß er Anteile erwirbt oder daß er Anteile leerverkauft.

Nach Dow und Werlang (1992) beobachtet man in der Praxis ein anderes Verhalten. Der Händler investiert in eine Anlage, falls deren Preis unter einer gewissen Schranke c_B liegt und tätigt Leerverkäufe der Anlage, sobald deren Preis eine zweite Schranke c_S übersteigt, wobei i.a. $c_B < c_S$ ist. Es existiert somit ein Intervall derart, daß der Händler für Preise in diesem Intervall Anteile weder kauft nocht verkauft. SEU fordert jedoch, daß $c_B = c_S$ ist, d.h. daß beide Schranken identisch sind. Dow und Werlang (1992) erklären das in der Praxis beobachtete Verhalten durch die Annahme, daß der ET ein unsicherheitsaverser CEU-Maximierer ist. Der Kauf eines Anteils der Anlage wird durch die Zufallsvariable $X - p$ beschrieben, der Leerverkauf eines Anteils durch die Zufallsvariable $p - X$. Der subjektiv erwartete Gewinn aus dem Kauf eines Anteils ist dann für einen CEU-Maximierer durch $E_\mu(X) - p$ gegeben, der Gewinn aus einem Leerverkauf durch $p + E_\mu(-X)$. Man wird also erwarten, daß für $p < E_\mu(X)$ die Anlage gekauft, für $p > -E_\mu(-X)$ verkauft wird. Zunächst muß man sich fragen, unter welchen Bedingungen $-E_\mu(-X) \geq E_\mu(X)$ ist. Das folgende Lemma liefert eine hinreichende Bedingung hierfür.

Lemma 8.1 *Sei μ eine stark superadditive Kapazität auf einem Meßraum $(\Omega, \mathcal{A})$. Dann gilt für alle μ-integrierbaren Funktionen $X : \Omega \to \mathbb{R}$*

$$E_\mu(X) \leq -E_\mu(-X)$$

Beweis: Nach dem Subadditivitäts-Theorem (Satz 2.1) ist $0 = E_\mu(X + (-X)) \geq E_\mu(X) + E_\mu(-X)$, was wiederum zu $-E_\mu(-X) \geq E_\mu(X)$ äquivalent ist. ■

Der folgende Satz präzisiert nun die oben formulierte intuitive Vorstellung über das Verhalten eines CEU-Maximierers.

Satz 8.5 (Dow und Werlang 1992) *In einem CEU-Modell sei u zweimal differenzierbar, $u' \geq 0$, $u'' \leq 0$ und μ stark superadditiv. Dann kauft der ET die Anlage, wenn $p < E_\mu(X)$ ist, und nur dann, wenn $p \leq E_\mu(X)$ ist. Der ET verkauft die Anlage, wenn $p > -E_\mu(-X)$ ist, und nur dann, wenn $p \geq -E_\mu(-X)$ ist.*

Anhang A

Mathematische Grundbegriffe

A.1 Ordnungsrelationen

Eine binäre Relation R auf einer Menge X heißt

- *reflexiv*, wenn für alle $x \in X$ gilt xRx,
- *symmetrisch*, wenn für alle $x, y \in X$ gilt $xRy \Rightarrow yRx$,
- *transitiv*, wenn für alle $x, y, z \in X$ gilt $xRy \wedge yRz \Rightarrow xRz$,
- *antisymmetrisch*, wenn für alle $x, y \in X$ gilt $xRy \wedge yRx \Rightarrow x = y$,
- *vollständig*, wenn für alle $x, y \in X$ gilt $xRy \vee yRx$.

Eine reflexive, symmetrische und transitive Relation heißt *Äquivalenzrelation*. Ist $x \in X$ und R eine Äquivalenzrelation, so bezeichnen wir mit x^* die Menge

$$x^* = \{y \in X \mid xRy\}.$$

x^* heißt die zu x gehörige Äquivalenzklasse. Die Äquivalenzklassen bilden eine Zerlegung der Menge X, d.h. jedes Element $x \in X$ ist in genau einer Äquivalenzklasse enthalten. Die Menge aller Äquivalenzklassen von X bezeichnen wir mit X^*.

Eine Relation R auf einer Menge X heißt

- *Präordnung*, wenn sie reflexiv und transitiv ist,
- *partielle Ordnung*, wenn sie reflexiv, transitiv und antisymmetrisch ist,
- *schwache Ordnung*, wenn sie transitiv und vollständig ist,
- *(lineare) Ordnung*, wenn sie transitiv, vollständig und antisymmetrisch ist.

Für diese Relationen benutzt man oft die Symbole $\preceq$ oder $\leq$. Ist eine Relation vollständig, so ist sie auch reflexiv. Insbesondere sind also auch schwache Ordnungen und lineare Ordnungen reflexiv.

Ist $\preceq$ eine schwache Ordnung, so sind die Relationen $\succeq$, $\prec$, $\succ$ und $\sim$ wie folgt definiert:

- $x \succeq y \ :\Longleftrightarrow \ y \preceq x$,
- $x \prec y \ :\Longleftrightarrow \ x \preceq y \wedge \neg y \preceq x$,
- $x \succ y \ :\Longleftrightarrow \ y \prec x$,
- $x \sim y \ :\Longleftrightarrow \ x \preceq y \wedge y \preceq x$.

Die Relation $\sim$ ist eine Äquivalenzrelation. Definiert man auf der Menge X^* aller Äquivalenzklassen eine Relation durch

$$x^* \preceq^* y^* \ :\Longleftrightarrow \ x \preceq y,$$

so ist $\preceq^*$ wohldefiniert und eine lineare Ordnung auf X^*.

Sind $(X_i, \preceq_i)$, $i = 1, 2$, linear geordnete Räume und definiert man auf dem Kartesischen Produkt $X_1 \times X_2$ eine Relation $\preceq$ durch

$$(x_1, x_2) \preceq (y_1, y_2) \ :\Longleftrightarrow \ x_1 \preceq_1 y_1 \text{ und } x_2 \preceq_2 y_2,$$

so ist $\preceq$ eine partielle Ordnung auf $X_1 \times X_2$. $\preceq$ heißt die *von $\preceq_1$ und $\preceq_2$ induzierte partielle Ordnung auf* $X_1 \times X_2$.

Seien $(X, \preceq)$, $(X_i, \preceq_i)$, $i = 1, 2$, mit einer Präordnung versehene Räume.

- Eine Funktion $f : X_1 \to X_2$ heißt *isoton*, wenn für alle $x, y \in X_1$ gilt $x \preceq_1 y \Rightarrow f(x) \preceq_2 f(y)$.
- Eine Funktion $f : X_1 \to X_2$ heißt *antiton*, wenn für alle $x, y \in X_1$ gilt $x \preceq_1 y \Rightarrow f(y) \preceq_2 f(x)$.
- Eine Menge $A \subset X$ heißt *obere Menge*, wenn gilt $x \in A, x \preceq y \Rightarrow y \in X$.
- Eine Menge $A \subset X$ heißt *untere Menge*, wenn gilt $x \in A, y \preceq x \Rightarrow y \in X$.

Sei $\preceq$ eine binäre Relation auf einer Menge X. Eine Funktion $u : X \to \mathbb{R}$ *repräsentiert* $\preceq$, wenn gilt

$$x \preceq y \quad \Longleftrightarrow \quad u(x) \leq u(y).$$

Sei $(X, \preceq)$ ein linear geordneter Raum, $x \in X$ und $A \subset X$.

- x heißt *obere (untere) Schranke von A*, wenn $x \succeq a$ $(x \preceq a)$ für alle $a \in A$ gilt.

- x heißt *Supremum (Infimum) von* A, wenn x obere (untere) Schranke von A ist und für jede andere obere (untere) Schranke y von A gilt $x \preceq y$ ($x \succeq y$). Man schreibt dann $x = \sup A$ bzw. $x = \inf A$.
- x heißt *Maximum (Minimum) von* A, wenn x Supremum (Infimum) von A und $x \in A$ ist. Man schreibt dann $x = \max A$ bzw. $x = \min A$.
- A heißt beschränkt, wenn A sowohl eine untere als auch eine obere Schranke besitzt.

Für einen linear geordneten Raum $(X, \preceq)$ verwenden wir die folgende Notation: Hat X ein Maximum (Minimum), so bezeichnen wir dieses mit ∞ ($-\infty$). Für $a, b \in X$ ist $(a, b] = \{x \in X \,|\, a \prec x \preceq b\}$. Die Bezeichnungen $[a, b]$, $[a, b)$ und (a, b) sind analog zu verstehen. Hat X kein Maximum, so ist $(a, \infty] = (a, \infty) = \{x \in X \,|\, a \prec x\}$ und $[a, \infty]$, $[a, \infty)$ sind entsprechend definiert. Die Bezeichnungen $[-\infty, a)$, $(-\infty, a)$, $[-\infty, a]$ und $(-\infty, a]$ für Räume ohne Minimum sind analog zu interpretieren.

A.2 Maßtheorie

Sei Ω eine Menge. Die Menge aller Teilmengen von Ω heißt die *Potenzmenge* von Ω. Wir bezeichnen sie mit $\mathcal{P}(\Omega)$. Eine Familie $\mathcal{A}$ von Teilmengen von Ω heißt *Algebra* wenn sie die folgenden Eigenschaften besitzt:

A1: $\emptyset, \Omega \in \mathcal{A}$,

A2: $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$,

A3: $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

Insbesondere enthält eine Algebra $\mathcal{A}$ mit je zwei Mengen A, B auch deren Durchschnitt $A \cap B$. $\mathcal{A}$ heißt *σ-Algebra* wenn an Stelle von **A3** die stärkere Eigenschaft **A3′** gilt:

A3′: Für jede Folge $(A_n)_{n \in \mathbb{N}}$ von Mengen aus $\mathcal{A}$ liegt $\bigcup_{n=1}^{\infty} A_n$ in $\mathcal{A}$.

Es sei $(A_n)_{n \in \mathbb{N}}$ eine Folge von Mengen. Wir schreiben

- $A_n \nearrow$, wenn $A_1 \subset A_2 \subset \ldots$ ist,
- $A_n \searrow$, wenn $A_1 \supset A_2 \supset \ldots$ ist,
- $A_n \nearrow A$, wenn $A_n \nearrow$ und $A = \bigcup_{n=1}^{\infty} A_n$ ist,
- $A_n \searrow A$, wenn $A_n \searrow$ und $A = \bigcap_{n=1}^{\infty} A_n$ ist.

Ist $A_n \nearrow$, so nennen wir die Folge aufsteigend, ist $A_n \searrow$, so nennen wir sie absteigend. Ist $A_n \nearrow$ oder $A_n \searrow$, so schreiben wir auch $\lim_{n\to\infty} A_n = A$.

Sei Ω eine Menge. Ein Familie $\mathcal{A}$ von Teilmengen von Ω heißt *monotone Klasse*, wenn die Vereinigung jeder aufsteigenden Folge und der Durchschnitt jeder absteigenden Folge von Mengen in $\mathcal{A}$ wieder in $\mathcal{A}$ liegt. Insbesondere ist jede σ-Algebra eine monotone Klasse.

Ist Ω eine Menge und $\mathcal{A}$ eine σ-Algebra auf Ω, so nennt man $(\Omega, \mathcal{A})$ einen *Meßraum.*

Wir bezeichnen jede Funktion $\mu : \mathcal{A} \to \overline{\mathbb{R}}$ mit $\mu(\emptyset) = 0$ als Mengenfunktion. Eine Mengenfunktion μ heißt

- *normiert*, falls $\mu(\Omega) = 1$,
- *nichtnegativ*, falls $\mu(A) \geq 0$ für alle $A \in \mathcal{A}$,
- *monoton*, falls für alle $A, B \in \mathcal{A}$ gilt: $A \subset B \Rightarrow \mu(A) \leq \mu(B)$,
- *endlich additiv*, falls für alle $A, B \in \mathcal{A}$ mit $A \cap B = \emptyset$ und $A \cup B \in \mathcal{A}$ gilt $\mu(A \cup B) = \mu(A) + \mu(B)$,
- *σ-additiv*, falls für jede Folge $(A_n)_{n\in\mathbb{N}}$ paarweise disjunkter Mengen aus $\mathcal{A}$ mit $\cup_{n=1}^{\infty} A_n \in \mathcal{A}$ gilt: $\mu(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$.

Eine nichtnegative, endich additive Mengenfunktion μ auf einer Algebra $\mathcal{A}$ heißt *Inhalt*. Ein normierter Inhalt heißt *endlich additives Wahrscheinlichkeitsmaß*. Eine nichtnegative, σ-additive Mengenfunktion auf einer σ-Algebra $\mathcal{A}$ heißt *Maß*. Ein normiertes Maß heißt *Wahrscheinlichkeitsmaß* oder kurz *W-Maß*. Ist P ein Wahrscheinlichkeitsmaß und $(\Omega, \mathcal{A})$ ein Meßraum, so heißt das Tripel $(\Omega, \mathcal{A}, P)$ *Wahrscheinlichkeitsraum* oder kurz *W-Raum*.

Die kleinste σ-Algebra auf $\mathbb{R}^n$, die alle offenen Mengen enthält heißt die *Borelsche σ-Algebra auf* $\mathbb{R}^n$. Wir bezeichnen sie mit $\mathcal{B}^n$. Statt $\mathcal{B}^1$ schreibt man auch $\mathcal{B}$. $\mathcal{B}$ ist auch die kleinste σ-Algebra auf $\mathbb{R}$, die alle Intervalle enthält.

Für $i = 1, 2$ sei $\mathcal{A}_i$ eine Algebra auf Ω_i und $f : \Omega_1 \to \Omega_2$. f heißt $\mathcal{A}_1 - \mathcal{A}_2$*-meßbar*, wenn für jedes $A \in \mathcal{A}_2$ die Menge $f^{-1}(A) = \{\omega \in \Omega_1 | f(\omega) \in A\}$ in $\mathcal{A}_1$ liegt. Ist klar, welche Algebren gemeint sind, so sagt man auch kürzer f ist *meßbar*. Ist μ eine Mengenfunktion auf $(\Omega_1, \mathcal{A}_1)$ und ist f $\mathcal{A}_1 - \mathcal{A}_2$*-meßbar*, so wird durch $\mu^f(A) = \mu(f^{-1}(A))$ eine Mengenfunktion auf $(\Omega_2, \mathcal{A}_2)$ definiert. μ^f heißt das *Bild von μ unter der Abbildung f*.

Sei Ω eine Menge und $A \subset \Omega$. Die Funktion

$$1_A(\omega) = \begin{cases} 1, & \text{falls } \omega \in A, \\ 0, & \text{falls } \omega \notin A, \end{cases}$$

heißt die *Indikatorfunktion* von A.

A.3 Der Satz von Daniell-Stone

Sei Ω eine beliebige Menge und $\mathcal{F}$ eine Menge reeller Funktionen auf Ω. Ist $\mathcal{F}$ ein reeller Vektorraum und enthält $\mathcal{F}$ mit je zwei Funktionen f und g auch die Funktionen $\min\{f, g\}$ und $\max\{f, g\}$, so heißt $\mathcal{F}$ *Vektorverband reeller Funktionen*. Diese Bedingung ist insbesondere dann erfüllt, wenn $\mathcal{F}$ mit jeder Funktion f auch die durch $\omega \mapsto |f(\omega)|$ definierte Funktion $|f|$ enthält. Mit $\mathcal{F}_+$ bezeichnen wir die Menge der Funktionen $f \in \mathcal{F}$ mit $f \geq 0$. Mit $\mathcal{A}_{\mathcal{F}}$ bezeichnen wir die kleinste σ-Algebra auf Ω, bezüglich der alle $f \in \mathcal{F}$ meßbar sind.

Sei $\mathcal{F}$ ein Vektorverband und L eine Linearform auf $\mathcal{F}$ (d.h. eine lineare Abbildung $L : \mathcal{F} \to \mathbb{R}$). L heißt positiv, wenn $L(f) \geq 0$ für alle $f \in \mathcal{F}_+$.

Satz A.1 (Daniell-Stone) *Sei Ω eine Menge, $\mathcal{F}$ ein Vektorverband reeller Funktionen, der die konstanten Funktionen enthält, und L eine positive Linearform auf $\mathcal{F}$. Dann sind die folgenden beiden Aussagen äquivalent.*

(i) Es gibt genau ein Maß μ auf $\mathcal{A}_{\mathcal{F}}$ mit

$$L(f) = \int f \, d\mu \quad \text{für alle } f \in \mathcal{F}.$$

(ii) Für jede absteigende Folge $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ mit $f_n \searrow 0$ gilt

$$\lim_{n\to\infty} L(f_n) = 0.$$

Beweis: Siehe z.B. Ash (1972), S. 175.

A.4 Der Satz von Hahn-Banach

Sei X ein reeller Vektorraum. Eine Funktion $l : X \to \mathbb{R}$ heißt *sublinear*, falls $l(\alpha x) = \alpha l(x)$ für alle $\alpha \in \mathbb{R}_+$, $x \in X$, und $l(x + y) \leq l(x) + l(y)$ für alle $x, y \in X$.

Satz A.2 (Hahn-Banach) *Sei X ein reeller Vektorraum und $p : X \to \mathbb{R}$ sublinear. Sei Y ein Teilraum von X und $l : Y \to \mathbb{R}$ eine lineare Abbildung mit $l(x) \leq p(x)$ für alle $x \in y$. Dann gibt es eine lineare Abbildung $L : X \to \mathbb{R}$ mit $L(x) = l(x)$ für alle $x \in Y$ und $L(x) \leq p(x)$ für alle $x \in X$.*

Beweis: Siehe z.B. Ash (1972), S. 139.

Anhang B

Axiomatisierung von CEU

In diesem Abschnitt geben wir eine kurze Beschreibung von vier verschiedenen Axiomatisierungen von CEU-Modellen. Zur Formulierung der Axiome benutzen wir die folgende Notation. Sind f, g Aktionen und ist $A \subset S$, so bezeichnen wir mit $f_A g_{-A}$ die Aktion, für die gilt

$$f_A g_{-A}(s) = \begin{cases} f(s), & \text{falls } s \in A, \\ g(s), & \text{falls } s \in A^c. \end{cases}$$

Wir identifizieren die Konsequenz $x \in \mathcal{C}$ mit der Aktion, die in jedem Zustand die Konsequenz x ergibt. Notationen wie $x_A y_{-A}$, $f_{-A} x_A$, $f_{-A} f(s)_A$ oder $f_{-A} x_{A-B} y_B$, wobei $f \in \mathcal{F}$ und $x, y \in \mathcal{C}$, sind daher in analoger Weise zu verstehen.

B.1 Ein Anscombe-Aumann-Ansatz: Schmeidler (1989)

Es sei $((S, \mathcal{A}), (\mathcal{C}, \mathcal{D}), \mathcal{F}_b, \preceq)$ ein Entscheidungsproblem unter Unsicherheit. Der Konsequenzenraum $\mathcal{C}$ sei die Menge aller Lotterien mit endlichem Träger auf einem Raum Γ. $\mathcal{F}_b$ sei die Menge aller beschränkten[1], $\mathcal{A}$-meßbaren Funktionen, $\mathcal{F}_0$ die Menge aller $\mathcal{A}$-meßbaren Funktionen mit endlichem Wertebereich. Für $u : \Gamma \to \mathbb{R}$ sei $EU_u(x)$ der Erwartungsnutzen der Lotterie $x \in \mathcal{C}$.

S1: Schwache Ordnung. Die Relation $\preceq$ ist eine schwache Ordnung auf $\mathcal{F}_0$.

S2: Komonotone Unabhängigkeit. Für alle komonotonen[2] Aktionen f, g und h in $\mathcal{F}_0$ und für alle $\alpha \in (0,1)$ gilt:

$$f \prec g \;\Rightarrow\; \alpha f + (1-\alpha)h \prec \alpha g + (1-\alpha)h\,.$$

[1]Eine Aktion $f : S \to \mathcal{C}$ heißt beschränkt, wenn es $x, y \in \mathcal{C}$ gibt, so daß $x \preceq f(s) \preceq y$ für alle $s \in S$.

[2]Zur Definition von Komonotonie siehe Definition 2.5.

S3: Stetigkeit. Für alle f, g und h in $\mathcal{F}_0$ gilt: Ist $f \prec g \prec h$, so gibt es $\alpha, \beta \in (0,1)$, so daß $\alpha f + (1-\alpha)h \prec g \prec \beta f + (1-\beta)h$.

S4: Monotonie. Für alle f und g in $\mathcal{F}_0$ gilt: Ist $f(s) \preceq g(s)$ für alle $s \in S$, so ist $f \preceq g$.

S5: Nichttrivialität. Es gibt f, g in $\mathcal{F}$ mit $f \prec g$.

Satz B.1 (Schmeidler 1989) *Erfüllt eine Relation $\preceq$ die Axiome* **S1**–**S5***, so gibt es eine Funktion $u : \Gamma \to \mathbb{R}$ und eine Choquet-Kapazität μ auf $(S, \mathcal{A})$, so daß $\preceq$ durch*

$$f \mapsto \int_S EU_u(f) d\mu$$

repräsentiert wird. Dabei ist μ eindeutig. u ist eindeutig bis auf positive affine Transformationen.

B.2 Ein Savage-Ansatz: Gilboa (1987)

Es sei $((S, \mathcal{P}(S)), (\mathcal{C}, \mathcal{D}), \mathcal{F}_b, \preceq)$ ein Entscheidungsproblem unter Unsicherheit.

Definition B.1 (f-konvex) *Sei $f \in \mathcal{F}$. Ein Ereignis $A \subset S$ heißt f-*konvex*, wenn gilt: Ist $f(r) \prec f(s) \prec f(t)$ und sind $r, t \in A$, so ist auch $s \in A$.*

G1: $\preceq$ ist eine schwache Ordnung.

G2: Für alle $A, B \subset S$, $f, g, u, v \in \mathcal{F}$, $\alpha, \beta, \gamma, \delta \in \mathcal{C}$, $\alpha \prec \gamma$, $\beta \prec \delta$, für die sowohl $f_{-A}\alpha_A$, $f_{-A}\gamma_A$, $g_{-A}\beta_A$, $g_{-A}\delta_A$ als auch $u_{-B}\alpha_B$, $u_{-B}\gamma_B$, $v_{-B}\beta_B$, $v_{-B}\delta_B$ jeweils paarweise komonoton sind, gilt

$$\begin{array}{lll} & f_{-A}\alpha_A \sim u_{-B}\alpha_B & \wedge \quad g_{-A}\beta_A \preceq v_{-B}\beta_B \\ \wedge & f_{-A}\gamma_A \preceq u_{-B}\gamma_B & \\ \Rightarrow & & \quad\; g_{-A}\delta_A \preceq v_{-B}\delta_B \end{array}$$

G3: Für alle $x, y \in \mathcal{C}$, $f \in \mathcal{F}$, $A \subset S$ gilt: Ist $x \prec y$, so ist auch $f_{-A}x_A \preceq f_{-A}y_A$.

G4: Es gibt $x, y, z \in \mathcal{C}$ mit $x \prec y \prec z$.

G5: Für alle $x, y \in \mathcal{C}$, $f, g \in \mathcal{F}$ und $A \subset S$ gilt: Sind $f_{-A}x_A$ und $f_{-A}y_A$ komonoton und ist $f_{-A}x_A \prec g \prec f_{-A}y_A$, so gibt es $B \subset A$, so daß $f_{-A}x_{A-B}y_B \sim g$.

G6: Seien $(f_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{F}$, $x, y \in \mathcal{C}$, $y \prec x$ und $A \subset S$. Ist $f_n(s) \prec y$ für alle $n \in \mathbb{N}$ und alle $s \in S$ und ist $f_{n+1} \sim (f_n)_{-A}x_A$ für alle $n \in \mathbb{N}$, so ist A ein Null-Ereignis.

G7: Für alle $f, g \in \mathcal{F}$ und alle f-konvexen Ereignisse $A \subset S$ gilt:
Ist $f_{-A}f(s)_A \succeq g$ für alle $s \in A$, so ist $f \succ g$.
Ist $f_{-A}f(s)_A \preceq g$ für alle $s \in A$, so ist $f \prec g$.

Satz B.2 (Gilboa 1987) *Die folgenden Aussagen sind äquivalent:*

(i) Die Relation $\preceq$ erfüllt die Axiome **G1**–**G7**.

(ii) Es gibt eine beschränkte Funktion $u : \mathcal{C} \to \mathbb{R}$ und eine Choquet-Kapazität μ auf $\mathcal{P}(S)$, so daß $\preceq$ durch

$$f \mapsto \int_S (u \circ f) d\mu$$

repräsentiert wird.

Dabei ist μ eindeutig. u ist eindeutig bis auf positive affine Transformationen.

B.3 Ein topologischer Ansatz: Wakker (1989)

Es sei $((S, \mathcal{P}(S)), (\mathcal{C}, \mathcal{D}), \mathcal{F}, \preceq)$ ein Entscheidungsproblem unter Unsicherheit. Der Zustandsraum S sei endlich. Der Konsequenzenraum $\mathcal{C}$ sei ein zusammenhängender, separabler topologischer Raum. $\mathcal{F} = \mathcal{C}^S$ die Menge aller Funktionen von S nach $\mathcal{C}$. $\mathcal{F}$ sei mit der Produkttopologie versehen.

Eine binäre Relation $\preceq$ auf einem topologischen Raum $\mathcal{F}$ heißt *stetig*, wenn für alle $f \in \mathcal{F}$ die Mengen $\{g \in \mathcal{F} \mid g \prec f\}$ und $\{g \in \mathcal{F} \mid g \succ f\}$ offen sind.

W1: Schwache Ordnung. $\preceq$ ist eine schwache Ordnung.

W2: Stetigkeit. $\preceq$ ist stetig.

W3: Nichtdegeneriertheit. Es gibt ein Ereignis, das weder null noch universell ist.

W4: Komonotone Tradeoff-Konsistenz. Für alle $s, t \in S$, $f, g, u, v \in \mathcal{F}$, $\alpha, \beta, \gamma, \delta \in \mathcal{C}$, für die sowohl $f_{-s}\alpha_s$, $g_{-s}\beta_s$, $f_{-s}\gamma_s$, $g_{-s}\delta_s$ als auch $u_{-t}\alpha_t$, $v_{-t}\beta_t$, $u_{-t}\gamma_t$, $v_{-t}\delta_t$ jeweils paarweise komonoton sind, gilt

$$\begin{array}{rll} & f_{-s}\alpha_s \preceq g_{-s}\beta_s \quad \wedge & u_{-t}\alpha_t \succeq v_{-t}\beta_t \\ \wedge & f_{-s}\gamma_s \succeq g_{-s}\delta_s & \\ \Rightarrow & & u_{-t}\gamma_t \succeq v_{-t}\delta_t \end{array}$$

Satz B.3 (Wakker 1989a,b) *Die folgenden Aussagen sind äquivalent:*

(i) Die Relation $\preceq$ erfüllt die Axiome **W1**–**W4**.

(ii) Es gibt eine stetige Funktion $u : \mathcal{C} \to \mathbb{R}$ und eine Choquet-Kapazität μ auf $\mathcal{P}(S)$, so daß $\preceq$ durch

$$f \mapsto \int_S (u \circ f) d\mu$$

repräsentiert wird.

Dabei ist in (ii) μ eindeutig. u ist eindeutig bis auf positive affine Transformationen.

B.4 Ein algebraischer Ansatz: Nakamura (1990)

Es sei $((S, \mathcal{P}(S)), (\mathcal{C}, \mathcal{D}), \mathcal{F}, \preceq)$ ein Entscheidungsproblem unter Unsicherheit. Der Zustandsraum S sei endlich.

N1: Schwache Ordnung. $\preceq$ ist eine schwache Ordnung.

N2: Beschränktheit. Für jedes $f \in \mathcal{F}$ gibt es $x, y \in \mathcal{C}$, so daß $x \preceq f \preceq y$.

N3: Nichttrivialität. Es gibt ein Ereignis $A \subset S$, das weder null noch universell ist.

N4: Lösbarkeit. Für alle $f \in \mathcal{F}$, $x, y, z \in \mathcal{C}$, $A \subset S$ gilt: Ist $x_A z_{-A} \preceq f \preceq y_A z_{-A}$, so gibt es $a \in \mathcal{C}$, so daß $f \sim a_A z_{-A}$.

N5: Monotonie. Für alle $A \subset S$, A nicht null, und $x, y, z \in \mathcal{C}$ mit $x, y \preceq z$ gilt: $x \preceq y \iff x_A z_{-A} \preceq y_A z_{-A}$.

Für alle $A \subset S$, A nicht universell, und $x, y, z \in \mathcal{C}$ mit $z \preceq x, y$ gilt: $x \preceq y \iff z_A x_{-A} \preceq z_A y_{-A}$.

N6: Monotonie. Für alle $x, y \in \mathcal{C}$, $A, B \subset S$ gilt: Ist $x \preceq y$ und $A \subset B$, so ist $x_B y_{-B} \preceq x_A y_{-A}$.

N7: Archimedisches Axiom. Jede streng beschränkte Standardfolge[3] ist endlich.

N8: Symmetrie. Für alle Partitionen $\{A_1, \ldots, A_n\}$ von S, $A \subset S$, x, y, $x_1, \ldots, x_n$, $y_1, \ldots, y_n$, $z_1, \ldots, z_n \in \mathcal{C}$ gilt: Ist $x_1 \preceq \ldots \preceq x_n$, $y_1 \preceq \ldots \preceq y_n$, $x_i \preceq y_i$, $i = 1, \ldots n$,

$$z_i \sim \begin{bmatrix} y_i & x_i \\ A^c & A \end{bmatrix}, \; i = 1, \ldots, n,$$

$$x \sim \begin{bmatrix} x_1 & \cdots & x_n \\ A_1 & \cdots & A_n \end{bmatrix} \quad \text{und} \quad y \sim \begin{bmatrix} y_1 & \cdots & y_n \\ A_1 & \cdots & A_n \end{bmatrix},$$

so gilt auch

$$\begin{bmatrix} y & x \\ A^c & A \end{bmatrix} \sim \begin{bmatrix} z_1 & \cdots & z_n \\ A_1 & \cdots & A_n \end{bmatrix}$$

Satz B.4 (Nakamura 1990) *Erfüllt die Relation $\preceq$ auf $\mathcal{F}$ die Axiome* **N1**–**N8**, *so gibt es eine Funktion $u : \mathcal{C} \to \mathbb{R}$ und eine Kapazität μ auf $\mathcal{P}(S)$, so daß $\preceq$ durch*

$$f \mapsto \int_S (u \circ f) d\mu$$

repräsentiert wird. Dabei ist μ eindeutig. u ist eindeutig bis auf positive affine Transformationen.

[3] Zur Definition einer Standardfolge siehe Definition 7.1. Eine Standardfolge $\{a_i\}$ heiß streng beschränkt, wenn es $b, c \in \mathcal{C}$ gibt mit $b \prec a_i \prec c$ für alle i.

Literaturverzeichnis

Allais, M. (1953a). Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'ecole américaine. *Colloques Internationaux du Centre National de la Recherche Scientifique, Vol.* XL*, Econométrie*, Paris. 257–332. Englischsprachiger Nachdruck: The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American school. In: *Expected Utility Hypotheses and the Allais Paradox* (Allais, M. und Hagen, O., 1979, eds.). Reidel, Dordrecht. 27–145.

Allais, M. (1953b). Le comportement de l'homme rationnel devant le risque. Critique des postulats et axiomes de l'école américaine. *Econometrica* **21**, 503–546; gekürzte Fassung von Allais (1953a).

Allais, M. (1979). The so-called Allais paradox and rational decisions under uncertainty. In: *Expected Utility Hypotheses and the Allais Paradox* (Allais, M. und Hagen, O., eds.). Reidel, Dordrecht. 437–681.

Anger, B (1977). Representation of capacities. *Mathematische Annalen* **229**, 245–258.

Anscombe, F.J. und Aumann, R.J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics* **34**, 199–205.

Arrow, K.J. (1965). *Aspects of the Theory of Risk-Bearing*. Academic Bookstore, Helsinki.

Arrow, K.J. (1971). *Essays in the Theory of Risk-Bearing*. North-Holland, Amsterdam.

Ash, R.B. (1972). *Real Analysis and Probability*. Academic Press, New York.

Bassanezi, R.C. und Greco, G.H. (1984). Sull'additività dell'integrale. *Rendiconti del Seminario Matematico dell'Università di Padova* **72**, 249–275.

Bergmann, R. (1991). Stochastic orders and their applications to a unified approach to various concepts of dependence and association. In: *Stochastic Orders and Decision under Risk* (Mosler, K. und Scarsini, M., eds.). Institute of Mathematical Statistics, Hayward, CA. 48–73.

Bernoulli, D. (1738). Specimen theoria novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae* **5**, 175–192. Englischsprachige Übersetzung von Sommer, L. (1954). Exposition of a new theory on the measurement of risk. *Econometrica* **12**, 23–36.

Billingsley, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

Bühler, W. (1975). Characterization of the extreme points of a class of special polyhedra. *Z. für Operations Research* **19**, 131–137.

Buja, A. (1984). Simultaneously least favorable experiments. Part I: Upper standard functionals and sufficiency. *Z. für Wahrscheinlichkeitstheorie und verwandte Gebiete / Probability Theory and related Fields* **65**, 367–384.

Camerer, C. und Weber, M. (1992). Recent developments in modelling preferences: Uncertainty and ambiguity. *J. of Risk and Uncertainty* **5**, 325–370.

Chateauneuf, A. (1991). On the use of capacities in modeling uncertainty aversion and risk aversion. *J. of Mathematical Economics* **20**, 343–369.

Chateauneuf, A. (1993). Modeling attitudes towards uncertainty and risk through the use of Choquet integral. Erscheint in *Annals of Operations Research*.

Chateauneuf, A. und Cohen, M. (1990). Risk seeking with diminishing marginal utility in a non-expected utility model. Working Paper, Université de Paris I.

Chateauneuf, A. und Jaffray, J.Y. (1989). Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences* **17**, 263–283.

Chateauneuf, A., Kast, R. und Lapied, A. (1992a). Choquet pricing for financial markets with frictions. Working Paper, Université de Paris I, GREQE Marseille, Université de Toulon.

Chateauneuf, A., Kast, R. und Lapied, A. (1992b). Pricing in slack markets. Working Paper, Université de Paris I, GREQE Marseille, Université de Toulon.

Chew, S.H. (1989). An axiomatic generalization of the quasilinear mean and Gini mean with application to decision theory. Working Paper, Johns Hopkins University and Tulane University.

Chew, S.H. und Karni, E. (1991). Choquet expected utility with a finite state space. Erscheint in *J. of Economic Theory*.

Chew, S.H., Karni, E. und Safra, Z. (1987). Risk aversion in the theory of expected utility with rank dependent probabilities. *J. of Economic Theory* **42**, 370–381.

Choquet, G. (1953/54). Theory of capacities. *Annales de l'Institut Fourier* **5**, 131–295.

Debreu, G. (1976). Least concave utility functions. *J. of Mathematical Economics* **3**, 121–129.

Dellacherie, C. (1971). Quelques commentaires sur les prolongements de capacités. In: *Seminaire de Probabilités V Strasbourg*. Springer, Berlin. 77–81.

Dempster, A.P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics* **38**, 325–339.

Denneberg, D. (1989). Verzerrte Wahrscheinlichkeiten in der Versicherungsmathematik. quantilsabhängige Prämienprinzipien. Preprint Nr. 34, Fachbereich Mathematik/Informatik, Universität Bremen.

Denneberg, D. (1990). Subadditive measure and integral. Preprint Nr. 39, Fachbereich Mathematik/Informatik, Universität Bremen.

Denneberg, D. (1992). Lectures on non-additive measure and integral. Preprint Nr. 42, Fachbereich Mathematik/Informatik, Universität Bremen.

Dharmadhikari, S.W. und Joag-Dev, K. (1988). *Unimodality, Convexity and Applications*. Academic Press, Boston, MA.

Dow, J. und Werlang, S. (1992). Uncertainty aversion, risk aversion, and the optimal choice of portfolios. *Econometrica* **60**, 197–204.

Dyckerhoff, R. (1994). Decomposition of utility functions in non-additive expected utility theory. *J. of Multi-Criteria Decision Analysis* **3**, 41–58.

Dyckerhoff, R. und Mosler, K. (1993). Stochastic dominance with nonadditive probabilities. *Z. für Operations Research* **37**, 231–256.

Eaton, M.L. (1987). *Lectures on Topics in Probability Inequalities*. Centrum voor Wiskunde en Informatica, Amsterdam.

Edwards, W. (1955). The prediction of decisions among bets. *J. of Experimental Psychology* **50**, 201–214.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review* **69**, 109–135.

Ellsberg, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly J. of Economics* **75**, 643–669.

Epstein, L.G. und Wang, T. (1992). Intertemporal asset pricing under Knightian uncertainty. Working Paper, University of Toronto.

Epstein, L.G. und Tanny, S.M. (1980). Increasing generalized correlation: A definition and some economic consequences. *Canadian J. of Economics* **13**, 16–34.

Farquhar, P.H. (1984). Utility assessment methods. *Management Science* **30**, 1283–1300.

Fellner, W. (1961). Distortion of subjective probabilities as a reaction to uncertainty. *Quarterly J. of Economics* **75**, 670–689.

Fishburn, P.C. (1964). *Decision and Value Theory*. Wiley, New York.

FISHBURN, P.C. (1965). Analysis of decisions with incomplete knowledge of probabilities. *Operations Research* **13**, 217–237.

FISHBURN, P.C. (1978). On Handa's „New theory of cardinal utility" and the maximization of expected return. *J. of Political Economy* **86**, 321–324.

FISHBURN, P.C. (1982). *The Foundations of Expected Utility*. Reidel, Dordrecht.

FISHBURN, P.C. (1988). *Nonlinear Preference and Utility Theory*. Johns Hopkins University Press, Baltimore, MD.

GILBOA, I. (1987). Expected utility with purely subjective non-additive probabilities. *J. of Mathematical Economics* **16**, 65–88.

GILBOA, I. (1989). Duality in non-additive expected utility. *Annals of Operations Research* **19**, 405–414.

GILBOA, I. und SCHMEIDLER, D. (1989). Maxmin expected utility with non-unique prior. *J. of Mathematical Economics* **18**, 141–153.

GILBOA, I. und SCHMEIDLER, D. (1992a). Additive representations of non-additive measures and the Choquet integral. Erscheint in *Annals of Operations Research*.

GILBOA, I. und SCHMEIDLER, D. (1992b). Canonical representations of set functions. Preprint, Kellogg Graduate School of Management, Northwestern University.

GRECO, G.H. (1977). Integrale monotono. *Rendiconti del Seminario Matematico dell'Università di Padova* **57**, 149–166.

GRECO, G.H. (1981). Sur la mesurabilité d'une fonction numérique par rapport a une famille d'ensembles. *Rendiconti del Seminario Matematico dell'Università di Padova* **65**, 163–176.

HAGEN, O. (1979). Towards a positive theory of preferences under risk. In: *Expected Utility Hypotheses and the Allais Paradox* (Allais, M. und Hagen, O., eds.). Reidel, Dordrecht. 271–302.

HALMOS, P. (1974). *Measure Theory*. 2nd printing. Springer, New York.

HANDA, J. (1977). Risk, probabilities and a new theory of cardinal utility. *J. of Political Economy* **85**, 97–122.

HEILMANN, W.R. (1988). *Fundamentals of Risk Theory*. Verlag Versicherungswirtschaft, Karlsruhe.

HODGES, JR., J.L. und LEHMANN, E.L. (1952). The use of previous experience in reading statistical decisions. *Annals of Mathematical Statistics* **23**, 396–407.

HOLZ, H. und MOSLER, K. (1992). An interactive decision procedure with multiple attributes under risk. Erscheint in *Annals of Operations Research*.

HUBER, P.J. (1981). *Robust Statistics*. Wiley, New York.

HUBER, P.J. und STRASSEN, V. (1973). Minimax tests and the Neyman-Pearson lemma for capacities. *Annals of Statistics* **1**, 251–263.

HURWICZ, L. (1951). Optimality criteria for decision making under ignorance. Cowles Commission Discussion Paper, Statistics, No. 370.

KAHNEMAN, D. und TVERSKY, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica* **47**, 263–291.

KAMAE, T., KRENGEL, U. und O'BRIEN, G.L. (1977). Stochastic inequalities on partially ordered spaces. *Annals of Probability* **5**, 899–912.

KARMARKAR, U.S. (1978). Subjectively weighted utilty: A descriptive extension of the expected utility model. *Organizational Behavior and Human Performance* **21**, 61–72.

KARMARKAR, U.S. (1979). Subjectively weighted utility and the Allais paradox. *Organizational Behavior and Human Performance* **24**, 67–72.

KEENEY, R.L. und RAIFFA, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Wiley, New York.

KINDLER, J. (1986). A Mazur-Orliz type theorem for submodular set functions. *J. of Mathematical Analysis and Applications* **120**, 533–546.

KISCHKA, P. und PUPPE, C. (1992). Decisions under risk and uncertainty: A survey of recent developments. *Z. für Operations Research* **36**, 125–147.

KOFLER, E. (1974). Entscheidungen bei teilweise bekannter Verteilung der Zustände. *Z. für Operations Research* **18**, 141–157.

KOFLER, E. und MENGES, G. (1976). *Entscheidungen bei unvollständiger Information*. Springer, Berlin.

KRANTZ, D.H., LUCE, R.D., SUPPES, P. und TVERSKY, A. (1971). *Foundations of Measurement, Volume I (Additive and Polynomial Representations)*. Academic Press, New York.

KRUMBHOLZ, W. (1982). Die Bestimmung einfacher Attributprüfpläne unter Berücksichtigung von unvollständiger Vorinformation. *Allgemeines Statistisches Archiv* **66**, 240–253.

LEHMANN, E.L. (1955). Ordered families of distributions. *Annals of Mathematical Statistics* **26**, 399–419.

LEVHARI, D., PAROUSH, J. und PELEG, B. (1975). Efficiency analysis for multivariate distributions. *Review of Economic Studies* **42**, 87–91.

LEVY, H. und LEVY, A. (1984). Multivariate decision-making. *J. of Economic Theory* **32**, 36–51.

MACCRIMMON, K.R. (1968). Descriptive and normative implications of the decision-theory postulates. In: *Risk and Uncertainty* (Borch, K. und Mossin, J., eds.). Macmillan, New York. 3–32.

MACCRIMMON, K.R. und LARSSON, S. (1979). Utility theory: Axioms versus „paradoxes“. In: *Expected Utility Hypotheses and the Allais Paradox* (Allais, M. und Hagen, O., eds.). Dordrecht, Reidel. 333–409.

Marshall, A.W. und Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic Press, New York.

McCord, M. und de Neufville, R. (1986). “Lottery equivalent”: Reduction of the certainty effect problem in utility assessment. *Management Science* **32**, 56–60.

Miyamoto, J.M. (1988). Generic utility theory: Measurement foundations and applications in multiattribute utility theory. *J. of Mathematical Psychology* **32**, 357–404.

Moskowitz, H. (1974). Effects of problem representation and feedback on rational behavior in Allais and Morlat-type problems. *Decision Sciences* **5**, 225–242.

Mosler, K. (1982). *Entscheidungsregeln bei Risiko: Multivariate stochastische Dominanz*. Springer, Berlin.

Mosler, K. (1984). Stochastic dominance decision rules when the attributes are utility independent. *Management Science* **30**, 1311–1322.

Mosler, K. (1987). Estimation under G-invariant quasi-convex loss. *J. of Multivariate Analysis* **22**, 137–143.

Mosler, K. und Scarsini, M., Eds. (1991a). *Stochastic Orders and Decision under Risk*. Institute of Mathematical Statistics, Hayward, CA.

Mosler, K. und Scarsini, M. (1991b). Some theory of stochastic dominance. In: *Stochastic Orders and Decision under Risk* (Mosler, K. und Scarsini, M., eds.). Institute of Mathematical Statistics, Hayward, CA. 261–284.

Muliere, P. und Scarsini, M. (1989). Multivariate decisions with unknown price vector. *Economics Letters* **29**, 13–19.

Nakamura, Y. (1990). Subjective expected utility with non-additive probabilities on finite spaces. *J. of Economic Theory* **51**, 346–366.

von Neumann, J. und Morgenstern, O. (1947). *Theory of Games and Economic Behavior*. 2nd edition. Princeton University Press, Princeton, NJ.

Neveu, J. (1969). *Mathematische Grundlagen der Wahrscheinlichkeitstheorie*. Oldenbourg, München.

von Nitzsch, R. und Weber, M. (1986). Die verläßliche Bestimmung von Nutzenfunktionen. *Z. für betriebswirtschaftliche Forschung* **38**, 844–862.

von Nitzsch, R. und Weber, M. (1988). Utility function assessment on a microcomputer: An interactive procedure. *Annals of Operations Research* **16**, 149–160.

Pfanzagl, J. und Pierlo, W. (1966). *Compact Systems of Sets*. Springer, Berlin.

Pratt, J.W. (1964). Risk aversion in the small and the large. *Econometrica* **32**, 122–136.

Quiggin, J. (1982). A theory of anticipated utility. *J. of Economic Behavior and Organization* **3**, 323–343.

Quiggin, J. (1985). Subjective utility, anticipated utility and the Allais paradox. *Organizational Behavior and Human Decision Processes* **35**, 94–101.

Quiggin, J. (1987). Decision weights in anticipated utility theory. *J. of Economic Behavior and Organization* **8**, 641–645.

Quiggin, J. (1993). *Generalized Expected Utility Theory. The Rank-Dependent Model*. Kluwer Academic Publishers, Norwell, MA.

Quiggin, J. und Wakker, P. (1992). The axiomatic basis of anticipated utility: A clarification. CenTER Discussion Paper No. 9203, University of Brabant, Tilburg.

Ramsey, F.P. (1931). Truth and probability. In: *The Foundations of Mathematics and Other Logical Essays*. Routledge and Kegan Paul, London. 156–198.

Reich, A. (1985). Eine Charakterisierung des Standardabweichungsprinzips. *Blätter der Deutschen Gesellschaft für Versicherungsmathematik* **18**, 93–102.

Richard, S. (1975). Multivariate risk aversion, utility independence and separable utility functions. *Management Science* **22**, 12–21.

Röell, A. (1987). Risk aversion in Quiggin and Yaari's rank-order model of choice under uncertainty. *Economic J.* **97**, 143–159.

Rolski, T. (1976). Order relations in the set of probability distribution functions and their applications in queueing theory. *Dissertationes Mathematicae* **132**, 52ff.

Rothschild, M. und Stiglitz, J.E. (1970). Increasing risk: I. A definition. *J. of Economic Theory* **2**, 225–243.

Rothschild, M. und Stiglitz, J.E. (1971). Increasing risk: II. Its economic consequences. *J. of Economic Theory* **3**, 66–84.

Rüschendorf, L. (1981). Stochastically ordered distributions and monotonicity of the OC-function of sequential probability ratio tests. *Mathematische Operationsforschung und Statistik. Series Statistics* **12**, 327–338.

Samuelson, P.A. (1967). General proof that diversification pays. *J. of Financial and Quantitative Analysis* **2**, 1–13.

Savage, L.J. (1954). *The Foundations of Statistics*. Wiley, New York.

Scarsini, M. (1985). Stochastic dominance with pair-wise risk aversion. *J. of Mathematical Economics* **14**, 187–201.

Scarsini, M. (1988). Dominance conditions for multivariate utility functions. *Management Science* **34**, 454–460.

Scarsini, M. (1992). Dominance conditions in non-additive expected utility theory. *J. of Mathematical Economics* **21**, 173–184.

SCHMEIDLER, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society* **97**, 255–261.

SCHMEIDLER, D. (1989). Subjective probability and expected utility without additivity. *Econometrica* **57**, 571–587.

SCHNEEWEISS, H. (1964). Eine Entscheidungsregel für den Fall partiell bekannter Wahrscheinlichkeiten. *Unternehmensforschung* **8**, 86–95.

SEGAL, U. (1987). Some remarks on Quiggin's anticipated utility. *J. of Economic Behavior and Organization* **8**, 145–154.

SEGAL, U. (1989). Anticipated utility: A measure representation approach. *Annals of Operations Research* **19**, 359–373.

SEGAL, U. (1993). The measure representation approach: A correction. *J. of Risk and Uncertainty* **6**, 99–107.

SHAFER, G.A. (1976). *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ.

ŠIPOŠ, JÁN (1979). Integral with respect to a pre-measure. *Mathematica Slovaca* **29**, 141–155.

SLOVIC, P. und TVERSKY, A. (1974). Who accepts Savage's axiom? *Behavioral Science* **19**, 368–373.

VON STENGEL, B. (1993). Closure properties of independence concepts for continuous utilities. *Mathematics of Operations Research* **18**, 346–389.

STOER, J. (1979). *Einführung in die Numerische Mathematik I*. 3. Auflage. Springer, Berlin.

TOPSØE, F. (1978). On construction of measures. In: *Proceedings of the Conference on Topology and Measure, I (Zinnowitz 1974)*. Ernst-Moritz-Arndt Universität, Greifswald. 343–381.

TVERSKY, A und KAHNEMAN, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *J. of Risk and Uncertainty* **5**, 297–323.

WAKKER, P. (1989a). *Additive Representations of Preferences. A New Foundation of Decision Analysis*. Kluwer Academic Publishers, Dordrecht.

WAKKER, P. (1989b). Continuous subjective expected utility with non-additive probabilities. *J. of Mathematical Economics* **18**, 1–27.

WAKKER, P. (1990a). Characterizing optimism and pessimism directly through comonotonicity. *J. of Economic Theory* **52**, 453–463.

WAKKER, P. (1990b). Under stochastic dominance Choquet-expected utility and anticipated utility are identical. *Theory and Decision* **29**, 119–132.

WAKKER, P. (1991). Additive representations of preferences, a new foundation of decision analyis; the algebraic approach. In: *Mathematical Psychology: Current Developments* (Doignon, J.P. und Falmagne, J.C., eds.). Springer, Berlin. 71–87.

WAKKER, P. (1993). Unbounded utility for Savage's 'Foundations of Statistics' and other models. *Mathematics of Operations Research* **18**, 446–485.

WAKKER, P. (1994). Separating marginal utility and probabilistic risk aversion. *Theory and Decision* **36**, 1–44.

WALD, A. (1950). *Statistical Decision Functions*. Wiley, New York.

WEBER, M. (1987). Decision making with incomplete information. *European J. of Operational Research* **28**, 44–57.

WILLARD, S. (1970). *General Topology*. Addison-Wesley, Reading, MA.

YAARI, M.E. (1987). The dual theory of choice under risk. *Econometrica* **55**, 95–115.