Z. Wahrscheinlichkeitstheorie verw. Geb. 12, 21–55 (1969)

Erschöpftheit und Invarianz beim Vergleich von Experimenten

HERBERT HEYER*

Summary. Let **E** be the collection of all experiments (in the sense of LeCam) with the same index set. Using the notion of ε-informativity a binary relation is introduced in **E**, which turns out to be equivalent with various other relations of comparison studied previously. For a certain subclass of **E** a characterization of ε-informativity via randomisation kernels yields the equivalence to Blackwell's sufficiency and in the case of contraction experiments also to sufficiency in the sense of Halmos and Savage. Furthermore the concept of translation invariance is generalized to large subclasses of **E**. In the case of invariance the characterizing kernels can also be chosen invariant. Applications of the theory to allocation and testing problems as well as to Shannon- and Fisher-informativity illuminate the general set up. The special case of experiments with only a finite index set appears strongly related to the study of orderings in the set of all positive Radon measures on compact convex subsets of certain topological vector spaces.

Inhalt

1. Einleitung

In seiner grundlegenden Arbeit [2] definiert Blackwell ein statistisches Experiment und führt in einer gewissen Klasse von Experimenten Vergleichsrelationen ein: die Informativitätsrelation und die Erschöpftheitsrelation. Etwas genauer besteht bei Blackwell ein Experiment aus einer Menge X, einer σ-Algebra $\mathfrak{A}$ von Teilmengen von X sowie einer Familie $(P_i)_{i\in I}$ von endlich vielen Wahrscheinlichkeitsmaßen (W-Maßen) auf $\mathfrak{A}$. Bezeichnen wir mit $\mathscr{L}_{\mathfrak{A}} := \mathscr{L}(X, \mathfrak{A})$ die Menge der $\mathfrak{A}$-meßbaren, beschränkten reellen Funktionen auf X, so erhalten wir das Tripel $(X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$, welches aus einer Menge X, einem Banachverband von reellen beschränkten Funktionen auf X und einer Familie von positiven, normierten Linearformen auf $\mathscr{L}_{\mathfrak{A}}$ besteht.

Zum Experiment $(X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ definiert man die Menge $B := B\{(P_i)_{i\in I}, \mathscr{L}_{\mathfrak{A}}\}$ als den kleinsten linearen Teilraum des Dualraums $\mathscr{L}'_{\mathfrak{A}}$ von $\mathscr{L}_{\mathfrak{A}}$, welcher die

* Die vorliegende Arbeit wurde von der Naturwissenschaftlichen Fakultät der Friedrich-Alexander-Universität Erlangen-Nürnberg im SS 1968 als Habilitationsschrift angenommen.

Familie $(P_i)_{i\in I}$ enthält, bezüglich der Normtopologie in $\mathscr{L}'_{\mathfrak{A}}$ abgeschlossen ist und die folgende Eigenschaft besitzt: Für $\mu\in B$ und $\nu\in\mathscr{L}'_{\mathfrak{A}}$ mit $0\leqq|\nu|\leqq|\mu|$ gilt $\nu\in B$. B wird auch das kleinste, die Familie $(P_i)_{i\in I}$ enthaltende Band in $\mathscr{L}'_{\mathfrak{A}}$ oder der L-Raum des Experiments $(X,\mathscr{L}_{\mathfrak{A}},(P_i)_{i\in I})$ genannt.

Im Jahre 1964 publizierte LeCam eine Arbeit [11], in der von einer möglichst abstrakten Definition eines Experiments ausgegangen wird. Diese Definition verallgemeinert die von Blackwell in zwei Richtungen: einmal werden anstelle der Verbände $\mathscr{L}_{\mathfrak{A}}$ beliebige Banachverbände E reeller beschränkter Funktionen auf einer Menge X, welche die konstanten Funktionen enthalten, zugelassen, andererseits beliebige Familien $(P_i)_{i\in I}$ von positiven, normierten Linearformen auf E. Zudem wird die Vergleichsrelation, bei LeCam Defizienz genannt, weiter gefaßt, indem beim Vergleich von zwei Experimenten $\mathfrak{X}$ und $\mathfrak{Y}$ mit gleicher Indexmenge eine Defektfunktion $\varepsilon\geqq 0$ benützt wird. Dadurch entsteht der Begriff der ε-Defizienz. Die Defizienz zwischen Experimenten unterscheidet sich im Spezialfall der klassischen (nichtrandomisierten) Theorie von der Informativität Blackwells dadurch, daß zwei verschiedene Klassen von Entscheidungsfunktionen auftreten und daher keine Transitivität mehr vorliegt.

Zur Charakterisierung der Defizienz etwa im Falle zweier klassischer Experimente $(X,\mathscr{L}_{\mathfrak{A}},(P_i)_{i\in I})$ und $(Y,\mathscr{L}_{\mathfrak{B}},(Q_i)_{i\in I})$ zieht LeCam positive, normierte Bilinearformen auf dem Produktraum $\mathscr{L}_{\mathfrak{B}}\times B$ heran, welche er randomisierte Abbildungen (des Experiments $(X,\mathscr{L}_{\mathfrak{A}},(P_i)_{i\in I})$ in das Paar $(Y,\mathscr{L}_{\mathfrak{B}})$) nennt.

Der Ausgangspunkt der vorliegenden Arbeit ist der Rückgriff auf die intuitive Informativität Blackwells, nun sogar für abstrakte Experimente und zugleich für Fehlerfunktionen $\varepsilon\geqq 0$ verallgemeinert (ε-Informativität), sowie deren Charakterisierung mittels sog. Randomisierungskerne (R-Kerne), welche an die Stelle der von LeCam benötigten Bilinearformen treten. In obigem Beispiel sind die R-Kerne genau die positiven, normierten linearen Abbildungen von $\mathscr{L}_{\mathfrak{B}}$ in $\mathscr{L}_{\mathfrak{A}}$. Sei T ein R-Kern von $\mathscr{L}_{\mathfrak{B}}$ nach $\mathscr{L}_{\mathfrak{A}}$, so induziert dieser zwar vermöge $T(g,\mu):=\langle T(g),\mu\rangle$ eine Bilinearform auf $\mathscr{L}_{\mathfrak{B}}\times B$, welche zur Charakterisierung der ε-Informativität führt. Die Existenz eines derartigen R-Kerns ist jedoch nicht immer garantiert.

Somit dürften die nachfolgenden Ausführungen nur in einzelnen Situationen mit der in [11] dargelegten Theorie vergleichbar sein.

Ein weitgehendes Motiv für den Rückgriff auf die Informativität Blackwells liefert die Einbettung der klassischen Erschöpftheitsrelation in den allgemeinen Rahmen sowie deren Spezialisierung für gewisse Klassen von Unterexperimenten.

In den §§ 2 und 3 wird der technische Apparat für die nachfolgenden Untersuchungen bereitgestellt: die für abstrakte Experimente verallgemeinerten Begriffe von Verlust, Risiko und ε-Informativität sowie Definition und grundlegende Eigenschaften des R-Kerns. In der Menge $\mathscr{R}$ der R-Kerne zwischen gewissen Banachverbänden reeller beschränkter Funktionen werden zwei verschiedene Topologien erklärt und bezüglich dieser Topologien dichte Teilmengen von $\mathscr{R}$ untersucht. Hinzu tritt ein Abschnitt über dominierte klassische Experimente. § 4 enthält einen Charakterisierungssatz für die Relation der ε-Informativität zwischen einem klassischen Experiment $\mathfrak{X}$ und einem beliebigen Experiment $\mathfrak{Y}$. Es werden Spezialfälle diskutiert: der Fall strikter ε-Informativität (ε konstant) und der Fall der Informativität schlechthin ($\varepsilon\equiv 0$). Mit Hilfe dieses Satzes läßt sich ein Resultat beweisen, welches die ε-Informativität zwischen Experimenten $\mathfrak{X}$ und $\mathfrak{Y}$ kennzeichnet durch die Existenz eines Hilfsexperiments $\mathfrak{Y}^*$, bezüglich dessen $\mathfrak{X}$ informativer ist. Der Satz beinhaltet das, was gelegentlich die Stabilität des Prinzips der Informativität genannt wird.

Ist ein Experiment $\mathfrak{X}$ einer gewissen Klasse informativer als ein beliebiges Experiment $\mathfrak{Y}$, so existiert ein Produktexperiment, das aus $\mathfrak{X}$ und $\mathfrak{Y}$ konstruiert wird und zudem weniger informativ ist als $\mathfrak{X}$. Dieser Satz sowie seine Umkehrung ermöglichen zahlreiche Anwendungen.

Im § 5 nun erfolgt die Diskussion der Frage nach dem Zusammenhang zwischen den in der Theorie des Vergleichs von Experimenten und den in der mathematischen Statistik auftretenden Konzepten der Erschöpftheit. Der Hauptsatz liefert wenigstens für eine Klasse von dominierten Experimenten die Äquivalenz von Informativität und Erschöpftheit. Eine weitere gleichwertige Aussage stellt die Beziehung der Vergleichsrelationen zur Theorie der erwartungstreuen Schätzungen her.

Die Spezialisierung auf Kontraktionsexperimente liefert die Äquivalenz aller bekannten Versionen von Erschöpftheit, insbesondere die Äquivalenz zum Begriff der erschöpfenden σ-Algebra. Der Fall erschöpfender Kontraktionsexperimente wird ferner mit Hilfe der von der Informationstheorie her bekannten f-Divergenz beschrieben.

Es folgt § 8 über die Existenz invarianter R-Kerne für spezielle Klassen von Experimenten und zulässige Familien von Abbildungen, bezüglich deren der Invarianzbegriff definiert wird. Es wird zudem eine hinreichende Bedingung für die Existenz genau eines invarianten R-Kernes angegeben, welcher die Maße des Experiments $\mathfrak{X}$ in die des Experiments $\mathfrak{Y}$ überführt. Anwendung auf die Gruppe der Translationen einer lokalkompakten abelschen Gruppe G mit abzählbarer Basis ergibt die Existenz eines translationsinvarianten Markoff-Kerns für klassische Experimente mit Grundraum G, bei denen die zugehörigen W-Maße aus einem W-Maß durch Translation erzeugt werden.

Beispiele aus der Theorie der Placierungsexperimente, endlicher klassischer Experimente und aus der Testtheorie sollen die Existenz- und Äquivalenzsätze der §§ 4 und 5 beleuchten und deren Gültigkeitsbereich abgrenzen. Im Falle endlicher klassischer Experimente werden Zusammenhänge der vorliegenden Ausführungen zur Theorie der Integraldarstellungen in konvexen, kompakten Mengen sichtbar. Diese Beziehungen, die im Spezialfall schon auf Hardy, Littlewood und Polya (1934) zurückgehen, dürften am Anfang der von Blackwell begründeten Theorie des Vergleichs von Experimenten gestanden haben.

Eine Gegenüberstellung der bisher betrachteten Vergleichsrelationen für Experimente mit den Relationen von Shannon und Fisher zeigt die Schärfe der von uns betrachteten Informativitätsrelation.

Im Anhang werden einige Resultate der Arbeit, die Vergleichsrelationen in der Klasse aller Experimente mit gleicher Indexmenge betreffend, im Sinne der Theorie der geordneten bzw. quasi-geordneten Mengen interpretiert. Beispielsweise bildet die Klasse der dominierten topologischen Experimente ein bezüglich der Produktbildung von Experimenten quasi-geordnetes Gruppoid.

Der Verfasser dankt Herrn Diplom-Mathematiker J. Greiner für die sorgfältige Analyse einiger in der Arbeit von LeCam [11] enthaltener Ideen, welche für die vorliegende Arbeit relevant waren.

2. Vorbereitungen

2.1. Experimente

Es sei $(X_0, \mathfrak{A}_0, P_0)$ ein Wahrscheinlichkeitsraum (W-Raum), welcher mittels einer Zufallsvariablen ξ auf den Stichprobenraum $(X, \mathfrak{A})$ abgebildet wird. Eine Familie $(P_i)_{i\in I}$ von W-Maßen auf $(X, \mathfrak{A})$, die bei statistischen Untersuchungen als die Menge der möglichen Verteilungen von ξ auftritt, definiert ein statistisches Experiment $(X, \mathfrak{A}, (P_i)_{i\in I})$ im Sinne von Blackwell [2]. Dieser Begriff wird verallgemeinert in der folgenden

2.1.1. Definition. *Experiment* (mit Grundmenge X und Indexmenge I) heißt jedes Tripel $\mathfrak{X}=(X, E, (P_i)_{i\in I})$, wobei E ein Banachverband reeller beschränkter Funktionen auf der Menge X mit $1\in E$ und $(P_i)_{i\in I}$ eine nichtleere Familie reeller positiver, normierter Linearformen auf E ist.

Dabei trage E die Supremum-Norm, und für je zwei Elemente $f, g\in E$ seien das Supremum $f\vee g$ sowie das Infimum $f\wedge g$ punktweise erklärt. Im übrigen sind die auftretenden Banachverbände identisch mit den das Element 1 enthaltenden abgeschlossenen Unteralgebren der Algebra aller beschränkten reellen Funktionen auf X.

Ein Experiment heißt *klassisch*, wenn E gleich ist der Menge $\mathscr{L}_{\mathfrak{A}}:=\mathscr{L}(X, \mathfrak{A})$ aller reellen beschränkten, $\mathfrak{A}$-meßbaren Funktionen auf X, wobei $\mathfrak{A}$ eine σ-Algebra in X bezeichnet und die P_i W-Maße auf $\mathfrak{A}$ sind.

Ein klassisches Experiment heißt *diskret*, wenn für jedes $i\in I$ das W-Maß P_i diskret, d.h. von der Form $\sum_{k\geqq 1} r_k\, \varepsilon_{x_k}$ mit $r_k\geqq 0$ für $k\geqq 1$ und $\sum_{k\geqq 1} r_k=1$ ist, wobei die Folge $(x_k)_{k\geqq 1}$ in X liegt.

Im Falle eines klassischen Experiments $\mathfrak{X}$ mit $\mathfrak{A}:=\mathfrak{P}(X)$, wobei $\mathfrak{P}(X)$ die Potenzmenge von X bezeichnet, fällt $\mathscr{L}_{\mathfrak{A}}$ mit der Menge $\mathscr{L}_X$ aller beschränkten reellen Funktionen auf X zusammen.

Weiterhin heißt ein klassisches Experiment *topologisch*, wenn X ein lokalkompakter Raum mit abzählbarer Basis und jedes P_i $(i\in I)$ ein Radonsches W-Maß auf X ist. In diesem Falle können bekanntlich die P_i als Borelmaße auf der Borel-σ-Algebra $\mathscr{B}:=\mathscr{B}(X)$ über X aufgefaßt werden.

Schließlich nennen wir $\mathfrak{X}$ *endlich*, falls $|I|:=\operatorname{Kard} I<\infty$.

2.1.2. Definition. *Teilexperiment* des Experiments $\mathfrak{X}=(X, E, (P_i)_{i\in I})$ heißt jedes Tripel $\mathfrak{X}_J=(X, E, (P_i)_{i\in J})$, wobei J eine nichtleere Teilmenge von I bezeichnet.

In diesem Sinne ist $\mathfrak{X}=\mathfrak{X}_I$.

2.1.3. Definition. Ist $\mathfrak{X}=(X, E, (P_i)_{i\in I})$ ein Experiment und G ein abgeschlossener Teilvektorverband von E mit $1\in G$, so heißt das Tripel $(X, G, (\operatorname{Res}_G P_i)_{i\in I})$ *Kontraktionsexperiment* von $\mathfrak{X}$ mit Banachverband G (*G-Kontraktion* von $\mathfrak{X}$). Wenn es darauf ankommt, den das Kontraktionsexperiment definierenden Banachverband G hervorzuheben, so wird das Tripel auch durch $\mathfrak{X}^G$ abgekürzt.

Für ein klassisches Experiment $\mathfrak{X}=(X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ und vorgegebene σ-Unteralgebra $\mathfrak{B}$ von $\mathfrak{A}$ nennt man die $\mathscr{L}_{\mathfrak{B}}$-Kontraktion von $\mathfrak{X}$ auch $\mathfrak{B}$-Kontraktion von $\mathfrak{X}$ und benützt das Symbol $\mathfrak{X}^{\mathfrak{B}}$.

2.1.4. Wir bemerken zuvor, daß für zwei Vektorverbände A und B von beschränkten reellen Funktionen (mit konstanter Funktion 1) das Tensorprodukt von A mit B definiert ist als der von der Menge C der Funktionen $\sum_{j=1}^{n} a_j b_j$ mit $a_j \in A$ und $b_j \in B$ $(j=1, \ldots, n)$ erzeugte Vektorverband beschränkter reeller Funktionen auf dem cartesischen Produkt $X \times Y$ der Mengen X und Y. Hierbei ist jede Funktion $c := \sum_{j=1}^{n} a_j b_j$ auf $X \times Y$ definiert durch $c(x, y) := \sum_{j=1}^{n} a_j(x) b_j(y)$.

Sind nun A und B zudem Banachverbände, so definiert man das Tensorprodukt $A \otimes B$ als den (gleichmäßigen) Abschluß des von C erzeugten Vektorverbandes (und erhält damit wieder einen Banachverband). Für Linearformen p bzw. q auf A bzw. B ist das Tensorprodukt $p \otimes q$ in der üblichen Weise als Linearform auf $A \otimes B$ erklärt.

Definition. Es seien $\mathfrak{X} = (X, E, (P_i)_{i \in I})$ und $\mathfrak{Y} = (Y, F, (Q_i)_{i \in I})$ zwei Experimente (mit gleicher Indexmenge I) und $E \otimes F$ das Tensorprodukt der Banachverbände E und F. Dann heißt das Tripel $\mathfrak{Z} := (X \times Y, E \otimes F, (R_i)_{i \in I})$, wobei $(R_i)_{i \in I}$ eine Familie reeller positiver, normierter Linearformen auf $E \otimes F$ bezeichnet, ein *schwaches Produktexperiment.*

$$\mathrm{Pr}_1\, \mathfrak{Z} := (X, E, (\mathrm{pr}_E\, R_i)_{i \in I}) \quad \text{bzw.} \quad \mathrm{Pr}_2\, \mathfrak{Z} := (Y, F, (\mathrm{pr}_F\, R_i)_{i \in I})$$

wird *X-Projektionsexperiment* (*X-Projektion*) bzw. *Y-Projektionsexperiment* (*Y-Projektion*) von $\mathfrak{Z}$ genannt.

Das Experiment $\mathfrak{W} := (X \times Y, E \otimes F, (P_i \otimes Q_i)_{i \in I})$ heißt das (starke) *Produkt* der Experimente $\mathfrak{X}$ und $\mathfrak{Y}$, wobei $P_i \otimes Q_i$ für jedes $i \in I$ das Tensorprodukt der positiven, normierten Linearformen P_i und Q_i bezeichnet.

2.1.5. *Bemerkung.* Definition 2.1.1 geht auf LeCam [11] zurück. Definitionen 2.1.2 und 2.1.3 hingegen sind von Definition 10 in [11] wohl zu unterscheiden.

2.2. Verlust und Risiko

2.2.1. Für jedes Experiment $\mathfrak{X} = (X, E, (P_i)_{i \in I})$ und beliebige Menge D von *Entscheidungen* definiert man in bekannter Weise die *Verlustfunktion* V als reelle Funktion auf $I \times D$ und interpretiert den Wert $V(i, d)$ von V an der Stelle (i, d) als den *Verlust*, mit dem man bei der Entscheidung d zu rechnen hat, falls $\xi(P_0) = P_i$ zutrifft. Wir werden bei den folgenden Untersuchungen Entscheidungen d mit Vektoren $(V(i, d))_{i \in I}$ in $\mathbb{R}^I$ identifizieren und zudem voraussetzen, daß D stets eine nichtleere konvexe, kompakte Teilmenge von $\mathbb{R}^I$ ist. Es wird sich zeigen (in 2.2.3), daß es darüber hinaus genügt, nur nichtleere konvexe, kompakte Teilmengen von $[-1, +1]^I$ als Entscheidungsmengen zu betrachten. Deren Klasse bezeichnen wir mit $\mathscr{D}(I)$. Der Verlust bei der Entscheidung d wird sodann mittels der Projektionsabbildungen p_i durch $V(i, d) = p_i(d) = d_i$ $(i \in I)$ beschrieben.

2.2.2. Weiterhin erklärt man die zu $\mathfrak{X}$ und beliebigem D gehörige Menge $\mathfrak{E}(\mathfrak{X}, D)$ der *Entscheidungsfunktionen* als die Menge aller Abbildungen $f \colon X \to D$ mit $f = (f_i)_{i \in I}$, wobei $f_i \in E$ für alle $i \in I$.

Bezeichnen $\mathfrak{M}(X, A)$ bzw. $\mathscr{C}(X, A)$ die Vektorräume der bezüglich einer σ-Algebra $\mathfrak{A}$ in X meßbaren bzw. der stetigen Abbildungen von X in einen topologischen Raum A, so ergeben sich folgende *Beispiele* für $\mathfrak{E}(\mathfrak{X}, D)$:

(1) Man betrachte einen Meßraum $(X, \mathfrak{A})$ und setze $E := \mathscr{L}_{\mathfrak{A}}$. Dann gilt $\mathfrak{E}(\mathfrak{X}, D) = \mathfrak{M}(X, D)$.

(2) Man betrachte für einen kompakten Raum X den Meßraum $(X, \mathfrak{B})$, wobei $\mathfrak{B}$ die σ-Algebra der Borelschen Mengen über X ist, und setze $E := \mathscr{C}(X) = \mathscr{C}(X, \mathbb{R})$. Dann ist $\mathfrak{E}(\mathfrak{X}, D) = \mathscr{C}(X, D)$.

2.2.3. Die Menge $\mathfrak{E}(\mathfrak{X}, D)$ ist konvex, enthält die konstante Abbildung 1 und erfüllt für jedes $x \in \mathbb{R}^I$ die Mengengleichung

$$\mathfrak{E}(\mathfrak{X}, x + D) = x + \mathfrak{E}(\mathfrak{X}, D).$$

Darüber hinaus erhält man für jede eineindeutige Abbildung L von $\mathbb{R}^I$ in $\mathbb{R}^I$ der Gestalt $L(y) := (\alpha_i y_i)_{i \in I}$, wobei $y := (y_i)_{i \in I} \in \mathbb{R}^I$ und $\alpha_i \in \mathbb{R}$ für alle $i \in I$ ist:

$$\mathfrak{E}(\mathfrak{X}, L(D)) = \{L \circ f \colon f \in \mathfrak{E}(\mathfrak{X}, D)\}.$$

Schließlich gilt für $g \in \mathscr{C}(D)$ und $f \in \mathfrak{E}(\mathfrak{X}, D)$ stets $g \circ f \in E$.

2.2.4. Mittels einer Entscheidungsfunktion $f \in \mathfrak{E}(\mathfrak{X}, D)$ definiert man den *mittleren Verlust* $R(i, f)$ als die Zahl $\langle f_i, P_i \rangle$, falls $\xi(P_0) = P_i$ zutrifft. Hält man f fest, so heißt die Abbildung $R_f \colon I \to \mathbb{R}$, definiert durch $R_f(i) := R(i, f)$, die zu f gehörige *Risikofunktion.*

2.3. ε-Informativität

Es soll nun eine Vergleichsrelation für Experimente eingeführt werden. Nachfolgende Definition geht im Spezialfall endlicher klassischer Experimente verschwindender Fehlerfunktion ε auf Blackwell [2] zurück.

Gegeben seien zwei Experimente $\mathfrak{X} = (X, E, (P_i)_{i \in I})$ und $\mathfrak{Y} = (Y, F, (Q_i)_{i \in I})$ mit gleicher Indexmenge I sowie eine Abbildung $\varepsilon \colon I \to \mathbb{R}_+$.

2.3.1. Definition. $\mathfrak{X}$ heißt *ε-informativer* als $\mathfrak{Y}$ $(\mathfrak{X} \supset_\varepsilon \mathfrak{Y})$, wenn es zu jedem $D \in \mathscr{D}(I)$ und jedem $g \in \mathfrak{E}(\mathfrak{Y}, D)$ ein $f \in \mathfrak{E}(\mathfrak{X}, D)$ gibt mit der Eigenschaft

$$R_f \leqq R_g + \varepsilon.$$

Genauer heißt dies: $R_f(i) \leqq R_g(i) + \varepsilon(i)$ für alle $i \in I$, wobei zu beachten ist, daß zwar $R_f(i) = \langle f_i, P_i \rangle$, jedoch $R_g(i) = \langle g_i, Q_i \rangle$ für jedes $i \in I$ gesetzt wird. Ist ε die konstante Funktion auf I, so heißt $\mathfrak{X}$ *streng ε-informativer* als $\mathfrak{Y}$ $(\mathfrak{X} \sqsupset_\varepsilon \mathfrak{Y})$.

$\mathfrak{X}$ heißt *informativer* als $\mathfrak{Y}$ $(\mathfrak{X} \supset \mathfrak{Y})$, falls $\mathfrak{X}$ streng 0-informativer ist als $\mathfrak{Y}$.

Die durch $\supset$ definierte Relation ist eine Quasiordnung in der Klasse aller Experimente mit gleicher Indexmenge I: $\mathfrak{X}$ heißt zudem *äquivalent* zu $\mathfrak{Y}$, falls sowohl $\mathfrak{X} \supset \mathfrak{Y}$ als auch $\mathfrak{Y} \supset \mathfrak{X}$ gilt. Wir schreiben $\mathfrak{X} \sim \mathfrak{Y}$. Offensichtlich hat man die folgende

2.3.2. Proposition. *$\mathfrak{X} \supset_\varepsilon \mathfrak{Y}$ genau dann, wenn $\mathfrak{X}_J \supset_{\varepsilon_J} \mathfrak{Y}_J$ gilt für jede nichtleere Teilmenge J von I.*

2.3.3. Definition. $\mathfrak{X}$ heißt *bedingt ε-informativer* als $\mathfrak{Y}$ $(\mathfrak{X} >_\varepsilon \mathfrak{Y})$, wenn für jede *endliche* nichtleere Teilmenge J von I gilt $\mathfrak{X}_J \supset_{\varepsilon_J} \mathfrak{Y}_J$.

Analog zur Äquivalenz definiert man die *bedingte Äquivalenz.*

3. Randomisierung

3.1. R-Kerne

3.1.1. Unter einem *Randomisierungskern* (*R-Kern*) für zwei (abstrakte) Banachverbände A und B mit gleichbezeichnetem Einselement 1 versteht man eine positive, normierte lineare Abbildung von A in B, d.h. eine lineare Abbildung T von A in B mit $Ta \geqq 0$ für alle $a \geqq 0$ $(a \in A)$ und $T1 = 1$. Einzelheiten der Theorie der abstrakten Banachverbände werden etwa in [18], Chapter V, §1 und 7 ausgeführt. Die Menge aller R-Kerne für die Banachverbände A und B wird mit $\mathscr{R}(A, B)$ abgekürzt.

Es werde nun ein Experiment $\mathfrak{X} = (X, E, (P_i)_{i \in I})$ zugrunde gelegt. Zu gegebener Entscheidungsmenge $D \in \mathscr{D}(I)$ betrachten wir den Banachverband $\mathscr{C}(D)$. Dann läßt sich die Menge $\mathfrak{E}(\mathfrak{X}, D)$ in der folgenden Weise mit Hilfe von Elementen aus $\mathscr{R}(\mathscr{C}(D), E)$ darstellen:

3.1.2. Proposition. $\mathfrak{E}(\mathfrak{X}, D) = \{(Tp_i)_{i \in I} \colon T \in \mathscr{R}(\mathscr{C}(D), E)\}$.

Beweis. Nach dem Trennungssatz für konvexe, kompakte Teilmengen in lokalkonvexen topologischen Vektorräumen, angewendet auf den Raum $\mathbb{R}^I$, definiert eine Familie $(f_i)_{i \in I}$ in E genau dann eine Abbildung $f := (f_i)_{i \in I}$ in $\mathfrak{E}(\mathfrak{X}, D)$, falls für jede stetige Linearform L auf $\mathbb{R}^I$ mit $L \leqq \beta$ auf D (β reell) auch $L \circ f = \sum \alpha_i f_i \leqq \beta$ gilt[1]. Es folgt somit wegen

$$L \circ [(Tp_i)_{i \in I}] = \sum \alpha_i Tp_i = T(\sum \alpha_i p_i) \leqq T\beta = \beta$$

sofort $(Tp_i)_{i \in I} \in \mathfrak{E}(\mathfrak{X}, D)$.

Ferner liefert jedes $f \in \mathfrak{E}(\mathfrak{X}, D)$ aufgrund von 2.2.3 mittels $T_f g := g \circ f$ für alle $g \in \mathscr{C}(D)$ einen R-Kern T_f. Wegen $T_f p_i = p_i \circ f = f_i$ für alle $i \in I$ erhält man $f = (T_f p_i)_{i \in I}$.

3.1.3. Definition. Für jedes $T \in \mathscr{R}(\mathscr{C}(D), E)$ ist die *Risikofunktion* R_T *(des R-Kerns T)* definiert durch

$$R_T(i) := \langle Tp_i, P_i \rangle = R_{(Tp_j)_{j \in I}}(i)$$

für alle $i \in I$.

3.1.4. *Bemerkung.* Es genügt, anstelle von Elementen in $\mathscr{R}(\mathscr{C}(D), E)$ Algebrahomomorphismen von $\mathscr{C}(D)$ in E zu betrachten. Diese erweisen sich nach einem Satz von Phelps (Theorem 1.1 in [15]) als die Extremalpunkte der konvexen Menge $\mathscr{R}(\mathscr{C}(D), E)$.

3.1.5. Seien $\mathfrak{X} = (X, E, (P_i)_{i \in I})$ und $\mathfrak{Y} = (Y, F, (Q_i)_{i \in I})$ zwei Experimente mit gleicher Indexmenge I. Die ordnungstheoretischen Dualräume E' bzw. F' von E bzw. F sind sogar topologische Dualräume. Zu jedem $T \in \mathscr{R}(F, E)$ existiert der adjungierte Operator T' als Abbildung von E' in F', definiert durch $\langle g, T'\pi \rangle = \langle Tg, \pi \rangle$ für alle $g \in F$ und $\pi \in E'$. Bekanntlich ist T' linear, positiv und isometrisch.

3.1.6. Wir kommen zu einer ersten hinreichenden Bedingung dafür, daß das Experiment $\mathfrak{X}$ ε-informativer ist als das Experiment $\mathfrak{Y}$.

1. Hierbei wird nur über endlich viele $i \in I$ summiert.

Proposition. *Es existiere ein R-Kern $T\in\mathscr{R}(F,E)$ mit $\|T'P_i-Q_i\|\leqq\varepsilon(i)$ für alle $i\in I$. Dann gilt*

$$\mathfrak{X}\supset_\varepsilon\mathfrak{Y}.$$

Beweis. Zu beliebigem $D\in\mathscr{D}(I)$ und $V\in\mathscr{R}(\mathscr{C}(D),F)$ wählen wir

$$U:=T\circ V\in\mathscr{R}(\mathscr{C}(D),E).$$

Nach Voraussetzung erhalten wir wegen $\|p_i\|\leqq 1$ und $\|Vp_i\|\leqq 1$ für jedes $i\in I$ die gewünschte Ungleichung:

$$R_U(i)=\langle Up_i,P_i\rangle=\langle Vp_i,T'P_i\rangle\leqq\langle Vp_i,Q_i\rangle+\varepsilon(i)=R_V(i)+\varepsilon(i).$$

3.2. Dominierte klassische Experimente

3.2.1. Es sei $\mathfrak{X}=(X,\mathscr{L}_{\mathfrak{A}},(P_i)_{i\in I})$ ein klassisches Experiment. Die Familie $(P_i)_{i\in I}$ werde *dominiert* durch ein σ-endliches Maß μ auf $(X,\mathfrak{A})$, d.h. es gilt $P_i\ll\mu$ für alle $i\in I$. Nach Halmos und Savage ([10], Lemma 7) existiert sogar ein W-Maß P auf $(X,\mathfrak{A})$, welches zur Familie $(P_i)_{i\in I}$ *äquivalent* ist in dem Sinne, daß für $A\in\mathfrak{A}$ die Relation $P_i(A)=0$ für alle $i\in I$ genau dann gilt, wenn $P(A)=0$ ist. Wir schreiben $(P_i)_{i\in I}\sim P$. Unter dieser Voraussetzung heißt das *Experiment $\mathfrak{X}$ μ-dominiert.*

3.2.2. Für die weiteren Untersuchungen ist es wichtig, einen Typ eines *verallgemeinerten Experimentes* zuzulassen, nämlich das einem μ-dominierten Experiment $\mathfrak{X}$ zugeordnete *μ-assoziierte* (verallgemeinerte) *Experiment*

$$\mathfrak{X}(\mu):=(X,L_{\mathfrak{A}},(P_i)_{i\in I}),$$

wobei $L_{\mathfrak{A}}:=L^\infty(X,\mathfrak{A},\mu)$ den Banachverband der (Äquivalenz-)Klassen μ-wesentlich beschränkter $\mathfrak{A}$-meßbarer Funktionen auf X bezeichnet. $\mathfrak{X}(\mu)$ kann mit Experimenten (im engeren Sinne) verglichen werden (bezüglich der Relationen in 2.3), falls man für jedes $D\in\mathscr{D}(I)$ als Menge $\mathfrak{E}(\mathfrak{X}(\mu),D)$ von (nunmehr verallgemeinerten) Entscheidungsfunktionen die Menge $\mathscr{R}(\mathscr{C}(D),L_{\mathfrak{A}})$ wählt. Es gilt zunächst $\mathfrak{X}(\mu)\supset\mathfrak{X}$. Zum Beweis wende man einfach 3.1.6 auf die Restklassenabbildung von $\mathscr{L}_{\mathfrak{A}}$ auf $L_{\mathfrak{A}}$ an.

Darüber hinaus erhält man die folgende

Proposition. *Ist für das μ-dominierte Experiment $\mathfrak{X}$ die Indexmenge I höchstens abzählbar, so gilt*

$$\mathfrak{X}(\mu)\sim\mathfrak{X}.$$

Beweis. Nach der vorausgeschickten Bemerkung ist noch $\mathfrak{X}(\mu)\subset\mathfrak{X}$ zu zeigen. Es seien $D\in\mathscr{D}(I)$ sowie $U\in\mathscr{R}(\mathscr{C}(D),L_{\mathfrak{A}})$ vorgegeben. Hierzu wählen wir für jedes $i\in I$ einen Repräsentanten f_i' aus der Klasse Up_i und Folgen $(L_n)_{n\geqq 1}$ von stetigen Linearformen auf $\mathbb{R}^I$ sowie Folgen $(\alpha_n)_{n\geqq 1}$ in $\mathbb{R}$, welche D beschreiben[2]. Für jedes $n\geqq 1$ erhalten wir außerhalb einer μ-Nullmenge N_n die Ungleichung $L_n\circ(f_i')_{i\in I}\leqq\alpha_n$. Da $N:=\bigcup_{n\geqq 1}N_n$ wieder eine μ-Nullmenge ist, ergibt sich nach Abänderung von f_i' auf N zu f_i $(i\in I)$, daß $f:=(f_i)_{i\in I}\in\mathfrak{E}(\mathfrak{X},D)$ gilt. Offenbar hat man zudem $R_U=R_f$.

2. Im Sinne von 3.1.2.

Zusatz. Für jedes klassische Experiment $\mathfrak{X}$ (mit beliebiger Indexmenge I) ist $\mathfrak{X}(\mu)$ bedingt äquivalent zu $\mathfrak{X}$.

3.2.3. Bekanntlich heißt ein Tripel $(X, \mathfrak{A}, P)$ *Maßraum mit Lifting*, wenn es eine positive, normierte lineare Abbildung T von $L^{\infty}(X, \mathfrak{A}, P)$ in $\mathscr{L}_{\mathfrak{A}}$ (also ein $T \in \mathscr{R}(L^{\infty}(X, \mathfrak{A}, P), \mathscr{L}_{\mathfrak{A}}))$ gibt, so daß $T(f) \in f$ erfüllt ist. Für vollständige W-Räume oder lokalkompakte W-Räume ist die Existenz eines sogar isometrischen Liftings gesichert (vgl. [14], 154–156).

Wir nennen ein μ-dominiertes Experiment $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ *Experiment mit Lifting (L-Experiment)* (mit dominierendem Maß μ), falls $(X, \mathfrak{A}, \mu)$ ein Maßraum mit Lifting ist.

Proposition. *Für jedes L-Experiment (mit dominierendem Maß μ) gilt $\mathfrak{X}(\mu) \sim \mathfrak{X}$.*

Beweis. Ergibt sich sofort aus 3.1.6.

3.2.4. Nach Vervollständigung der σ-Algebra $\mathfrak{A}$ des μ-dominierten Experiments $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ zur σ-Algebra $\mathfrak{A}' := \mathfrak{A}^{\mu}$ gewinnt man unmittelbar das *μ-vervollständigte Experiment* $\mathfrak{X}^{\mu} := (X, \mathscr{L}_{\mathfrak{A}'}, (P_i)_{i \in I})$, welches wegen $\mathscr{L}_{\mathfrak{A}} \subset \mathscr{L}_{\mathfrak{A}'}$ informativer ist als $\mathfrak{X}$. Es gilt darüber hinaus die folgende

Proposition. $\mathfrak{X}(\mu) \sim \mathfrak{X}^{\mu}$.

Beweis. Nach 3.2.3 haben wir $\mathfrak{X}^{\mu} \sim \mathfrak{X}^{\mu}(\mu)$, wegen der Isomorphie von $L_{\mathfrak{A}}$ und $L_{\mathfrak{A}'}$ zudem $\mathfrak{X}^{\mu}(\mu) \sim \mathfrak{X}(\mu)$, also das Resultat.

3.3. Dichte Teilmengen von R-Kernen

Es soll für einen zunächst beliebigen Banachverband A reeller beschränkter Funktionen auf einer Menge (mit $1 \in A$) die Menge $\mathscr{R} := \mathscr{R}(A, L_{\mathfrak{A}})$ mit einer Topologie $\mathscr{T}$ ausgerüstet werden, so daß $\mathscr{R}$ eine konvexe, kompakte Menge wird.

3.3.1. Die Topologie $\mathscr{T}$ wird definiert als die gröbste Topologie in $\mathscr{R}$ bzw. in dem hiervon erzeugten Vektorraum $\mathscr{V}$, welche alle Abbildungen $U \to \langle U g, f \rangle$ mit $g \in A$ und $f \in L^1_{\mathfrak{A}} := L^1(X, \mathfrak{A}, \mu)$ sowie $U \in \mathscr{R}$ stetig macht.

$\mathscr{V}$ wird damit zu einem lokalkonvexen Hausdorff-Raum.

$\mathscr{R}$ ist offenbar konvex. Wegen $L_{\mathfrak{A}} = (L^1_{\mathfrak{A}})'$ sind die abgeschlossenen Kugeln von $L_{\mathfrak{A}}$ schwach kompakt, also auch die Menge

$$K := \prod_{g \in A} \{h \in L_{\mathfrak{A}} \colon \|h\| \leqq \|g\|\}.$$

$\mathscr{R}$ läßt sich außerdem mit einer abgeschlossenen Teilmenge von K identifizieren, so daß $\mathscr{R}$ kompakt wird.

Wir erhalten somit die folgende

Proposition. *$\mathscr{R}$ ist eine bezüglich der Topologie $\mathscr{T}$ kompakte, konvexe Menge.*

3.3.2. Analog zum Beweis von Theorem 1.1 in [15] zeigt man, daß die Menge $\mathscr{R}_e$ der Extremalpunkte von $\mathscr{R}$ mit der Menge der Algebrahomomorphismen von $\mathscr{R}$ übereinstimmt.

Da die Abbildung $U \to U g_1 \, U g_2$ $(g_1, g_2 \in A)$ von $\mathscr{R}$ in $L_{\mathfrak{A}}$ stetig ist, wenn man in $\mathscr{R}$ die Topologie $\mathscr{T}$ betrachtet, ist $\mathscr{R}_e$ sogar abgeschlossen.

Nach dem Satz von Krein-Milman gilt die

Proposition. *Die konvexe Hülle von $\mathcal{R}_e$ ist bezüglich der Topologie $\mathcal{T}$ dicht in $\mathcal{R}$.*

3.3.3. In Anlehnung an ein Resultat von LeCam ([11], Theorem 1) ergibt sich eine weitere dichte Teilmenge von $\mathcal{R}$ für den Fall, daß A die Menge $\mathscr{C}(S)$ aller stetigen reellen Funktionen auf einem kompakten Raum S ist.

Wir bezeichnen die Menge $\mathcal{R}(\mathscr{C}(S), L_{\mathfrak{A}})$ wieder mit $\mathcal{R}$ und definieren die Menge

$$\mathcal{R}' := \left\{ U \in \mathcal{R} \colon \; U := \sum_{k=1}^{n} h_k \, \varepsilon_{s_k}, \, s_k \in S, \, h_k \in L_{\mathfrak{A}} \right.$$
$$\left. \text{mit } h_k \geqq 0 \text{ für alle } k \geqq 1 \text{ und } \sum_{k=1}^{n} h_k = 1 \right\}.$$

Hierbei ist ε_s für $s \in S$ das Punktfunktional in s. Es werde nun $\mathcal{R}$ mit der kompakt-offenen Topologie Σ ausgerüstet. Dann gilt die

Proposition. *$\mathcal{R}'$ ist bezüglich der Topologie Σ dicht in $\mathcal{R}$.*

Speziell liefert diese Aussage für klassische Experimente $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ *mit* $|I| := n < \infty$ das

Korollar. *Für jedes $D \in \mathscr{D}(I)$, $U \in \mathcal{R}(\mathscr{C}(D), L_{\mathfrak{A}})$ und $\delta > 0$ existiert ein $V \in \mathcal{R}'$ mit der Eigenschaft*

$$\sup_{i=1,\ldots,n} |R_U(i) - R_V(i)| \leqq \delta.$$

Der *Beweis des Korollars* ergibt sich nach Anwendung der Proposition auf die Menge $K := \{p_1, \ldots, p_n\}$ unter Beachtung der Ungleichung

$$|R_U(i) - R_V(i)| \leqq \| U p_i - V p_i \| \qquad \text{für alle } i = 1, \ldots, n.$$

4. Charakterisierung der ε-Informativität

4.1. Verallgemeinertes Minimax-Theorem

Wir werden einen Satz formulieren, der Theorem 2 in [11] enthält, und einige für das weitere wichtige Folgerungen ziehen.

Es seien hierzu X ein lokalkompakter Raum, $\mathfrak{K}$ das System der kompakten Teilmengen von X, $\mathcal{M}_K := \mathcal{M}_K(X)$ die Menge aller Radonschen W-Maße auf X mit kompaktem Träger. Die Menge $\mathscr{C}(X)$ werde mit der Topologie der kompakten Konvergenz versehen und damit zu einem lokalkonvexen Hausdorff-Raum.

Für jede Teilmenge C von $\mathscr{C}(X)$ heißt die durch $\chi_C(\mu) := \inf_{f \in C} \mu(f)$ definierte numerische Funktion χ_C auf $\mathcal{M}_K$ die *charakteristische Einhüllende* von C. Betrachtet man die Menge

$$\alpha(C) := C + \mathscr{C}_+(X) = \{ g \in \mathscr{C}(X) \colon \text{Es existiert ein } f \in C \text{ mit } f \leqq g \},$$

so gilt $\chi_{\alpha(C)} = \chi_C$. Hierbei steht offenbar $\mathscr{C}_+(X)$ für die Menge $\{f \in \mathscr{C}(X) \colon f \geqq 0\}$.

Eine Teilmenge C von $\mathscr{C}(X)$ heißt *subkonvex*, wenn es zu $f, g \in C$ sowie $\alpha \in]0, 1[$ ein $h \in C$ gibt derart daß $h \leqq \alpha f + (1-\alpha) g$ gilt.

Konvexe Mengen und absteigend filtrierende Mengen sind stets subkonvex. Ferner ist für jede subkonvexe Teilmenge C von $\mathscr{C}(X)$ die Menge $\alpha(C)$ konvex.

4.1.1. Proposition. *Seien X ein lokalkompakter, Y ein kompakter Raum, f eine reellwertige Funktion auf $X \times Y$ mit den folgenden Eigenschaften:*

(i) *Die Abbildung $x \to f(x, y)$ von X in $\mathbb{R}$ ist für jedes $y \in Y$ stetig.*

(ii) *Für jedes $x \in X$ ist die Abbildung $y \to f(x, y)$ von Y in $\mathbb{R}$ nach unten halbstetig.*

Dann ist für $C_f := \{f(\cdot, y): y \in Y\}$ die Menge $\alpha(C_f)$ abgeschlossen in $\mathscr{C}(X)$.

Beweis. Wir zeigen $\overline{\alpha(C_f)} \subset \alpha(C_f)$. Seien $g \in \overline{\alpha(C_f)}$, $K \in \mathfrak{K}$ und $\gamma \in \mathbb{R}$ mit $\gamma > 0$. Dann gibt es ein $h \in \alpha(C_f)$ mit $|h(x) - g(x)| \leqq \gamma$ für alle $x \in K$ und daher ein $y = y(K, \gamma) \in Y$ mit $f(x, y) \leqq h(x) \leqq g(x) + \gamma$ für alle $x \in K$. Die nichtleere Menge

$$F_{K,\gamma} := \{y \in Y: f(x, y) \leqq g(x) + \gamma \text{ für alle } x \in K\}$$

ist wegen (ii) abgeschlossen in Y. Da Y kompakt ist, ergibt sich

$$F := \bigcap_{\substack{K \in \mathfrak{K} \\ \gamma \in \mathbb{R}, \gamma > 0}} F_{K,\gamma} \neq \emptyset.$$

Für jedes $y \in F$ gilt also $f(x, y) \leqq g(x)$ für alle $x \in X$ und damit $g \in \alpha(C_f)$.

4.1.2. Spezialfall. Setzt man in obiger Proposition $Y := C$ und $f(x, g) := g(x)$ für alle $x \in X$, so folgt für kompakte Teilmengen C von $\mathscr{C}(X)$ stets die Abgeschlossenheit von $\alpha(C)$. Es ist nur zu beachten, daß die Abbildung $(x, g) \to f(x, g)$ stetig ist und $C = C_f$ gilt.

4.1.3. Satz. *Es seien C und C' nichtleere subkonvexe Teilmengen von $\mathscr{C}(X)$. Für die Funktionen χ_C und $\chi_{C'}$, wobei $\chi_{C'}$ reellwertig ist, sind folgende Aussagen äquivalent:*

(1) $\chi_C \leqq \chi_{C'}$.

(2) $C' \subset \overline{\alpha(C)}$.

(3) *Zu jedem $f \in C'$ existiert ein $g \in \overline{\alpha(C)}$ mit $g \leqq f$.*

Beweis. [11], 1428.

4.1.4. Korollar. *Ist $\alpha(C)$ zudem abgeschlossen in $\mathscr{C}(X)$, so ist die Bedingung* (3) *des Satzes äquivalent zu* (3'). *Zu jedem $f \in C'$ existiert ein $g \in C$ mit $g \leqq f$.*

Beweis. Klar.

4.1.5. Als Spezialfall von 4.1.3 erhält man das berühmte *Minimax-Theorem*, nämlich den folgenden

Satz. *Vorgegeben seien ein lokalkonvexer Hausdorff-Raum M sowie ein Vektorraum N. X sei eine konvexe, kompakte Teilmenge von M, Y eine konvexe Teilmenge von N. Ferner bezeichne f eine reelle Funktion auf $X \times Y$ mit den folgenden Eigenschaften:*

(i) *Für jedes $y \in Y$ ist die Abbildung $x \to f(x, y)$ stetig und konkav auf X.*

(ii) *Für jedes $x \in X$ ist die Abbildung $y \to f(x, y)$ konvex auf Y.*

(iii) $\sup\limits_{x \in X} \inf\limits_{y \in Y} f(x, y) \in \mathbb{R}$.

Dann existiert mindestens ein $x_0 \in X$ mit

$$\sup_{x \in X} \inf_{y \in Y} f(x, y) = \inf_{y \in Y} \sup_{x \in X} f(x, y) = \inf_{y \in Y} f(x_0, y).$$

Beweis. Sei $\delta := \sup_{x\in X} \inf_{y\in Y} f(x,y) < \infty$.

Es ist zunächst zu zeigen: $\inf_{y\in Y} \sup_{x\in X} f(x,y) \leqq \delta$. Wir betrachten die Mengen $C = C_f := \{f(\cdot, y): y\in Y\}$ sowie $C' = \{\delta\}$. C ist wegen (ii) subkonvex. Bezeichnet x_μ für $\mu\in\mathcal{M}_K$ den Schwerpunkt von μ, so gewinnt man die Ungleichung:

$$\chi_C(\mu) := \inf_{g\in C} \mu(g) \leqq \inf_{y\in Y} f(x_\mu, y) \leqq \sup_{x\in X} \inf_{y\in Y} f(x,y) =: \chi_{C'}(\mu).$$

Nach 4.1.3(3) existiert zu $\delta\in C'$ ein $g\in\overline{\alpha(C)}$ mit $g\leqq\delta$. Zu $\varepsilon>0$ gibt es ein $y := y(\varepsilon)\in Y$ mit $f(\cdot, y)\leqq\delta+\varepsilon$, also auch $\inf_{y\in Y}\sup_{x\in X} f(x,y)\leqq\delta+\varepsilon$, woraus obige Teilbehauptung folgt.

Als nach oben halbstetige Funktion auf dem kompakten Raum X nimmt die reelle Funktion $x\to\inf_{y\in Y} f(x,y)$ ihr Supremum auf X an, etwa im Punkte $x_0\in X$.

4.2. Charakterisierungssatz

Wir greifen die Situation von 3.2 erneut auf:

$\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ sei ein klassisches und $\mathfrak{Y} = (Y, F, (Q_i)_{i\in I})$ ein beliebiges Experiment. μ bezeichne ein σ-endliches Maß auf $(X, \mathfrak{A})$. Dann gilt der folgende

4.2.1. Satz. *Vorgelegt seien ein μ-dominiertes Experiment $\mathfrak{X}$ sowie das μ-assoziierte Experiment $\mathfrak{X}(\mu)$. Für jedes Experiment $\mathfrak{Y}$ sind die folgenden Aussagen äquivalent:*

(1) $\mathfrak{X} >_\varepsilon \mathfrak{Y}$.

(2) $\mathfrak{X}(\mu) >_\varepsilon \mathfrak{Y}$.

(3) $\mathfrak{X}(\mu) \supset_\varepsilon \mathfrak{Y}$.

(4) *Es existiert ein $T\in\mathscr{R}(F, L_{\mathfrak{A}})$ mit $\|T' P_i - Q_i\| \leqq \varepsilon(i)$ für alle $i\in I$.*

Beweis. (1)$\Leftrightarrow$(2) folgt aus 3.2.2 (Zusatz), (3)$\Rightarrow$(2) aus 2.3.2, (4)$\Rightarrow$(3) aufgrund von 3.1.6.

Zu zeigen bleibt (2)$\Rightarrow$(4): Eine für den Beweis dieser Inklusion wesentliche Idee entstammt der Arbeit [11] (Beweis von Theorem 3). Als Entscheidungsmenge D betrachten wir speziell die Menge $[-1, +1]^I$. I trage die diskrete Topologie. Es werde nun auf $\mathfrak{E}(\mathfrak{Y}, D)\times\mathscr{R}(F, L_{\mathfrak{A}})\times\mathcal{M}_K(I)$ die reellwertige Funktion Φ definiert durch:

$$\begin{aligned}\Phi(g, U, \pi) &:= \int [\langle U g_i, P_i\rangle - \langle g_i, Q_i\rangle]\, \pi(di)\\ &= \int [R_{(U g_j)_{j\in I}}(i) - R_g(i)]\, \pi(di).\end{aligned}$$

Da der Träger J von π eine nichtleere *endliche* Teilmenge von I ist, existiert zu $G := (g_j)_{j\in J}\in\mathfrak{E}(\mathfrak{Y}_J, D)$ und zur abgeschlossenen konvexen Hülle D' von $G(Y)(\in\mathscr{D}(J))$ nach (2) ein $V\in\mathscr{R}(\mathscr{C}(D'), L_{\mathfrak{A}})$ mit der Eigenschaft:

$$R_V \leqq R_G + \varepsilon_J.$$

Ist $\gamma\in\mathbb{R}$, $\gamma>0$ vorgegeben, so folgt hieraus

$$R_V < R_G + \varepsilon_J + \gamma.$$

Wegen 3.3.3 (Korollar) gibt es ein $W \in \mathscr{R}'(\mathscr{C}(D'), L_{\mathfrak{A}})$ der Gestalt

$$W := \sum_{k=1}^{n} h_k \, \varepsilon_{x_k},$$

wobei $x_k \in D'$, $h_k \in L_{\mathfrak{A}}$ mit $h_k \geqq 0$ für alle $k \geqq 1$ und

$$\sum_{k=1}^{n} h_k = 1$$

ist, so daß $R_W < R_G + \varepsilon_J + \gamma$ erfüllt bleibt. Man bemerkt, daß die Menge $\{x_1, \ldots, x_n\}$ sogar in der konvexen Hülle von $G(Y)$ gewählt werden darf. Wird also für jedes $k \geqq 1$

$$x_k := \sum_{l=1}^{m} \rho_{l\,k} \, y_l$$

gesetzt, wobei $y_l \in G(Y)$ für alle l sowie $\rho_{l\,k} \in \mathbb{R}$, $\rho_{l\,k} \geqq 0$ mit $\sum_{l=1}^{m} \rho_{l\,k} = 1$ für alle l, k ($l = 1, \ldots, m$; $k = 1, \ldots, n$) gilt, so ergibt sich für alle $j \in J$ die Gleichung:

$$W p_j = \sum_{k=1}^{n} h_k \, p_j \Big(\sum_{l=1}^{m} \rho_{l\,k} \, y_l \Big) = \sum_{l=1}^{m} \Big(\sum_{k=1}^{n} \rho_{l\,k} \, h_k \Big) \, p_j(y_l).$$

Wählt man nun zu $y_l \in G(Y)$ einen Punkt $z_l \in G^{-1}(y_l)$ ($l = 1, \ldots, m$), so liefert die Definition

$$U := \sum_{l=1}^{m} f_l \, \varepsilon_{z_l} \quad \text{mit} \quad f_l := \sum_{k=1}^{n} \rho_{l\,k} \, h_k \in L_{\mathfrak{A}},$$

$f_l \geqq 0$ für alle $l = 1, \ldots, m$, ein Element von $\mathscr{R}'(F, L_{\mathfrak{A}})$. Für jedes $j \in J$ gilt:

$$U g_j = \sum_{l=1}^{m} f_l \, g_j(z_l) = \sum_{l=1}^{m} f_l \, p_j(y_l) = W p_j,$$

also

$$R_{(U g_j)_{j \in J}} = R_W < R_G + \varepsilon_J + \gamma,$$

mithin nach Restriktion auf J schließlich:

$$\Phi(g, U, \pi) < \int \varepsilon \, d\pi + \gamma$$

oder

$$\inf_{U \in \mathscr{R}(F, L_{\mathfrak{A}})} \Phi(g, U, \pi) \leqq \int \varepsilon \, d\pi.$$

Für festes $\pi \in \mathscr{M}_K(I)$ ist Φ in den beiden ersten Variablen affin-linear und außerdem in der zweiten Variablen stetig. Wir führen die Abkürzungen $\mathfrak{E} := \mathfrak{E}(\mathfrak{Y}, D)$ sowie $\mathscr{R} := \mathscr{R}(F, L_{\mathfrak{A}})$ ein, wenden 4.1.5 auf die Funktion $f := -\Phi$ an und erhalten die Existenz eines $U_0 := U_\pi \in \mathscr{R}$ mit den Eigenschaften:

$$\begin{aligned} \sup_{g \in \mathfrak{E}} \Phi(g, U_0, \pi) &= \inf_{U \in \mathscr{R}} \sup_{g \in \mathfrak{E}} \Phi(g, U, \pi) \\ &= \sup_{g \in \mathfrak{E}} \inf_{U \in \mathscr{R}} \Phi(g, U, \pi) \\ &\leqq \int \varepsilon \, d\pi. \end{aligned}$$

Zu $\gamma\in\mathbb{R}, \gamma>0$ und $i\in I$ wählen wir jetzt $g_i\in F$ mit $\|g_i\|\leqq 1$, so daß gilt:

$$(1-\gamma)\ \|U_0' P_i - Q_i\| \leqq \langle g_i, U_0' P_i - Q_i\rangle.$$

Setzt man endlich $g:=(g_i)_{i\in I}\in\mathfrak{E}(\mathfrak{Y}, D)$, so liefert Integration nach π die Abschätzung:

$$(1-\gamma)\int \|U_0' P_i - Q_i\|\ \pi(di) \leqq \int \langle g_i, U_0' P_i - Q_i\rangle\ \pi(di) = \Phi(g, U_0, \pi) \leqq \int \varepsilon\, d\pi.$$

Die Abbildung $r\colon I\times\mathscr{R}\to\mathbb{R}$, definiert durch

$$r(i, U) := \|U' P_i - Q_i\|,$$

ist als Funktion von U auf der nach 3.3.1 $\mathscr{T}$-kompakten, konvexen Menge $\mathscr{R}$ nach unten halbstetig, also $\alpha(C_r)$ nach 4.1.1 abgeschlossen, und auch subkonvex. Für $C' := \{\varepsilon\}\subset\mathscr{C}(I)$ gilt

$$\chi_{C_r} \leqq \chi_{C'}.$$

Man erhält also gemäß 4.1.4 zu $\varepsilon\in C'$ ein $T\in\mathscr{R}$ mit $r(i, T)\leqq\varepsilon(i)$ für alle $i\in I$. Dies ist aber die behauptete Aussage (4).

4.2.2. Korollar. *Für jede Funktion* $\varepsilon\colon I\to\mathbb{R}_+$ *ist die Menge* $\mathscr{N}_\varepsilon := \{U\in\mathscr{R}(F, L_{\mathfrak{A}})\colon \|U' P_i - Q_i\| \leqq \varepsilon(i)$ *für alle* $i\in I\}$ $\mathscr{T}$*-kompakt und konvex.*

Beweis. Da die Abbildung $U\to r(i, U) := \|U' P_i - Q_i\|$ von $\mathscr{R} := \mathscr{R}(F, L_{\mathfrak{A}})$ in $[0, 2]$ nach unten halbstetig ist, ist für jedes $i\in I$ die Menge $\mathscr{N}_i := \{U\in\mathscr{R}\colon r(i, U)\leqq \varepsilon(i)\}$ $\mathscr{T}$-kompakt, also auch $\mathscr{N}_\varepsilon := \bigcap_{i\in I}\mathscr{N}_i$.

Analog zeigt man die Konvexität von $\mathscr{N}_\varepsilon$.

4.2.3. Korollar. *Für die Experimente* $\mathfrak{X}$ *und* $\mathfrak{Y}$ *des Satzes* 4.2.1 *sind folgende Aussagen äquivalent:*

(1) $\mathfrak{X}(\mu)\sqsupset_\varepsilon\mathfrak{Y}$ *für jedes* $\varepsilon>0$.

(2) *Es existiert ein* $T\in\mathscr{R}(F, L_{\mathfrak{A}})$ *mit* $T' P_i = Q_i$ *für alle* $i\in I$.

Beweis. Zu zeigen bleibt (1)$\Rightarrow$(2): Wegen 4.2.1 gilt $\mathscr{N}_\varepsilon \neq \emptyset$ für jedes $\varepsilon\in\mathbb{R}, \varepsilon>0$. Nach 4.2.2 besitzt das absteigend filtrierende System $\{\mathscr{N}_\varepsilon\colon \varepsilon\in\mathbb{R}, \varepsilon>0\}$ nichtleerer kompakter Teilmengen von $\mathscr{R}(F, L_{\mathfrak{A}})$ einen nichtleeren Durchschnitt $\mathscr{N}$. Jedes $T\in\mathscr{N}$ hat die Eigenschaft $T' P_i = Q_i$ für alle $i\in I$.

Die folgenden zwei Sätze sind Folgerungen aus 4.2.1 für den Fall, daß das Experiment $\mathfrak{X}$ ein L-Experiment ist.

4.2.4. Satz. *Es seien* $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ *ein L-Experiment (mit dominierendem Maß* μ*),* $\mathfrak{Y}$ *ein beliebiges Experiment. Dann sind folgende Aussagen äquivalent:*

(1) $\mathfrak{X}\supset_\varepsilon\mathfrak{Y}$.

(2) *Es existiert ein* $T\in\mathscr{R}(F, \mathscr{L}_{\mathfrak{A}})$ *mit* $\|T' P_i - Q_i\| \leqq \varepsilon(i)$ *für alle* $i\in I$.

Beweis. Offensichtlich nach 3.2.3.

4.2.5. Satz. *Es seien* $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ *ein L-Experiment (mit dominierendem Maß* μ*),* $\mathfrak{Y}$ *ein beliebiges Experiment. Dann sind folgende Aussagen äquivalent:*

(1) $\mathfrak{X}\supset\mathfrak{Y}$.

(2) *Zu jedem* $D\in\mathscr{D}(I)$ *und jedem* $g\in\mathfrak{E}(\mathfrak{Y}, D)$ *existiert ein* $f\in\mathfrak{E}(\mathfrak{X}, D)$ *mit* $R_f = R_g$.

Beweis. Wegen 4.2.4 kann man ein $T \in \mathscr{R}(F, \mathscr{L}_{\mathfrak{A}})$ wählen mit $T' P_i = Q_i$ für alle $i \in I$. Definiert man zu $g \in \mathfrak{E}(\mathfrak{Y}, D)$ die Funktion $f := (Tg_i)_{i \in I} \in \mathfrak{E}(\mathfrak{X}, D)$, so erhält man sofort für alle $i \in I$:

$$R_f(i) = \langle f_i, P_i \rangle = \langle g_i, T' P_i \rangle = \langle g_i, Q_i \rangle = R_g(i).$$

4.3. *Stabilitätssatz*

Es soll das Problem der Konstruktion von Hilfsexperimenten untersucht werden, mittels welcher ε-Informativität auf Informativität schlechthin zurückgeführt werden kann.

4.3.1. Satz. *Es seien $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ ein L-Experiment (mit dominierendem Maß μ) und $\mathfrak{Y}$ ein beliebiges Experiment $(Y, F, (Q_i)_{i \in I})$. Dann sind folgende Aussagen äquivalent:*

(1) $\mathfrak{X} \supset_{\varepsilon} \mathfrak{Y}$.

(2) *Es existiert ein Experiment $\mathfrak{Y}^* := (Y, F, (Q_i^*)_{i \in I})$ mit den folgenden Eigenschaften:*

(i) $\|Q_i - Q_i^*\| \leqq \varepsilon(i)$ *für alle* $i \in I$.

(ii) $\mathfrak{X} \supset \mathfrak{Y}^*$.

Beweis. (1)$\Rightarrow$(2): Nach 4.2.4 existiert ein $T \in \mathscr{R}(F, \mathscr{L}_{\mathfrak{A}})$ mit $\|T' P_i - Q_i\| \leqq \varepsilon(i)$ für alle $i \in I$. Setzt man also $Q_i^* := T' P_i$ für jedes $i \in I$, so ist (i) bereits erfüllt. Trivialerweise gilt dann auch (ii), wiederum wegen 4.2.4.

(2)$\Rightarrow$(1): Die Bedingung (ii) von (2) liefert bei Spezialisierung von 4.2.4 für den Fall $\varepsilon \equiv 0$ die Existenz eines $T \in \mathscr{R}(F, \mathscr{L}_{\mathfrak{A}})$ mit $T' P_i = Q_i$ für alle $i \in I$. Aus der Abschätzung

$$\|T' P_i - Q_i\| \leqq \|T' P_i - Q_i^*\| + \|Q_i^* - Q_i\|$$

ergibt sich, erneut vermöge 4.2.4, die Behauptung (1).

4.3.2. *Bemerkung.* Bedingung (i) in Aussage (2) von 4.3.1 bedeutet wenigstens im Spezialfall strikter ε-Informativität, daß die Linearformen Q_i und Q_i^* für jedes $i \in I$ „praktisch ununterscheidbar" sind. Im Sinne der statistischen Selektionstheorie ergibt sich hieraus die Stabilität des „Prinzips der Informativität" (vgl. [11], 1451). Die Beziehungen des Satzes zur Theorie der robusten statistischen Verfahren sollen an anderer Stelle diskutiert werden.

4.4. *Informativität für Produktexperimente*

Satz. *Es seien $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ ein L-Experiment (mit dominierendem Maß μ), $\mathfrak{Y} = (Y, F, (Q_i)_{i \in I})$ ein beliebiges Experiment. Dann besteht die Relation $\mathfrak{X} \supset \mathfrak{Y}$ genau dann, wenn ein schwaches Produktexperiment $\mathfrak{Z} := (X \times Y, \mathscr{L}_{\mathfrak{A}} \otimes F, (R_i)_{i \in I})$ existiert mit den nachstehenden Eigenschaften:*

(1) $\mathfrak{X} = \mathrm{Pr}_1\, \mathfrak{Z}$ *und* $\mathfrak{Y} = \mathrm{Pr}_2\, \mathfrak{Z}$.

(2) $\mathfrak{X} \supset \mathfrak{Z}$.

Beweis. (1)$\Rightarrow$(2): Die Voraussetzung $\mathfrak{X} \supset \mathfrak{Y}$ impliziert wegen 4.2.4 die Existenz eines $T \in \mathscr{R}(F, \mathscr{L}_{\mathfrak{A}})$ mit $T' P_i = Q_i$ für alle $i \in I$. Zu jedem $i \in I$ definieren wir eine positive, normierte Linearform R_i auf dem Tensorprodukt G_0 der Vektorräume $\mathscr{L}_{\mathfrak{A}}$ und F durch

$$\langle g, R_i \rangle := \sum_{j=1}^{n} \langle u_j T v_j, P_i \rangle$$

für alle $g \in G_0$ der Gestalt

$$g := \sum_{j=1}^{n} u_j v_j \qquad (u_j \in \mathscr{L}_{\mathfrak{A}}, v_j \in F; j = 1, \ldots, n).$$

Bekanntlich läßt sich R_i zu einer positiven, normierten Linearform auf $G := \mathscr{L}_{\mathfrak{A}} \otimes F$ eindeutig fortsetzen; die Fortsetzung heiße wiederum R_i $(i \in I)$.

Offenbar gelten die beiden Aussagen in (1). Zu zeigen bleibt (2): Hierzu definieren wir eine lineare Abbildung M von G in $\mathscr{L}_{\mathfrak{A}}$ durch

$$M g = M\left(\sum_{j=1}^{n} u_j v_j\right) := \sum_{j=1}^{n} u_j T v_j \, (g \in G_0).$$

M ist positiv und normiert, wie man leicht sieht, also ein Element von $\mathscr{R}(G, \mathscr{L}_{\mathfrak{A}})$. Es folgt für jedes $i \in I$ und alle $g \in G$ sofort:

$$\langle g, M' P_i \rangle = \langle M g, P_i \rangle = \sum_{j=1}^{n} \langle u_j T v_j, P_i \rangle = \langle g, R_i \rangle,$$

also $M' P_i = R_i$ für alle $i \in I$. Erneute Anwendung von 4.2.4 liefert die Behauptung.

(2)$\Rightarrow$(1): Angenommen es existiert ein schwaches Produktexperiment $\mathfrak{Z} := (X \times Y, G, (R_i)_{i \in I})$ mit $G := \mathscr{L}_{\mathfrak{A}} \otimes F$ und den Eigenschaften (1) und (2). Das heißt insbesondere: Es gibt einen Kern $M \in \mathscr{R}(G, \mathscr{L}_{\mathfrak{A}})$ mit $M' P_i = R_i$ sowie $P_i = \mathrm{pr}_E R_i$ und $Q_i = \mathrm{pr}_F R_i$ für alle $i \in I$. Mit q werde die Adjungierte von pr_F bezeichnet. Dann setze man $T := M \circ q$. Offenbar gilt $T \in \mathscr{R}(F, \mathscr{L}_{\mathfrak{A}})$ und auch

$$T' P_i = (M \circ q) P_i = (q' \circ M') P_i = q'(M' P_i) = q' R_i = \mathrm{pr}_F R_i = Q_i \qquad \text{für alle } i \in I.$$

Wegen 4.2.4 ergibt sich also $\mathfrak{X} \supset \mathfrak{Y}$.

Bemerkung. Es ist möglich, hinreichende Bedingungen insbesondere an die Familie $(Q_i)_{i \in I}$ zu stellen, so daß die im schwachen Produktexperiment $\mathfrak{Z}$ auftretenden Linearformen R_i sogar *W-Maße* auf der Produkt-σ-Algebra $\mathfrak{A} \otimes \mathfrak{B}$ sind $(i \in I)$, wobei $\mathfrak{B}$ die von dem Banachverband F erzeugte σ-Algebra in Y bezeichnet. Der Beweis kann aus [11], 1443–1444 auf die vorliegende Situation übertragen werden.

5. Informativität und Erschöpftheit

5.1. Erschöpftheit für Experimente

5.1.1. Es seien $(X, \mathfrak{A})$ und $(Y, \mathfrak{B})$ zwei Meßräume und T eine Abbildung von $X \times \mathfrak{B}$ in $\mathbb{R}$. Dann heißt T ein *Markoff-Kern* von $(X, \mathfrak{A})$ nach $(Y, \mathfrak{B})$, wenn die Abbildung $B \to T(x, B)$ für jedes $x \in X$ ein *W*-Maß auf $\mathfrak{B}$ und die Abbildung

$x \to T(x, B)$ $\mathfrak{A}$-meßbar ist für alle $B \in \mathfrak{B}$. Ist μ ein W-Maß auf $(X, \mathfrak{A})$, so definiert man durch die Festsetzung

$$T\mu(B) := \int T(x, B)\, \mu(dx)$$

ein W-Maß auf $\mathfrak{B}$. Ferner ist für jedes $f \in \mathscr{L}_{\mathfrak{B}}$ die Abbildung Tf von X in $\mathbb{R}$, definiert durch

$$Tf(x) := \int f(y)\, T(x, dy),$$

ein Element von $\mathscr{L}_{\mathfrak{A}}$.

Ist speziell X eine konvexe, kompakte Teilmenge eines lokal-konvexen Hausdorff-Raums und $\mathscr{B}$ die Borel-σ-Algebra über X, so heißt ein Markoff-Kern T von $(X, \mathscr{B})$ in sich (auf $(X, \mathscr{B})$) eine *Dilatation*, wenn $Tf = f$ gilt für alle affin-linearen reellen Funktionen auf X. Eine Dilatation T (gelegentlich auch Diffusion genannt) kann auch als meßbare Abbildung von X in die Menge $\mathscr{M}_1 := \mathscr{M}_1(X)$ der Borelschen W-Maße auf X aufgefaßt werden mit der folgenden Eigenschaft: Für jedes $y \in X$ ist der Schwerpunkt $x_{T(y)}$ des Maßes $T(y)$ gleich y. Also ist T eine Abbildung der Menge $\{\varepsilon_y \colon y \in X\}$ in $\mathscr{M}_1$ mit $T\varepsilon_y = T(y)$ und kann somit zu einer Abbildung von $\mathscr{M}_1$ in $\mathscr{M}_1$ fortgesetzt werden.

5.1.2. Definition. Vorgegeben seien zwei klassische Experimente $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ und $\mathfrak{Y} = (Y, \mathscr{L}_{\mathfrak{B}}, (Q_i)_{i \in I})$. $\mathfrak{X}$ heißt *erschöpfend* für $\mathfrak{Y}$ $(\mathfrak{X} \succ \mathfrak{Y})$, wenn es einen Markoff-Kern T von $(X, \mathfrak{A})$ nach $(Y, \mathfrak{B})$ gibt mit

$$TP_i = Q_i \qquad \text{für alle } i \in I.$$

5.1.3. Wir wollen einige *Spezialfälle* von Paaren von Experimenten kurz diskutieren:

(a) Zu $\mathfrak{Y} = (Y, \mathscr{L}_{\mathfrak{B}}, (Q_i)_{i \in I})$ betrachte man das Experiment $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$, wobei für jedes $i \in I$ das Maß P_i so definiert ist: Es existiert ein Markoff-Kern M von $(Y, \mathfrak{B})$ nach $(X, \mathfrak{A})$ mit $P_i = MQ_i$ für alle $i \in I$.

(b) Zu $\mathfrak{Y} = (Y, \mathscr{L}_{\mathfrak{B}}, (Q_i)_{i \in I})$ betrachte man das Experiment $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$, wobei für jedes $i \in I$ das W-Maß P_i definiert ist als Bildmaß $M(Q_i)$ von Q_i unter einer meßbaren Abbildung M von $(Y, \mathfrak{B})$ in $(X, \mathfrak{A})$.

(c) $\mathfrak{Y} = (Y, \mathscr{L}_{\mathfrak{B}}, (Q_i)_{i \in I})$ sei ein vorgegebenes Experiment. Ferner sei $\mathfrak{A}$ eine σ-Unteralgebra von $\mathfrak{B}$.

Dann gewinnt man als zweites Experiment das Tripel

$$\mathfrak{X} = (Y, \mathscr{L}_{\mathfrak{A}}, (\operatorname{Res}_{\mathfrak{A}} Q_i)_{i \in I}).$$

Die hier aufgeführten Spezialfälle treten in der Arbeit [7] von Csiszár für den Fall $|I| = 1$ als die drei Typen *indirekter Beobachtungen* auf.

Die Fälle (b) und (c) können im wesentlichen als identisch angesehen werden. Man hat etwa in (c) nur die Identität von Y zu betrachten, um (b) zu erhalten.

Im Fall (b) hingegen kann der Meßraum $(X, \mathfrak{A})$ durch den maß-isomorphen Meßraum $(Y, M^{-1}(\mathfrak{A}))$ ersetzt werden, um (c) zu gewinnen.

Diese beiden Spezialfälle sind die in der mathematischen Statistik am häufigsten auftretenden. Der Fall (a) dagegen entstammt der Informationstheorie. Man spricht im Falle $|I| = 1$ von einem Beobachtungskanal (Y, M, X).

Um den Fall (b) und damit auch (c) als Sonderfälle von (a) zu erkennen, setzt man in (b) einfach

$$M(y,A):=\varepsilon_{M(y)}=\begin{cases}1 & \text{für } M(y)\in A\\ 0 & \text{sonst}\end{cases}$$

$(y\in Y, A\in\mathfrak{A})$.

Andererseits kann (a) in die Form von (c) gebracht werden. Man betrachte nämlich anstelle des Experiments $\mathfrak{Y}$ das schwache Produktexperiment $\mathfrak{Z}:=(Y\times X, \mathscr{L}_{\mathfrak{B}}\otimes\mathscr{L}_{\mathfrak{A}}, (R_i)_{i\in I})$, wobei R_i für jedes $i\in I$ zunächst definiert ist durch

$$R_i(B\times A):=\int_B M(y,A)\,Q_i(dy)$$

für alle $B\in\mathfrak{B}$ und $A\in\mathfrak{A}$ und schließlich fortgesetzt werden kann auf $\mathfrak{B}\otimes\mathfrak{A}$. Setzt man $\mathfrak{A}_0:=\{Y\times A\colon A\in\mathfrak{A}\}$, so entspricht für jedes $i\in I$ $\operatorname{Res}_{\mathfrak{A}_0}R_i$ gerade dem Maß P_i (im Sinne isomorpher Maßräume).

5.1.4. Der in 5.1.2 definierte Begriff der Erschöpftheit eines Experiments $\mathfrak{X}$ für ein Experiment $\mathfrak{Y}$ nimmt in obigen Spezialfällen (a), (b) und (c), falls $\mathfrak{Y}$ als μ-dominiert und topologisch vorausgesetzt wird, die folgende Gestalt an:

Zu jedem $B\in\mathfrak{B}$ existiert eine $\mathfrak{A}$-meßbare Funktion φ_B auf X, so daß für alle $A\in\mathfrak{A}$ und alle $i\in I$ gilt:

$$\int_A \varphi_B(x)\,P_i(dx)=\eta_i(A,B), \tag{E}$$

wobei $\eta_i(A,B)$ entsprechend den Fällen (a), (b) und (c) durch $R_i(B\times A)$, $Q_i(B\cap M^{-1}(A))$ und $Q_i(A\cap B)$ zu ersetzen ist.

Die Implikation $(\mathrm{E})\Rightarrow\mathfrak{X}\succ\mathfrak{Y}$ ergibt sich nach bekannten Schlüssen[3], da $\mathfrak{Y}$ topologisch ist. Den Beweis der Umkehrung, für den übrigens die Voraussetzung, daß $\mathfrak{Y}$ topologisch ist, entbehrt werden kann, führen wir mit Hilfe des Begriffs der *f-Divergenz* durch, welcher auf Csiszár [6] zurückgeht:

Da nach Halmos-Savage paarweise Erschöpftheit (bezüglich der Familie $(Q_i)_{i\in I}$) Erschöpftheit impliziert ([10], Theorem 3), genügt es, den Fall $|I|=2$ zu betrachten.

Wir benötigen einige *Vorbereitungen* aus [7] und [6]:

Es sei $(Z,\mathfrak{C})$ zunächst ein beliebiger Meßraum. Auf $(Z,\mathfrak{C})$ sei ein Paar $\{\mu_1,\mu_2\}$ von W-Maßen gegeben, welches durch ein σ-endliches Maß ν auf $(Z,\mathfrak{C})$ dominiert werde. Es gilt $\mu_i=q_i\,\nu$, wobei q_i die Dichte von μ_i bezüglich ν bezeichnet $(i=1,2)$.

Ferner sei f eine streng konvexe reelle Funktion auf $]0,\infty[$ mit

$$f(0):=\lim_{u\to+0} f(u)<\infty \quad\text{und}\quad \lim_{u\to\infty}\frac{f(u)}{u}<\infty.$$

Dann ist das Integral

$$\mathscr{I}_f(\mu_1,\mu_2):=\int q_2(z)\,f\left[\frac{q_1(z)}{q_2(z)}\right]\nu(dz)$$

wohldefiniert, und es gilt $\mathscr{I}_f(\mu_1,\mu_2)\geqq f(1)$.

3. Vgl. etwa Beweis in 5.2.2.

Csiszár zeigt in [7], 310 bzw. [6], Satz 1 sowie Folgerungen 3 und 4, daß für Paare von Experimenten vom Typ (a), (b) und (c) (bei $|I|=2$) stets die Ungleichung

$$\mathscr{I}_f(Q_1, Q_2) \geqq \mathscr{I}_f(P_1, P_2) \tag{$*$}$$

erfüllt ist, und daß

$$\mathscr{I}_f(Q_1, Q_2) = \mathscr{I}_f(P_1, P_2) \tag{$**$}$$

genau dann gilt, wenn (E) zutrifft.

Liegt nun eine der Situationen (a), (b) und (c) vor, so haben wir einerseits ($*$), andererseits aber auch

$$\mathscr{I}_f(Q_1, Q_2) \leqq \mathscr{I}_f(P_1, P_2),$$

da es nach Voraussetzung einen Markoff-Kern T von $(X, \mathfrak{A})$ nach $(Y, \mathfrak{B})$ gibt mit $TP_i = Q_i$ für $i=1,2$ und somit ($*$) für den Fall (a) aufgeschrieben werden kann.

Damit ist ($**$) erfüllt, und es gilt (E).

Offenbar stimmen die drei Forderungen unter (E) mit den Definitionen von erschöpfendem Kanal, erschöpfender Statistik und erschöpfender σ-Algebra überein.

Das hier erörterte Äquivalenz-Problem wird in etwas weitergefaßtem Rahmen auch in [17] behandelt*.

5.2. *Charakterisierung der Erschöpftheit*

Bei einem Vergleich der Erschöpftheitsrelation mit der zur Informativität äquivalenten Formulierung, die sich aus 4.2.4 (2) ergibt, drängt sich die Frage auf, unter welchen Bedingungen aus einem R-Kern ein Markoff-Kern gewonnen werden kann und umgekehrt. Damit gleichwertig ist die Frage nach Voraussetzungen über zwei klassische Experimente $\mathfrak{X}$ und $\mathfrak{Y}$, unter denen Informativität und Erschöpftheit äquivalent sind. Eine Antwort liefert der folgende

5.2.1. Satz. *Es seien $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ ein L-Experiment (mit dominierendem Maß μ) und $\mathfrak{Y} = (Y, \mathscr{L}_{\mathfrak{B}}, (Q_i)_{i\in I})$ ein topologisches Experiment. Dann sind nachstehende Aussagen äquivalent:*

(1) $\mathfrak{X} \supset \mathfrak{Y}$.

(2) $\mathfrak{X} \succ \mathfrak{Y}$.

Beweis. Die Implikation (2)$\Rightarrow$(1) ergibt sich sofort aus 3.1.6, wenn man 5.1.1 beachtet. Denn jeder Markoff-Kern T von $(X, \mathfrak{A})$ nach $(Y, \mathfrak{B})$ mit $TP_i = Q_i$ für alle $i \in I$ läßt sich zu einer positiven, normierten linearen Abbildung von $\mathscr{L}_{\mathfrak{B}}$ in $\mathscr{L}_{\mathfrak{A}}$ mit $T' P_i = Q_i$ fortsetzen.

Um die Umkehrung (1)$\Rightarrow$(2) zu beweisen, müssen wir gemäß 4.2.4 von einem R-Kern $T \in \mathscr{R}(\mathscr{L}_{\mathfrak{B}}, \mathscr{L}_{\mathfrak{A}})$ mit $T' P_i = Q_i$ für alle $i \in I$ ausgehen und zeigen, daß ein Markoff-Kern M von $(X, \mathfrak{A})$ nach $(Y, \mathfrak{B})$ existiert, so daß für alle $i \in I$ gilt: $M P_i = Q_i$.

Wir definieren zu jedem $x \in X$ eine positive Linearform T_x auf $\mathscr{L}_{\mathfrak{B}}$ durch die Festsetzung:

$$T_x(g) := (Tg)(x) \qquad \text{für alle } g \in \mathscr{L}_{\mathfrak{B}}.$$

---

* *Zusatz bei Korrektur.* Genauer heißt dies, daß man das soeben bewiesene Resultat auch mittels eines Ergebnisses aus [17] gewinnen kann. Dies wurde in der Note: „Zum Erschöpftheitsbegriff von D. Blackwell", welche in Metrika erscheinen wird, ausgeführt. Es konnte zudem gezeigt werden, daß die Regularitätsvoraussetzungen (an $\mathfrak{Y}$) nicht ersatzlos fortgelassen werden können.

Die Abbildung $x \to T_x(g)$ von X in R ist nach Definition von T $\mathfrak{A}$-meßbar und beschränkt. Es gilt ferner $0 \leqq T_x(g) \leqq 1$ für alle $g \in \mathscr{L}_{\mathfrak{B}}$ mit $0 \leqq g \leqq 1$.

Als positive Linearform auf dem Teilraum $\mathscr{K}(Y)$ aller stetigen reellen Funktionen auf Y mit kompaktem Träger von $\mathscr{L}_{\mathfrak{B}}$ kann T_x somit als positives Radon-Maß auf Y aufgefaßt werden. Zu zeigen bleibt die Normiertheit des Maßes T_x, d.h. die Gleichung $T_x(1)=1$, wenigstens nach Abänderung auf einer Nullmenge von X.

Hierzu sei $(g_n)_{n \geqq 1}$ eine isotone Folge in $\mathscr{K}_+(Y)$ (d.h. in $\mathscr{K}(Y)$ mit $g_n \geqq 0$ für alle $n \geqq 1$), so daß $\sup\limits_{n \geqq 1} g_n = 1$ gilt.

Dann erhalten wir für jedes $i \in I$:

$$\int T_x(1)\, P_i(dx) = \sup_{n \geqq 1} \int T_x(g_n)\, P_i(dx)$$

und weiterhin wegen $T' P_i = Q_i$ sogar

$$\int T_x(1)\, P_i(dx) = \sup_{n \geqq 1} \int g_n\, dQ_i = \int 1\, dQ_i = 1.$$

Da die Familie $(P_i)_{i \in I}$ als μ-dominiert vorausgesetzt wurde, existiert nach Halmos-Savage ([10], Lemma 7) ein W-Maß P mit $(P_i)_{i \in I} \sim P$. Nach obiger Gleichung ist also die Menge $N := \{x \in X\colon\ T_x(1) < 1\}$ für jedes $i \in I$ eine P_i-Nullmenge und somit eine P-Nullmenge. Wir definieren den gesuchten Markoff-Kern M durch

$$M(x, \cdot) = \begin{cases} T_x & \text{für } x \in \complement N \\ \varepsilon_y & \text{für } x \in N, \end{cases}$$

wobei y beliebig in Y gewählt werden kann.

Für jedes $i \in I$ stimmen die Maße MP_i und Q_i auf $\mathscr{K}(Y)$ überein, es gilt also $MP_i = Q_i$.

5.2.2. Um Aussage (1) in 5.2.1 noch mit der Theorie der erwartungstreuen Schätzungen in Beziehung zu bringen, stellen wir eine diesbezügliche hierzu äquivalente Aussage bereit: Es seien ein Meßraum $(X, \mathfrak{A})$ sowie eine Familie $(P_i)_{i \in I}$ von W-Maßen auf $(X, \mathfrak{A})$ (also ein klassisches Experiment) gegeben. Jede Abbildung φ von I in $\mathbb{R}$ heißt *Parameter. Schätzfunktion (Schätzung)* auf X für (den Parameter) φ wird sodann jede meßbare Abbildung s von X in $\mathbb{R}$ genannt. Eine Schätzung s für φ heißt *erwartungstreu*, wenn gilt:

(i) s ist P_i-integrierbar für alle $i \in I$,

(ii) $E_i(s) := \int s\, dP_i = \varphi(i)$ für alle $i \in I$.

Die konvexe Menge der erwartungstreuen Schätzungen auf X für φ werde mit $\mathscr{S}(\varphi)$, die ihrer nichtnegativen Elemente mit $\mathscr{S}_+(\varphi)$ bezeichnet.

Die Familie $(P_i)_{i \in I}$ von W-Maßen auf $(X, \mathfrak{A})$ heißt *vollständig*, falls für jede $\mathfrak{A}$-meßbare reelle Funktion f auf X gilt:

$$\int f\, dP_i = 0 \text{ für alle } i \in I \Rightarrow f = 0$$

P_i – f.s. für alle $i \in I$.

Satz. *Vorgegeben seien ein L-Experiment* $\mathfrak{X}=(X,\mathscr{L}_{\mathfrak{A}},(P_i)_{i\in I})$ *mit vollständiger Familie* $(P_i)_{i\in I}$ *sowie ein topologisches Experiment* $\mathfrak{Y}=(Y,\mathscr{L}_{\mathfrak{B}},(Q_i)_{i\in I})$. *Dann sind folgende Aussagen äquivalent:*

(1) $\mathfrak{X}\supset\mathfrak{Y}$.

(2) *Für jedes* $i\in I$ *und festes* $B\in\mathfrak{B}$ *existiert ein* $s\in\mathscr{S}_+(Q_i(B))$.[4]

Beweis. (1)$\Rightarrow$(2) ist evident; denn gemäß 5.2.1 existiert ein Markoff-Kern T von $(X,\mathfrak{A})$ nach $(Y,\mathfrak{B})$ mit $TP_i=Q_i$ für alle $i\in I$, d.h. aber

$$Q_i(B)=\int T(x,B)\,P_i(dx) \qquad \text{für alle } i\in I$$

($B\in\mathfrak{B}$). Dann ist für jedes $i\in I$ und festes $B\in\mathfrak{B}$ in der Tat $T\in\mathscr{S}_+(Q_i(B))$.

(2)$\Rightarrow$(1): Es werde angenommen: Für jedes $B\in\mathfrak{B}$ existiert ein $s_B\in\mathscr{S}_+(Q_i(B))$ für alle $i\in I$, also

$$E_i^B(s):=\int s_B(x)\,P_i(dx)=Q_i(B). \tag{$*$}$$

Nach 5.2.1 genügt es zu zeigen, daß die Funktion s auf $X\times\mathfrak{B}$, definiert durch $s(x,B):=s_B(x)$, in $(*)$ ersetzt werden kann durch einen Markoff-Kern T von $(X,\mathfrak{A})$ nach $(Y,\mathfrak{B})$.

Zunächst folgt wegen der Vollständigkeit von $(P_i)_{i\in I}$ aus $(*)$, daß für s P_i-f.s. ($i\in I$) gilt:

(i) $s(x,\emptyset)=0$ sowie $s(x,Y)=1$.

(ii) $0\leqq s(x,B)\leqq 1$ für alle $B\in\mathfrak{B}$.

(iii) $s(x,\bigcup_{j\geqq 1}B_j)=\sum_{j\geqq 1}s(x,B_j)$ für jede Folge $(B_j)_{j\geqq 1}$ paarweise disjunkter Mengen in $\mathfrak{B}$.

Hierbei hängen die P_i-Ausnahmemengen noch von den Mengen B sowie B_j ($j\geqq 1$) ab. Damit liegt eine Situation vor, die der im Beweis der Existenz der (regulären) bedingten W-Verteilung analog ist: Da Y ein lokalkompakter Raum mit abzählbarer Basis ist, existiert zu jedem $i\in I$ eine P_i-Nullmenge, außerhalb derer (i) bis (iii) für alle B sowie B_j ($j\geqq 1$) erfüllt sind. Nun ist aber $\mathfrak{X}$ ein L-Experiment und damit speziell dominiert durch ein σ-endliches Maß auf $(X,\mathfrak{A})$. Also existiert nach Halmos-Savage ([10], Lemma 7) ein W-Maß P mit $(P_i)_{i\in I}\sim P$ und also eine P-Nullmenge N, so daß die Eigenschaften (i) bis (iii) für alle $x\in\complement N$ erfüllt sind. Schließlich läßt sich s zu einem Markoff-Kern T von $(X,\mathfrak{A})$ nach $(Y,\mathfrak{B})$ abändern, für welchen gilt:

$$TP_i=\int T(x,B)\,P_i(dx)=Q_i(B)$$

für alle $B\in\mathfrak{B}$ und alle $i\in I$. Wegen 5.2.1 folgt die Behauptung.

5.2.3. *Bemerkung.* Interessiert man sich in 5.2.2 nur für die Implikation (2)$\Rightarrow$(1), so kann man auf die Voraussetzung der Vollständigkeit der Familie $(P_i)_{i\in I}$ verzichten, falls man anstelle von (2) die folgende Aussage betrachtet:

(2′) Für jedes $i\in I$ und festes $B\in\mathfrak{B}$ existiert *genau* ein $s\in\mathscr{S}_+(Q_i(B))$.

---

4. Genauer müßte anstelle von $Q_i(B)$ die Abbildung $i\to Q_i(B)$ stehen.

5.3. *Erschöpftheit für Produktexperimente*

Es seien für $i=1,2$ $(X_i, \mathfrak{A}_i)$ sowie $(Y_i, \mathfrak{B}_i)$ Meßräume und A_i Markoff-Kerne von $(X_i, \mathfrak{A}_i)$ nach $(Y_i, \mathfrak{B}_i)$. Betrachtet man den Meßraum $(X_1 \times X_2, \mathfrak{B}_1 \otimes \mathfrak{B}_2)$, so definiert man den *Produktkern* $A_1 \otimes A_2$ von $(X_1 \times X_2, \mathfrak{A}_1 \otimes \mathfrak{A}_2)$ nach $(Y_1 \times Y_2, \mathfrak{B}_1 \otimes \mathfrak{B}_2)$, indem man für jedes Paar $(x_1, x_2) \in X_1 \times X_2$ $A_1 \otimes A_2((x_1, x_2), \cdot)$ als das Produktmaß von $A_1(x_1, \cdot)$ mit $A_2(x_2, \cdot)$ auf $\mathfrak{B}_1 \otimes \mathfrak{B}_2$ auffaßt. Die Definition findet ihre Rechtfertigung in dem Resultat, daß die Abbildung $A_1 \otimes A_2$ von $(X_1 \times X_2) \times (\mathfrak{B}_1 \otimes \mathfrak{B}_2)$ in $\mathbb{R}$ ein Markoff-Kern ist von $(X_1 \times X_2, \mathfrak{A}_1 \otimes \mathfrak{A}_2)$ nach $(Y_1 \times Y_2, \mathfrak{B}_1 \otimes \mathfrak{B}_2)$.

5.3.1. Satz. *Für jedes $k=1, \ldots, n$ seien $\mathfrak{X}_k$ bzw. $\mathfrak{Y}_k$ Experimente $(X_k, \mathscr{L}_{\mathfrak{A}_k}, (P_i^k)_{i \in I})$ bzw. $(Y_k, \mathscr{L}_{\mathfrak{B}_k}, (Q_i^k)_{i \in I})$. Gilt nun $\mathfrak{X}_k \succ \mathfrak{Y}_k$ für alle $k=1, \ldots, n$, so auch*

$$\mathfrak{X}_1 \otimes \cdots \otimes \mathfrak{X}_n \succ \mathfrak{Y}_1 \otimes \cdots \otimes \mathfrak{Y}_n .$$

Beweis. Es genügt, den Fall $n=2$ zu betrachten. Nach Voraussetzung existiert für jedes $k=1,2$ ein Markoff-Kern T^k von $(X_k, \mathfrak{A}_k)$ nach $(Y_k, \mathfrak{B}_k)$ mit $T^k P_i^k = Q_i^k$ für alle $i \in I$. Wir betrachten die Produktexperimente $\mathfrak{X}_1 \otimes \mathfrak{X}_2$ und $\mathfrak{Y}_1 \otimes \mathfrak{Y}_2$. Zu zeigen ist die Existenz eines Markoff-Kerns T von $(X_1 \times X_2, \mathfrak{A}_1 \otimes \mathfrak{A}_2)$ nach $(Y_1 \times Y_2, \mathfrak{B}_1 \otimes \mathfrak{B}_2)$ mit der Eigenschaft:

$$T(P_i^1 \otimes P_i^2) = Q_i^1 \otimes Q_i^2 \qquad \text{für alle } i \in I. \tag{$*$}$$

Als den gesuchten Markoff-Kern T nehme man den Produktkern $T^1 \otimes T^2$, definiert durch:

$$T(\cdot, B_1 \times B_2) := T^1(\cdot, B_1) \cdot T^2(\cdot, B_2)$$

für alle $B_1 \in \mathfrak{B}_1$ und $B_2 \in \mathfrak{B}_2$.

Um die Forderung $(*)$ zu bestätigen, genügt es, Mengen $C \in \mathfrak{B}_1 \otimes \mathfrak{B}_2$ der Gestalt $B_1 \times B_2$ zu betrachten. Damit erhalten wir für jedes $i \in I$ die Gleichung:

$$\begin{aligned}
Q_i^1 \otimes Q_i^2 (B_1 \times B_2) &= Q_i^1(B_1) \cdot Q_i^2(B_2) \\
&= \int T^1(x_1, B_1)\, P_i^1(dx_1) \int T^2(x_2, B_2)\, P_i^2(dx_2) \\
&= \iint T((x_1, x_2), B_1 \times B_2)\, P_i^1(dx_1)\, P_i^2(dx_2) \\
&= \int_{X_1 \times X_2} T((x_1, x_2), B_1 \times B_2)\, P_i^1 \otimes P_i^2(d(x_1, x_2)) \\
&= T(P_i^1 \otimes P_i^2)(B_1 \times B_2),
\end{aligned}$$

woraus per definitionem $\mathfrak{X}_1 \otimes \mathfrak{X}_2 \succ \mathfrak{Y}_1 \otimes \mathfrak{Y}_2$ folgt.

5.3.2. Korollar. *Für jedes $k=1, \ldots, n$ seien $\mathfrak{X}_k$ bzw. $\mathfrak{Y}_k$ L-Experimente bzw. topologische Experimente. Dann folgt aus $\mathfrak{X}_k \supset \mathfrak{Y}_k$ für alle $k=1, \ldots, n$ unmittelbar:*

$$\mathfrak{X}_1 \otimes \cdots \otimes \mathfrak{X}_n \supset \mathfrak{Y}_1 \otimes \cdots \otimes \mathfrak{Y}_n .$$

Beweis. Klar nach 5.2.1.

5.4. *Folgerungen für Kontraktionsexperimente*

Wir werden nun Sätze für Kontraktionsexperimente beweisen, welche alle bekannten Formulierungen der Erschöpftheit von σ-Algebren enthalten. Insbesondere soll die Rolle der Voraussetzung der μ-Dominanz für das vorgegebene Experiment diskutiert werden.

5.4.1. Ausganspunkt der Überlegungen sind zwei klassische Experimente $\mathfrak{X}=(X,\mathscr{L}_{\mathfrak{A}},(P_i)_{i\in I})$ und $\mathfrak{Y}=(X,\mathscr{L}_{\mathfrak{B}},(Q_i)_{i\in I})$ mit gleichem Grundraum X (und gleicher Indexmenge I). Wir setzen ferner voraus, daß $\mathfrak{A}$ eine σ-Unteralgebra von $\mathfrak{B}$ ist und daß für jedes $i\in I$ gilt: $P_i=\operatorname{Res}_{\mathfrak{A}} Q_i$. $\mathfrak{X}$ ist also $\mathfrak{A}$-Kontraktion von $\mathfrak{Y}$, gemäß 2.1.2 gilt also $\mathfrak{X}=\mathfrak{Y}^{\mathfrak{A}}$. (Man vergleiche hierzu Typ (c) unter 5.1.3.) Mit $\mathfrak{F}$ werde die kleinste σ-Algebra über I bezeichnet, so daß alle Funktionen $i\to Q_i(f)$ für alle $f\in\mathscr{L}_{\mathfrak{B}}$ $\mathfrak{F}$-meßbar sind. Zu jedem W-Maß π auf dem Meßraum $(I,\mathfrak{F})$ werde λ_π definiert als W-Maß auf $\mathfrak{B}\otimes\mathfrak{F}$ durch die Festsetzung:

$$\lambda_\pi(B\times C):=\int_C Q_i(B)\,\pi(di)$$

für alle $B\in\mathfrak{B}$ und $C\in\mathfrak{F}$.

$\mathfrak{A}'$, $\mathfrak{B}'$ und $\mathfrak{F}'$ bezeichnen die σ-Algebren der $\mathfrak{A}$-, $\mathfrak{B}$- und $\mathfrak{F}$-Zylindermengen auf $X\times I$.

Satz. *Es seien $\mathfrak{Y}=(X,\mathscr{L}_{\mathfrak{B}},(Q_i)_{i\in I})$ ein μ-dominiertes topologisches Experiment, $\mathfrak{A}$ eine σ-Unteralgebra von $\mathfrak{B}$ und $\mathfrak{Y}^{\mathfrak{A}}$ die $\mathfrak{A}$-Kontraktion von $\mathfrak{Y}$. Dann sind die folgenden Aussagen äquivalent:*

(1) $\mathfrak{Y}^{\mathfrak{A}}\supset\mathfrak{Y}$.

(2) $\mathfrak{Y}^{\mathfrak{A}}\succ\mathfrak{Y}$.

(3) *$\mathfrak{A}$ ist erschöpfende σ-Unteralgebra von $\mathfrak{B}$ bezüglich der Familie $(P_i)_{i\in I}$, d.h.: Zu jedem $B\in\mathfrak{B}$ existiert eine $\mathfrak{A}$-meßbare reelle Funktion φ_B auf X mit*

$$Q_i^{\mathfrak{A}}(B)=\varphi_B \qquad Q_i-f.s. \text{ für alle } i\in I.$$

(4) *Für jedes W-Maß π auf $(I,\mathfrak{F})$ mit endlichem Träger sind die σ-Algebren $\mathfrak{F}'$ und $\mathfrak{B}'$ bedingt unabhängig bezüglich $\mathfrak{A}'$ für das Maß λ_π (im Sinne von* [14], 30*).*

(5) *Zu jedem $C'\in\mathfrak{F}'$ existiert eine $\mathfrak{A}'$-meßbare reelle Funktion φ_C auf $X\times I$ mit*

$$\lambda_\pi^{\mathfrak{B}'}(C')=\varphi_{C'} \qquad \lambda_\pi-f.s.$$

für alle W-Maße π auf $(I,\mathfrak{F})$ mit endlichem Träger.

(6) *Es sei π ein W-Maß auf $(I,\mathfrak{F})$ mit endlichem Träger* $\operatorname{Tr}\pi$. *Dann existiert zu jedem $B\in\mathfrak{B}$ eine $\mathfrak{A}$-meßbare reelle Funktion φ_B auf X mit der Eigenschaft:*

$$Q_i^{\mathfrak{A}}(B)=\varphi_B \qquad Q_i-f.s. \text{ für alle } i\in\operatorname{Tr}\pi.$$

Beweis. (1)$\Leftrightarrow$(2) folgt als Spezialfall von 5.2.1.

(2)$\Leftrightarrow$(3) wurde bereits in 5.1.4 (Fall (c)) abgeleitet.

(4)$\Leftrightarrow$(5) ergibt sich nach Modifikation der Schlüsse in [11], 1440. Das einschlägige Argument entnimmt man der Theorie der Markoff-Prozesse, indem man $\mathfrak{F}'$ als Vergangenheit, $\mathfrak{A}'$ als Gegenwart und $\mathfrak{B}'$ als Zukunft eines Prozesses interpretiert. Analog sieht man (5)$\Leftrightarrow$(6) ein. Die Äquivalenz (4)$\Leftrightarrow$(1) erschließt man schließlich mit Hilfe von 4.2.1 (Beweis).

5.4.2. *Bemerkung.* Es seien $\mathfrak{Y}=(Y,\mathscr{L}_{\mathfrak{B}},(Q_i)_{i\in I})$ ein beliebiges topologisches Experiment und $\mathfrak{A}_0$ sowie $\mathfrak{A}$ σ-Unteralgebren von $\mathfrak{B}$ mit $\mathfrak{A}_0\subset\mathfrak{A}$. Dann ergibt sich, falls $\mathfrak{Y}$ μ-dominiert ist (durch ein σ-endliches Maß μ auf $(Y, \mathfrak{B})$), aus der Theorie der erschöpfenden σ-Algebren ([1], Theorem 6.4):

$$\mathfrak{Y}^{\mathfrak{A}_0} \supset \mathfrak{Y} \Rightarrow \mathfrak{Y}^{\mathfrak{A}} \supset \mathfrak{Y}.$$

Hierbei kann die μ-Dominanz der Familie $(Q_i)_{i \in I}$ nicht ersatzlos fortgelassen werden. Für nichtdominierte Experimente ist die Aussage falsch, wie man einem Gegenbeispiel von Burkholder ([4], 1192) entnimmt: Man betrachte nämlich das Experiment $\mathfrak{Y} = (\mathbb{R}, \mathscr{L}_{\mathscr{B}}, (Q_i)_{i \in I})$ mit symmetrischen $Q_i (i \in I)$ sowie die σ-Unteralgebren $\mathfrak{A}_0$ und $\mathfrak{A}$ von $\mathscr{B}$ definiert durch

$$\mathfrak{A}_0 := \{B \in \mathscr{B}: \ B = -B\}$$

und

$$\mathfrak{A} \ := \{B \cup A_0: \ B \in \mathscr{B}, A_0 \in \mathfrak{A}_0 \text{ mit } B \subset S\},$$

wobei S eine symmetrische, die Null enthaltende nicht-Borelsche Teilmenge von $\mathbb{R}$ ist. Dann ist zwar $\mathfrak{Y}^{\mathfrak{A}_0} \supset \mathfrak{Y}$, nicht aber $\mathfrak{Y}^{\mathfrak{A}} \supset \mathfrak{Y}$.

5.4.3. Als Letztes in diesem Paragraphen werde noch der Zusammenhang zwischen der Erschöpftheitsrelation für Kontraktionsexperimente und der f-Divergenz hergestellt, welche wir bereits in 5.1.4 eingeführt haben.

Wir behalten die Bezeichnung von 5.1.4 bei.

Dann gilt der

Satz. *Es sei* $\mathfrak{Y} = (Y, \mathscr{L}_{\mathfrak{B}}, (Q_i)_{i \in I})$ *ein durch ein* σ*-endliches Maß* μ *auf* $(Y, \mathfrak{B})$ *dominiertes topologisches Experiment,* $\mathfrak{A}$ *eine* σ*-Unteralgebra von* $\mathfrak{B}$ *und* $\mathfrak{Y}^{\mathfrak{A}}$ *die* $\mathfrak{A}$*-Kontraktion von* $\mathfrak{Y}$*. Es sind nachstehende Aussagen* (*und natürlich auch die Aussagen* (2) *bis* (6) *von* 5.4.1) *äquivalent:*

(1) $\mathfrak{Y}^{\mathfrak{A}} \supset \mathfrak{Y}$.

(2) $\mathscr{I}_f(Q_{i_1}, Q_{i_2}) = \mathscr{I}_f(\mathrm{Res}_{\mathfrak{A}}\, Q_{i_1}, \mathrm{Res}_{\mathfrak{A}}\, Q_{i_2})$ *für alle Paare* $(i_1, i_2) \in I \times I$.

Beweis. Nach der Implikation (1)$\Leftrightarrow$(3) von 5.4.1 genügt es zu zeigen, daß obige Aussage (2) mit der Erschöpftheit der σ-Algebra $\mathfrak{A}$ bezüglich der Familie $(Q_i)_{i \in I}$ gleichwertig ist. Dies trifft aber zu nach 5.1.4 (Fall (c)), da nach Halmos-Savage ([10], Theorem 3) paarweise Erschöpftheit und Erschöpftheit bezüglich dominierter Familien von W-Maßen äquivalent sind.

6. Informativität und Invarianz

In diesem Paragraphen werden wir den Charakterisierungssatz 4.2.1 im Spezialfall strenger ε-Informativität (4.2.3) verschärfen derart, daß der zur Charakterisierung der Relation $\mathfrak{X} \sqsupset_{\varepsilon} \mathfrak{Y}$ herangezogene R-Kern zusätzlich invariant gewählt werden kann. Invarianz versteht sich in diesem Zusammenhang als Invarianz bezüglich gewisser Familien von linearen Abbildungen zwischen den relevanten Banachverbänden von $\mathfrak{X}$ und $\mathfrak{Y}$.

6.1. Zulässige Familien von linearen Abbildungen

Der Begriff der Zulässigkeit bezieht sich auf ein gegebenes Paar $(\mathfrak{X}, \mathfrak{Y})$ von Experimenten mit gleicher Indexmenge I, wobei $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ μ-dominiert ist durch ein σ-endliches Maß μ auf $(X, \mathfrak{A})$ und $\mathfrak{Y} = (Y, F, (Q_i)_{i \in I})$ vorerst als beliebig vorausgesetzt wird.

Es seien also A eine beliebige Menge sowie für jedes $\alpha \in A$ S_α bzw. V_α positive (beschränkte) lineare Abbildungen von $L^1_{\mathfrak{A}}$ in $L^1_{\mathfrak{A}}$ bzw. Kerne in $\mathscr{R}(F, F)$ mit den folgenden Eigenschaften:

(1) Für jedes $\alpha \in A$ bildet S_α bzw. V'_α die Familie $(P_i)_{i \in I}$ bzw. $(Q_i)_{i \in I}$ in sich ab.

(2) Für jedes $\alpha \in A$ und Paare $(i, j) \in I \times I$ gilt:

$$S_\alpha P_i = P_j \Rightarrow V'_\alpha Q_j = Q_i.$$

Wir betrachten nun für jedes $\varepsilon \geqq 0$ die Teilmenge

$$\mathscr{N}_\varepsilon := \{U \in \mathscr{R}(F, L_{\mathfrak{A}}): \ \|U' P_i - Q_i\| \leqq \varepsilon \text{ für alle } i \in I\}$$

von $\mathscr{R}(F, L_{\mathfrak{A}})$ sowie die Abbildungen $U \to S'_\alpha U V_\alpha$ von $\mathscr{N}_\varepsilon$ in $\mathscr{R}(F, L_{\mathfrak{A}})$ $(\alpha \in A)$. Diese ist linear und bildet $\mathscr{N}_\varepsilon$ in sich ab, wie folgende Abschätzung zeigt:

$$\begin{aligned}\|(S'_\alpha U V_\alpha)' P_i - Q_i\| &= \|(V'_\alpha U' S_\alpha) P_i - Q_i\| \\ &= \|V'_\alpha U' P_j - Q_i\| \leqq \|V'_\alpha U' P_j - V'_\alpha Q_j\| + \|V'_\alpha Q_j - Q_i\| \\ &= \|V'_\alpha (U' P_j - Q_j)\| \leqq \varepsilon \qquad \text{für alle } i \in I.\end{aligned}$$

Definition. Eine Familie $(\Gamma_\alpha)_{\alpha \in A}$ von Abbildungen Γ_α von $\mathscr{N}_0$ in sich, definiert durch $\Gamma_\alpha U := S'_\alpha U V_\alpha$ für alle $U \in \mathscr{N}_0$ $(\alpha \in A)$, wobei die Familien $(S_\alpha)_{\alpha \in A}$ und $(V_\alpha)_{\alpha \in A}$ obige Bedingungen (1) und (2) erfüllen, heißt *zulässig* (für das Paar $(\mathfrak{X}, \mathfrak{Y})$), falls $(\Gamma_\alpha)_{\alpha \in A}$ eine der nachstehenden Eigenschaften besitzt:

(i) $(\Gamma_\alpha)_{\alpha \in A}$ ist abelsch (d.h. die Γ_α sind paarweise vertauschbar).

(ii) $(\Gamma_\alpha)_{\alpha \in A}$ ist auflösbare (oder auch kompakte) Gruppe.

(iii) $(\Gamma_\alpha)_{\alpha \in A}$ ist Halbgruppe mit rechts- bzw. linksinvariantem Mittel.

Zur Eigenschaft (iii) vergleiche man etwa [8], 586.

6.2. Existenz invarianter R-Kerne

Ein Kern $T \in \mathscr{N}_0$ wird wie üblich *Γ_α-invariant* oder kurz *invariant* genannt, falls $\Gamma_\alpha T = T$ gilt für alle $\alpha \in A$.

Der folgende Satz liefert eine Verschärfung von 4.2.3:

6.2.1. Satz. *Es seien $\mathfrak{X}$ und $\mathfrak{Y}$ Experimente wie zu Beginn des Paragraphen, $\mathfrak{X}(\mu)$ das μ-assoziierte Experiment von $\mathfrak{X}$ sowie $(\Gamma_\alpha)_{\alpha \in A}$ eine (für das Paar $(\mathfrak{X}, \mathfrak{Y})$) zulässige Familie. Dann sind die folgenden Aussagen äquivalent:*

(1) *$\mathfrak{X}(\mu) \sqsupset_\varepsilon \mathfrak{Y}$ für alle $\varepsilon > 0$.*

(2) *Es existiert ein $T \in \mathscr{R}(F, L_{\mathfrak{A}})$ mit $T' P_i = Q_i$ für alle $i \in I$.*

(3) *Es existiert ein invarianter Kern $T \in \mathscr{R}(F, L_{\mathfrak{A}})$ mit $T' P_i = Q_i$ für alle $i \in I$.*

Beweis. Die Äquivalenz (1)$\Leftrightarrow$(2) wurde bereits in 4.2.3 gezeigt. (3)$\Rightarrow$(2) ist evident. Zu zeigen bleibt (2)$\Rightarrow$(3): Nach 4.2.2 ist die Menge $\mathscr{N}_0$ konvex und $\mathscr{T}$-kompakt. Wegen (2) gilt $\mathscr{N}_0 \neq \emptyset$. Jede der Abbildungen Γ_α ist affin-linear und $\mathscr{T}$-stetig $(\alpha \in A)$: Die $\mathscr{T}$-Stetigkeit der Abbildung $T \to \langle S'_\alpha T V_\alpha g, f\rangle$ von $\mathscr{N}_0$ in $\mathbb{R}$ ergibt sich aus der Gleichung $\langle S'_\alpha T V_\alpha g, f\rangle = \langle T V_\alpha g, S_\alpha f\rangle$ für alle $g \in F$ und $f \in L^1_{\mathfrak{A}}$ $(\alpha \in A)$. Nach dem Fixpunktsatz von Markoff-Kakutani bzw. seinen Verschärfungen ([8], Theorem 1), angewendet auf die Menge $\mathscr{N}_0$ zusammen mit der Familie $(\Gamma_\alpha)_{\alpha \in A}$, existiert mindestens ein $T_0 \in \mathscr{N}_0$ mit $\Gamma_\alpha T_0 = T_0$ für alle $\alpha \in A$.

6.2.2. Korollar. *Besitzt $\mathfrak{X}$ zudem die Eigenschaft, daß die Familie $(P_i)_{i\in I}$ normtotal ist in der Menge $L'_{\mathfrak{A}}$, so sind folgende Aussagen äquivalent:*

(1) $\mathfrak{X}(\mu) \sqsupset_\varepsilon \mathfrak{Y}$ *für alle* $\varepsilon > 0$.

(2) *Es existiert genau ein invarianter Kern* $T \in \mathscr{R}(F, L_{\mathfrak{A}})$ *mit* $T' P_i = Q_i$ *für alle* $i \in I$.

Beweis. Angenommen es gäbe zwei invariante Kerne T_1 und T_2 in $\mathscr{R}(F, L_{\mathfrak{A}})$ mit $T_1' P_i = Q_i = T_2' P_i$ für alle $i \in I$. Da $(P_i)_{i\in I}$ (norm-)total ist in $L'_{\mathfrak{A}}$, gilt somit $T_1' = T_2'$ auf ganz $L'_{\mathfrak{A}}$. Die Abbildung $T \to T'$ ist aber ein isometrischer Isomorphismus von $\mathscr{R}(F, L_{\mathfrak{A}})$ in $\mathscr{R}(L'_{\mathfrak{A}}, F')$, also gilt $T_1 = T_2$.

Als Anwendung betrachten wir schließlich:

6.3. Translationsinvariante Experimente

Es seien $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ und $\mathfrak{Y} = (X, \mathscr{L}_{\mathfrak{A}}, (Q_i)_{i\in I})$ zwei topologische Experimente, wobei speziell $X = I := G$ eine lokalkompakte abelsche Gruppe mit abzählbarer Basis (etwa $G := \mathbb{R}^n$ für ein $n \geqq 1$) bezeichnet. G werde additiv geschrieben, das neutrale Element sei 0. Dann induziert die Rechtstranslation R_α definiert durch $R_\alpha(x) := x + \alpha$ für alle $x \in X$ (und festes $\alpha \in G$) analoge Abbildungen $f \to f_\alpha$ bzw. $\mu \to \mu_\alpha$, die Rechtstranslationen beliebiger reeller Funktionen bzw. von W-Maßen auf X. ω bezeichne das bis auf eine multiplikative Konstante eindeutig bestimmte Haarsche Maß auf G. Wir setzen weiterhin voraus, daß zwei W-Maße P und Q auf $(X, \mathfrak{A})$ gegeben seien mit $P \ll \omega$, so daß $P_i = R_{-i} P$ und $Q_i = R_{-i} Q$ für alle $i \in I$ gilt. Man definiere für jedes $\alpha \in A := G$ die in 6.1 erklärten Abbildungen S_α und V_α durch $S_\alpha := R_\alpha$ und $V_\alpha := R_\alpha$. Es ist leicht nachzurechnen, daß die Familie $(\Gamma_\alpha)_{\alpha\in A}$ mit $\Gamma_\alpha T = R_{-\alpha} T R_\alpha$ für alle $T \in \mathscr{N}_0$ $(\alpha \in A)$ zulässig ist für das Paar $(\mathfrak{X}, \mathfrak{Y})$. Denn zunächst gilt

$$R_\alpha P_i = P_{i-\alpha} \Rightarrow R_{-\alpha} Q_{i-\alpha} = Q_i$$

für alle $\alpha \in A$, $i \in I$. Ferner ist $(\Gamma_\alpha)_{\alpha\in A}$ isomorph zur Gruppe G, daher also insbesondere abelsch.

Sei nun $\mathfrak{X} \sqsupset \mathfrak{Y}$. Wegen 5.2.1 existiert ein Markoff-Kern T auf $(X, \mathfrak{A})$, dessen Fortsetzung zu einer Abbildung $T \in \mathscr{R}(\mathscr{L}_{\mathfrak{A}}, \mathscr{L}_{\mathfrak{A}})$ die Gleichung $T' P_i = Q_i$ für alle $i \in I$ und $\Gamma_\alpha T = T$ für alle $\alpha \in A$ erfüllt. Offenbar sind die folgenden Aussagen äquivalent:

(a) $\Gamma_\alpha T = T (\Leftrightarrow R_\alpha T R_{-\alpha} = T)$.

(b) $T(x+\alpha, B+\alpha) = T(x, B)$ für alle $\alpha \in A$, $B \in \mathfrak{A}$.

Denn zunächst gilt für Indikatorfunktionen 1_B auf X $(B \in \mathfrak{A})$:

$$\begin{aligned} R_\alpha T_x R_{-\alpha} 1_B &= T_{x+\alpha} R_{-\alpha} 1_B = T_{x+\alpha} (1_B)_{-\alpha} \\ &= \int (1_B)_{-\alpha}(y)\, T(x+\alpha, dy) = \int 1_B(y-\alpha)\, T(x+\alpha, dy) \\ &= \int 1_{B+\alpha}(y)\, T(x+\alpha, dy) = T_{x+\alpha} 1_{B+\alpha} \\ &= T(x+\alpha, B+\alpha) \end{aligned}$$

für alle $x \in X$, $\alpha \in A$. Ist nun $f \in \mathscr{L}_{\mathfrak{A}}$, so genügt es, f in Positiv- und Negativteil zu zerlegen und beide Summanden durch Elementarfunktionen zu approximieren.

Ausgehend von einem Γ_α-invarianten (d.h. translations-invarianten) Kern T auf $(X, \mathfrak{A})$ definiert $T(0, \cdot)$ ein W-Maß λ auf $(X, \mathfrak{A})$ mit $T(x, B) = T(0, B-x) = \lambda(B-x)$ für alle $x \in X$ und alle $B \in \mathfrak{A}$. Umgekehrt sei ein W-Maß λ auf $\mathfrak{A}$ vorgegeben, so daß die Abbildung $x \to \lambda(B-x)$ für alle $B \in \mathfrak{A}$ meßbar ist. Dann wird durch $T(x, B) := \lambda(B-x)$ für alle $x \in X$, $B \in \mathfrak{A}$ ein translationsinvarianter Markoff-Kern auf $(X, \mathfrak{A})$ definiert. Die Meßbarkeit der Abbildung $x \to \lambda(B-x)$ für alle $B \in \mathfrak{A}$ folgt z.B. aus der Voraussetzung $\lambda \ll \omega$.

In der Tat ist T ein *Faltungskern*, d.h. $Tf = f * \lambda^{\sim}$ für alle $f \in \mathscr{L}_{\mathfrak{A}}$, wobei λ ein W-Maß auf $(X, \mathfrak{A})$ und $\lambda^{\sim}$ die Spiegelung von λ am neutralen Element 0 der Gruppe G ist $\left(\lambda^{\sim}(B) = \lambda(-B)\right.$ für alle $\left.B \in \mathfrak{A}\right)$.

Es sei nämlich $f \in \mathscr{L}_{\mathfrak{A}}$. Zunächst gilt für alle $x \in X$, $B \in \mathfrak{A}$:

$$\begin{aligned} T 1_B(x) &= T(x, B) = \lambda(B-x) = \int 1_{B-x}(y)\, \lambda(dy) \\ &= \int 1_B(x+y)\, \lambda(dy) = \int 1_B(x-y)\, \lambda^{\sim}(dy). \end{aligned}$$

Approximation durch Elementarfunktionen und Zerlegung von f in Positiv- und Negativteil liefern:

$$Tf(x) = \int f(x-y)\, \lambda^{\sim}(dy), \qquad \text{also } Tf = f * \lambda^{\sim}.$$

Im Falle unseres Beispiels hatte der Markoff-Kern T auf $(X, \mathfrak{A})$ die Eigenschaft $TP_i = Q_i$ für alle $i \in I$, d.h. aber:

$$\begin{aligned} &\int T_x(f)\, P(dx) = Q(f) \qquad \text{für alle } f \in \mathscr{L}_{\mathfrak{A}} \\ &\quad \Leftrightarrow \iint f(x+y)\, \lambda(dy)\, P(dx) = Q(f) \qquad \text{für alle } f \in \mathscr{L}_{\mathfrak{A}} \\ &\quad \Leftrightarrow \lambda * P = Q. \end{aligned}$$

Das gesuchte W-Maß λ kann als Lösung dieser Faltungsgleichung gewonnen werden. Seien nämlich $\hat{P}$, $\hat{Q}$ und $\hat{\lambda}$ die verallgemeinerten Fouriertransformierten der Maße P, Q und λ. (Als Referenz für die relevanten Fakten aus der harmonischen Analyse lokalkompakter abelscher Gruppen konsultiere man etwa [16], insbesondere 1.3.3 und 1.7.3.) Es gilt für jeden Charakter χ von X die Gleichung:

$$\hat{\lambda}(\chi)\, \hat{P}(\chi) = \hat{Q}(\chi).$$

Bezeichne $\hat{\omega}$ das Haarsche Maß auf der Gruppe $\hat{X}$ aller stetigen Charaktere von X. Dann findet man, da die Menge $B := \{\chi \in \hat{X} : \hat{P}(\chi) = 0\}$ das $\hat{\omega}$-Maß 0 besitzt, für die Fouriertransformierte des Maßes λ:

$$\hat{\lambda}(\chi) = \begin{cases} \dfrac{\hat{Q}(\chi)}{\hat{P}(\chi)}, & \text{falls } \chi \in \complement B, \\[2ex] \lim\limits_{\substack{\psi \to \chi \\ \psi \in \complement B}} \dfrac{\hat{Q}(\chi)}{\hat{P}(\chi)}, & \text{falls } \chi \in B. \end{cases}$$

Wegen der Eineindeutigkeit der Zuordnung $\lambda \to \hat{\lambda}$ ist dadurch λ eindeutig festgelegt.

7. Anwendungen und Beispiele

7.1. Placierungsexperimente

Derartige Experimente, eingeführt von De Groot ([9], Definition 2.3), werden bei solchen statistischen Fragestellungen untersucht, bei denen es darum geht, eine feste Anzahl n von Beobachtungen optimal zu placieren in dem Sinne, daß man eine natürliche Zahl k sowie natürliche Zahlen $n_1, \dots, n_k$ wählt mit $\sum_{j=1}^{k} n_j = n$ und die zugehörigen Experimente studiert. Genauer wird jedem topologischen Experiment $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ auf folgende Weise ein weiteres Experiment $\tilde{\mathfrak{X}}$ zugeordnet: Es sei ξ die gemäß 2.1 auf $(X_0, \mathfrak{A}_0, P_0)$ definierte Zufallsvariable. Es werde angenommen, daß die Werte der zufälligen Stichprobe $(\xi_1, \dots, \xi_k)$ aus der nach P_i verteilten Grundgesamtheit nicht beobachtbar seien, statt dessen es aber möglich sei, für jedes $j = 1, \dots, k$ die Werte von n_j Zufallsvariablen $\xi_{j1}, \dots, \xi_{jn_j}$ zu messen, welche unter der Hypothese $[\xi_j = x]$ $(x \in X)$ unabhängig seien mit den bedingten Verteilungen $P_{\xi_{jl} | \xi_j = x} =: Q(\cdot, x)$ für alle $l = 1, \dots, n_j$. Aus der Unabhängigkeit der $\xi_1, \dots, \xi_k$ folgt die Unabhängigkeit der $\xi^{(1)}, \dots, \xi^{(k)}$ definiert durch

$$\xi^{(m)} := (\xi_{m1}, \dots, \xi_{mn_m}) \qquad \text{für } m = 1, \dots, k.$$

Zu jedem $i \in I$ berechnet man nun die gemeinsame Verteilung H_i der ξ_{jl} $(j = 1, \dots, k;$ $l = 1, \dots, n_j)$. Das *Placierungsexperiment* $\tilde{\mathfrak{X}}$ *von* $\mathfrak{X}$ *vom Grade* $n = \sum_{j=1}^{k} n_j$ ist dann als das Tripel $(X^n, \mathscr{L}_{\mathfrak{A}^n}, (H_i)_{i \in I})$ definiert und wird oft mit $\mathfrak{X}(n; k; n_1, \dots, n_k)$ bezeichnet.

Falls $\mathfrak{X}$ dominiert ist, so sind für je zwei Placierungsexperimente $\tilde{\mathfrak{X}}_1$ und $\tilde{\mathfrak{X}}_2$ von $\mathfrak{X}$ wegen 5.2.1 $\tilde{\mathfrak{X}}_1 \supset \tilde{\mathfrak{X}}_2$ und $\tilde{\mathfrak{X}}_1 \succ \tilde{\mathfrak{X}}_2$ äquivalent. Man kann somit eine Reihe von Resultaten aus [9] ohne Schwierigkeit auf die durch folgende Definition gegebene Situation übertragen: Innerhalb der Klasse $\mathbf{P} := \mathbf{P}(\mathfrak{X})$ aller Placierungsexperimente von $\mathfrak{X}$ vom Grade n werde $\tilde{\mathfrak{X}}_1$ *maximal informativ* genannt, wenn $\tilde{\mathfrak{X}}_1 \supset \tilde{\mathfrak{X}}_2$ gilt für alle $\tilde{\mathfrak{X}}_2 \in \mathbf{P}$.

Beispielsweise ergibt sich für das Experiment $\mathfrak{X} = (\mathbb{R}, \mathscr{L}_{\mathfrak{B}}, (P_i)_{i \in I})$ mit $P_i = \nu_{i, \sigma^2}$ für $i \in \mathbb{R}$ und $\sigma^2 > 0$ (dem Maß der Normalverteilung mit Mittelwert i und Varianz σ^2) falls $Q(\cdot, x) := \nu_{x, \tau^2}$ für $x \in \mathbb{R}$ und $\tau^2 > 0$, daß $\mathfrak{X}(n; n; 1, \dots, 1)$ maximal informativ ist in $\mathbf{P}$. (Siehe [9], 20.)

Der folgende *Einschub* wird später benötigt werden: Seien $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ und $\mathfrak{Y} = (Y, \mathscr{L}_{\mathfrak{B}}, (Q_i)_{i \in I})$ zwei diskrete Experimente mit $P_i := \sum_{k \geq 1} p_k^{(i)} \varepsilon_{x_k}$ und $Q_i := \sum_{k \geq 1} q_k^{(i)} \varepsilon_{y_k}$ für alle $i \in I$, wobei $(x_k)_{k \geq 1}$ bzw. $(y_k)_{k \geq 1}$ Folgen in X bzw. Y sind. Hierbei bezeichnen offenbar für jedes $i \in I$ $p_k^{(i)} := p^{(i)}(x_k)$ und $q_k^{(i)} := q^{(i)}(y_k)$ positive reelle Zahlen mit $\sum_{k \geq 1} p_k^{(i)} = 1$ und $\sum_{k \geq 1} q_k^{(i)} = 1$. Dann ist die Familie offenbar μ-dominiert (beispielsweise durch das σ-endliche Maß $\mu := \sum_{k \geq 1} \varepsilon_{x_k}$). Es sei nun $(P_i)_{i \in I}$ vollständig. Dann gilt $\mathfrak{X} \succ \mathfrak{Y}$ genau dann, wenn die folgende Bedingung erfüllt ist:

Zu jedem $y_k \in Y$ $(k \geq 1)$ existiert eine Folge $(a_j)_{j \geq 1}$ nichtnegativer reeller Zahlen $a_j := a_j(y_k)$ mit der Eigenschaft:

$$q^{(i)}(y_k) = \sum_{j \geq 1} a_j \, p^{(i)}(x_j). \qquad (*)$$

Dies sieht man sofort mittels 5.2.2 ein, wenn man die Vollständigkeit der Familie $(P_i)_{i\in I}$ so aufschreibt: Wenn immer $\sum_{j\geqq 1} a_j\, p^{(i)}(x_j)=0$ für alle $i\in I$, so folgt $a_j=0$ für $j\geqq 1$. Damit ist also eine Formel gefunden, deren Gültigkeit aufgrund von 5.2.1 eine Charakterisierung von $\mathfrak{X}\supset\mathfrak{Y}$ liefert.

Ein Experiment, das die an $\mathfrak{X}$ gestellten Voraussetzungen erfüllt, ist z.B. $\mathfrak{X}=(X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ mit $|I|=n$, $X=\{x_1,\dots,x_k\}$, $k\leq n$, $\mathscr{L}_{\mathfrak{A}}=\mathscr{L}_X$ und $P_i:=\sum_{j=1}^{k} p_j^{(i)}\,\varepsilon_{x_j}$ für jedes $i\in I$, wobei die Determinante der Matrix $(p_j^{(i)})_{i,j=1,\dots,k}$ von Null verschieden ist.

Für das Experiment $\mathfrak{X}=(X, \mathscr{L}_X, (P_i)_{i\in[0,1]})$ mit $X=\{x_1, x_2\}$, wobei $x_1, x_2\in]0,1[$, und $P_i:=i\,\varepsilon_{x_1}+(1-i)\,\varepsilon_{x_2}$ $(i\in[0,1])$ ist das Placierungsexperiment $\mathfrak{X}$ $(n;n;1,\dots,1)$ nicht maximal informativ innerhalb $\mathbf{P}$, da beispielsweise die Relation $\mathfrak{X}(n;n;1,\dots,1)\succ\mathfrak{X}(n;1;n)$ falsch ist. Dies liegt daran, daß die Bedingung $(*)$ für die beiden Placierungsexperimente nicht erfüllt ist.

7.2. *Endliche klassische Experimente*

Sie wurden bereits von Blackwell in seinen für die Theorie bahnbrechenden Arbeiten [2] und [3] behandelt. Es seien also $\mathfrak{X}=(X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ und $\mathfrak{Y}=(Y, \mathscr{L}_{\mathfrak{B}}, (Q_i)_{i\in I})$ Experimente mit $|I|=:n$. Jedem dieser Experimente kann eindeutig das *Standardexperiment* zugeordnet werden. Dies wird für $\mathfrak{X}$ etwa auf die folgende Weise gewonnen: Wegen $P_i\ll P':=\sum_{i\in I} P_i$ existieren reelle $\mathfrak{A}$-meßbare Dichten $f_i\geqq 0$ auf X mit $P_i:=f_i\,P'$. O.B.d.A. sei $\sum_{i\in I} f_i=1$ gesetzt. Also nimmt $f:=(f_i)_{i\in I}$ nur Werte im Simplex

$$K:=\{y\in\mathbb{R}^I\colon\; y_i\geqq 0\,(i\in I),\; \sum_{i\in I} y_i=1\}$$

an. Seien $\mathfrak{M}$ die Borel-σ-Algebra über K und $\pi_i:=f(P_i)$ für jedes $i\in I$. Dann heißt das Tripel $(K, \mathscr{L}_{\mathfrak{M}}, (\pi_i)_{i\in I})$ das zu $\mathfrak{X}$ gehörige Standardexperiment $\mathfrak{X}^*$. Setzt man noch $\pi':=f(P')$ und normiert man π', so erhält man das zu $\mathfrak{X}$ gehörige *Standardmaß* π. Es errechnet sich der Schwerpunkt x_π von π zu $(1/n)$. Umgekehrt bestimmt jedes W-Maß λ auf K mit $x_\lambda=(1/n)$ eindeutig ein Standardexperiment. Es gilt $\mathfrak{X}^*\sim\mathfrak{X}$.

Seien nun μ und ν die Standarmaße der Experimente $\mathfrak{X}$ und $\mathfrak{Y}$. Mit $\mathscr{M}_+:=\mathscr{M}_+(K)$ werde die Menge aller positiven Radonmaße auf K bezeichnet. Dann gilt der

Satz. *Folgende Aussagen sind äquivalent:*

(1) $\mathfrak{X}\supset\mathfrak{Y}$.

(2) $\mathfrak{X}^*\supset\mathfrak{Y}^*$.

(3) $\mathfrak{X}^*\succ\mathfrak{Y}^*$.

(4) $\mu\succ\nu$ *im Sinne von Bishop und de Leeuw, nämlich:* $\mu(g)\geqq\nu(g)$ *für alle konvexen, stetigen reellen Funktionen* g *auf* K.

(5) *Es existiert eine Dilatation* T *auf* K *mit* $T\nu=\mu$.

(6) *Zu jeder Zerlegung* $\nu=\sum_{j=1}^{m}\nu_j$ *mit* $\nu_j\in\mathcal{M}_+$ *existiert eine Zerlegung* $\mu=\sum_{j=1}^{m}\mu_j$ *mit* $\mu_j\in\mathcal{M}_+$, *so daß* μ_j *und* ν_j *gleiche Resultanten besitzen für jedes* $j=1,\ldots,m$.

(7) *Es existiert ein W-Maß* σ *auf* $K\times K$ *mit Marginalmaßen* μ,ν, *so daß für alle stetigen Funktionen g und alle affin-linearen, stetigen Funktionen l auf K gilt:*

$$\int_{K\times K} g(x)\,l(y)\,\sigma(dx,dy)=\int_K g(x)\,l(x)\,\nu(dx).$$

Beweis. (1)$\Leftrightarrow$(2) ist klar wegen $\mathfrak{X}^*\sim\mathfrak{X}$.

(2)$\Leftrightarrow$(3) ergibt sich als Spezialfall von 5.2.1.

(3)$\Leftrightarrow$(4) steht bereits in [2], Theorem 4. (Man vergleiche auch [3], Theorem 8.)

Die Äquivalenzen (4)$\Leftrightarrow$(5)$\Leftrightarrow$(6)$\Leftrightarrow$(7) entnimmt man der Arbeit [5]: In [5] wird (4)$\Leftrightarrow$(6)$\Leftrightarrow$(7) sogar für beliebige konvexe, kompakte Mengen K in lokalkonvexen Hausdorff-Räumen bewiesen. Ist K zusätzlich metrisierbar, so folgt in dieser Allgemeinheit auch (4)$\Leftrightarrow$(5).

Bemerkung. Es sei die Situation zu Beginn dieses Abschnitts gegeben. Wir setzen für jedes $i\in I$:

$$A_i:=[f_i=1,\ f_j=0 \text{ für } j\in I, j\neq i],$$

dann ist $\mathfrak{X}$ maximal informativ genau dann, wenn gilt:

$$P'\Big(X\setminus\bigcup_{i\in I}A_i\Big)=0.$$

In der Tat genügt es wegen $\mathfrak{X}\sim\mathfrak{X}^*$, das Standardexperiment $\mathfrak{X}^*$ zu betrachten. Aufgrund des Satzes ((2)$\Leftrightarrow$(4)) ist aber $\mathfrak{X}^*$ maximal informativ genau dann, wenn das zugehörige Standardmaß π maximal ist im Sinne von Bishop und de Leeuw. π ist dasjenige eindeutig bestimmte W-Maß auf K, welches von den Eckpunkten des Simplex K getragen wird, belegt also jeden Eckpunkt von K mit der Masse $1/n$. (Man vergleiche hierzu [14], 231.)

Als Beispiel eines maximal informativen Experiments nennen wir $\mathfrak{X}=(X,\mathscr{L}_{\mathfrak{A}},(P_i)_{i\in I})$ mit $P_i:=\varepsilon_{x_i}$ für alle $i\in I$, wobei $x_1,\ldots,x_n$ beliebige Elemente von X sind.

Die Spezialisierung der Voraussetzungen dieses Abschnitts auf den Fall $|I|=2$ liefert *Anwendungen auf die Testtheorie:* Es seien durch die Experimente $\mathfrak{X}=(X,\mathscr{L}_{\mathfrak{A}},(P_i)_{i=1,2})$ und $\mathfrak{Y}=(X,\mathscr{L}_{\mathfrak{A}},(Q_i)_{i=1,2})$, welche als μ-dominiert vorausgesetzt werden (durch ein σ-endliches Maß μ auf $(X,\mathfrak{A})$), zwei einfache Tests definiert: der eine für die Hypothese P_1 gegen die Alternative P_2 und der zweite für Q_1 gegen Q_2. Wir schreiben diese Testprobleme kurz $[P_1:P_2]$ und $[Q_1:Q_2]$. Mit β bezeichnen wir die zu $[P_1:P_2]$ gehörige Abbildung von $[0,1]$ in sich, welche definiert ist durch

$$\beta(\alpha):=\sup\{E_{P_2}(t):\quad t: X\to[0,1], t\in\mathscr{L}_{\mathfrak{A}}, E_{P_1}(t)\leqq\alpha\}.$$

Analog wird β' für $[Q_1:Q_2]$ erklärt.

Wir erhalten aufgrund des Satzes und der Definition 2.3.1: $[P_1:P_2]\supset[Q_1:Q_2]$ genau dann, wenn $\beta(\alpha)\geqq\beta'(\alpha)$ für alle $\alpha\in[0,1]$.

Ferner folgt aus $[P_1:P_2]\supset[Q_1:Q_2]$ sofort $[P_2:P_1]\supset[Q_2:Q_1]$.

Setzt man insbesondere $Q_2 := P_1$ und ist die Familie $\{P_1, P_2, Q_1\}$ vom exponentiellen Typ, so ergibt sich aus einem bekannten Satz über monotone Dichtequotienten ([19], 221) die Relation $[P_1 : P_2] \supset [P_1 : Q_1]$.

Wir geben nun eine noch speziellere Anwendung und zwar auf die 2×2-*Kontingenztafel:*

Eine Menge von Individuen werde daraufhin untersucht, ob gewisse Merkmale A und B an ihnen vorhanden sind oder nicht. Man interessiert sich für einen Test für die Unabhängigkeit von A und B. Die relativen Häufigkeiten von A und B werden mit $p := P(A)$ und $\pi := P(B)$ $(p, \pi > 0)$ bezeichnet. Wir betrachten die Merkmalkombinationen $AB, A\bar{B}, \bar{A}B$ und $\bar{A}\bar{B}$ unter der Hypothese

$$H_0 : A \text{ und } B \text{ sind unabhängig } \left(\text{d.h. } P(AB) = P(A)P(B)\right)$$

gegen die Alternative

$$H_1 : \; P(AB) = \rho \quad \text{mit} \quad \rho \neq p\pi (0 \leqq \rho \leqq 1).$$

Man kann nun der Gesamtheit eine Stichprobe von festem Umfang entnehmen und dabei auf vier verschiedene Weisen vorgehen. Zum Beispiel kann man die Stichprobe der Menge derjenigen Individuen entnehmen, die A besitzen, und durch Inspektion des einzelnen Individuums feststellen, ob B vorliegt oder nicht. Man hat es hierbei also mit einer binomialverteilten Stichprobe zu tun, für die

$$P_{H_0}(B|A) = \pi \quad \text{und} \quad P_{H_1}(B|A) = \frac{\rho}{p}$$

gilt.

Das hierdurch beschriebene diskrete Experiment werde mit $\mathfrak{X}_1$ bezeichnet. $\mathfrak{X}_2$ werde so gewonnen: Man geht von denjenigen Individuen aus, welche B besitzen, und erhält:

$$P_{H_0}(A|B) = p \quad \text{sowie} \quad P_{H_1}(A|B) = \frac{\rho}{\pi}.$$

Da $\mathfrak{X}_1$ und $\mathfrak{X}_2$ den Voraussetzungen des Einschubs in 7.1 genügen, gilt $\mathfrak{X}_1 \succ \mathfrak{X}_2$, also $\mathfrak{X}_1 \supset \mathfrak{X}_2$ (5.2.1). Die Bezeichnung ist dabei so gewählt, daß $p \leqq \pi \leqq \frac{1}{2}$ gilt. Man kann ferner zeigen ([12], 76/77), daß das unter den vier möglichen Experimenten maximal informative dadurch gewonnen wird, daß man das innerhalb der Grundgesamtheit seltenste Merkmal auswählt.

7.3. Shannon- und Fisher-Informativität für topologische Experimente mit Apriori-Verteilung

7.3.1. Es sei $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$ ein topologisches Experiment, welches μ-dominiert ist durch ein σ-endliches Maß μ auf $(X, \mathfrak{A})$. Ferner seien $\mathfrak{J}$ eine σ-Algebra von Teilmengen in I und π ein W-Maß auf $(I, \mathfrak{J})$. Ein derartiges Maß π heißt auch *Apriori-Verteilung* für $\mathfrak{X}$.

Wegen $P_i \ll \mu$ gilt $P_i = f_i\,\mu$ für alle $i \in I$, wobei die f_i die Dichten der P_i bezüglich μ sind. Wir definieren nun die reelle Funktion g auf X durch

$$g(x) := \int_I f_i(x)\,\pi(di) \qquad \text{für alle } x \in X$$

und schließlich den *mittleren Informationsgehalt* (die *Shannonsche Information*) von $\mathfrak{X}$ (bei Apriori-Verteilung π auf $(I, \mathfrak{F})$) als die reelle Zahl

$$\begin{aligned}\mathscr{I}_{\mathfrak{X}} := \mathscr{I}_{\mathfrak{X}, \pi} = E_{P_i} E_\pi \log \frac{f_i}{g} &= E_{P_i} \int_I \log \frac{f_i(x)}{g(x)} \pi(di) \\ &= \int_X \int_I \log \frac{f_i(x)}{g(x)} \pi(di) P_i(dx) \\ &= \int_I \int_X f_i(x) \log \frac{f_i(x)}{g(x)} \mu(dx) \pi(di),\end{aligned}$$

sofern die aufgeschriebenen Integrale einen Sinn haben. Mittels des *Informationsoperators* H, definiert für Dichtefunktionen f auf X durch

$$H(f) := \int_X f(y) \log f(y) \mu(dy),$$

erhalten wir die Formel

$$\mathscr{I}_{\mathfrak{X}, \pi} = E_\pi H(f_i) - H E_\pi(f_i).$$

In [13] zeigt Lindley die folgenden *Eigenschaften* von $\mathscr{I}_{\mathfrak{X}}$:

(1) $\mathscr{I}_{\mathfrak{X}} \geqq 0$.

(2) $\mathscr{I}_{\mathfrak{X}} = 0$ genau dann, wenn f_i unabhängig ist von i außerhalb einer π-Nullmenge $(i \in I)$.

(3) Für das Experiment $\mathfrak{Z} := \mathfrak{X} \otimes \mathfrak{Y}$ gilt $\mathscr{I}_{\mathfrak{X}} \leqq \mathscr{I}_{\mathfrak{Z}}$.

(4) Für $\mathfrak{X} \succ \mathfrak{Z}$ ergibt sich $\mathscr{I}_{\mathfrak{X}} = \mathscr{I}_{\mathfrak{Z}}$.

Definition. Zwei μ-dominierte topologische Experimente $\mathfrak{X}$ und $\mathfrak{Y}$ mit gleicher Indexmenge I stehen in der Relation $\mathfrak{X} \overset{S}{\supset} \mathfrak{Y}$, d.h. $\mathfrak{X}$ ist *Shannon-informativer (S-informativer)* als $\mathfrak{Y}$, wenn gilt:

(i) $\mathscr{I}_{\mathfrak{X}, \pi} \geqq \mathscr{I}_{\mathfrak{Y}, \pi}$ für alle W-Maße π auf $(I, \mathfrak{F})$.

(ii) Es existiert mindestens ein W-Maß π_0 auf $(I, \mathfrak{F})$ mit $\mathscr{I}_{\mathfrak{X}, \pi_0} = \mathscr{I}_{\mathfrak{Y}, \pi_0}$.

Gilt Gleichheit in (i) für *alle* W-Maße π auf $(I, \mathfrak{F})$, so heißt $\mathfrak{X}$ *S-äquivalent* zu $\mathfrak{Y}$ $(\mathfrak{X} \approx \mathfrak{Y})$.

Sinngemäß verwenden wir die Bezeichnung $\mathfrak{X} \overset{S}{\gtrsim} \mathfrak{Y}$.

Aus 5.2.1, 4.4 und obigen Eigenschaften (3) und (4) folgt der

Satz. *Unter den bisherigen Voraussetzungen gilt:*

$$\mathfrak{X} \supset \mathfrak{Y} \Rightarrow \mathfrak{X} \overset{S}{\gtrsim} \mathfrak{Y}.$$

Bemerkung. Die Umkehrung des Satzes ist falsch, wie aus folgendem Beispiel resultiert ([2], 101 und [13], 997): Man betrachte das Experiment $\mathfrak{X} = (X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i \in I})$, wobei $X := \{0, 1\}$, $\mathscr{L}_{\mathfrak{A}} = \mathscr{L}_X$, $I := \{\frac{1}{4}, \frac{3}{4}\}$ und für jedes $i \in I$ $P_i := i\varepsilon_0 + (1-i)\varepsilon_1$ gesetzt wird.

Als Apriori-Verteilungen π auf $(I, \mathfrak{F})$ wählen wir die Maße $\alpha\varepsilon_{\frac{1}{4}} + (1-\alpha)\varepsilon_{\frac{3}{4}}$ $(\alpha \in [0, 1])$. Dann läßt sich ein Experiment $\mathfrak{Y} = (X, \mathscr{L}_X, (Q_i)_{i \in I})$ mit $Q_i = \alpha_i \varepsilon_0 + (1-\alpha_i)\varepsilon_1$ und $\alpha_i \in [0, 1]$ für alle $i \in I$ angeben, so daß zwar $\mathfrak{X} \overset{S}{\supset} \mathfrak{Y}$, nicht aber $\mathfrak{X} \supset \mathfrak{Y}$ gilt. In der Tat sind in einem derartigen Fall $\mathfrak{X}$ und $\mathfrak{Y}$ bezüglich der Relation $\supset$ sogar unvergleichbar.

7.3.2. Bei Restriktion auf eine spezielle Klasse von Experimenten kann schließlich noch der von Fisher eingeführte Informationsbegriff mit den bisher erörterten in Beziehung gebracht werden. Hierzu gehen wir zunächst von einem topologischen Experiment $\mathfrak{X}=(X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ aus mit $X=I:=\mathbb{R}$, welches μ-dominiert ist durch das Lebesgue-Maß μ auf $(X, \mathfrak{A})$. Für jedes $i\in I$ gilt also $P_i=f_i\,\mu$, wobei f_i die μ-Dichte von P_i ist. Die durch die Festsetzung

$$\mathscr{I}^{\mathfrak{X}}(i):=\int_X f_i(x)\left[\frac{\partial}{\partial i}\log f_i(x)\right]^2 \mu(dx)$$

definierte Funktion $\mathscr{I}^{\mathfrak{X}}$ auf I heißt, sofern das auftretende Integral existiert, die *Fisher-Information (F-Information) von* $\mathfrak{X}$.

Definition. Zwei μ-dominierte Experimente $\mathfrak{X}$ und $\mathfrak{Y}$ der obigen Art (mit gleicher Indexmenge $I:=\mathbb{R}$) stehen in der Relation $\mathfrak{X}\overset{F}{\supseteq}\mathfrak{Y}$, d.h. $\mathfrak{X}$ ist *Fisher-informativer (F-informativer)* als $\mathfrak{Y}$, falls gilt:

$$\mathscr{I}^{\mathfrak{X}}(i)\geqq\mathscr{I}^{\mathfrak{Y}}(i) \qquad \text{für alle } i\in I.$$

In [20] gibt Stone hinreichende Bedingungen an für die Gültigkeit der Implikation:

$$\mathfrak{X}\overset{S}{\supseteq}\mathfrak{Y}\Rightarrow\mathfrak{X}\overset{F}{\supseteq}\mathfrak{Y}.$$

Es ergibt sich also in diesem Fall wegen des Satzes in 7.3.1 sogar:

$$\mathfrak{X}\supset\mathfrak{Y}\Rightarrow\mathfrak{X}\overset{F}{\supseteq}\mathfrak{Y}.$$

Ohne Zusatzbedingung muß diese Implikation falsch sein, da man leicht Experimente $\mathfrak{X}$ und $\mathfrak{Y}$ konstruieren kann, für die zwar $\mathfrak{X}\supset\mathfrak{Y}$ erfüllt ist, die Funktionen $\mathscr{I}^{\mathfrak{X}}$ und $\mathscr{I}^{\mathfrak{Y}}$ jedoch nicht einmal definiert zu sein brauchen. Im Fall von *Lokalisationsexperimenten* $\mathfrak{X}(\alpha)=(X, \mathscr{L}_{\mathfrak{A}}, (P_i)_{i\in I})$ mit $X=I:=\mathbb{R}$, $\mathscr{L}_{\mathfrak{A}}:=\mathscr{L}_{\mathscr{B}}$ und $P_i:=f_i\,\mu$, wobei μ ein σ-endliches Maß auf $(X, \mathfrak{A})$ und $f_i(x):=f(\alpha(x-i))$ gilt für alle $x\in X$ ($i\in I$ und festes $\alpha\in\mathbb{R}, \alpha>0$), errechnet sich die Funktion $\mathscr{I}^{\mathfrak{X}(\alpha)}$ unter der Voraussetzung ihrer Existenz zu:

$$\begin{aligned}\mathscr{I}^{\mathfrak{X}(\alpha)}(i)&=\int f(\alpha(x-i))\left[\frac{\partial}{\partial i}\log f(\alpha(x-i))\right]^2 \mu(dx)\\&=\alpha\int f(u)\left[\frac{d}{du}\log f(u)\right]^2 \mu(du).\end{aligned}$$

Es folgt somit aus der Isotonie der Abbildung $\alpha\to\mathscr{I}^{\mathfrak{X}(\alpha)}$ sofort $\alpha_1>\alpha_2(>0)\Rightarrow\mathfrak{X}(\alpha_1)\overset{F}{\supseteq}\mathfrak{X}(\alpha_2)$.

8. Anhang. Binäre Relationen in Klassen von Experimenten

Zusammenfassend sollen nun einige Resultate der Arbeit in die Sprache der Theorie der geordneten Mengen gekleidet werden. Mit der Definition der ε-Informativität (2.3.1) gewinnt man im Spezialfall konstanter Fehlerfunktion $\varepsilon\equiv 0$ eine Quasiordnung innerhalb der Gesamtheit **E** aller Experimente mit gleicher (fester) Indexmenge. Das heißt, für je drei Elemente $\mathfrak{X}$, $\mathfrak{Y}$, $\mathfrak{Z} \in \mathbf{E}$ gelten die folgenden Axiome:

(1) $\mathfrak{X} \supset \mathfrak{X}$ (Reflexivität).

(2) $\mathfrak{X} \supset \mathfrak{Y}$ und $\mathfrak{Y} \supset \mathfrak{Z} \Rightarrow \mathfrak{X} \supset \mathfrak{Z}$ (Transitivität).

Wir nennen diese Quasiordnung die *statistische* (oder kurz *S-)Quasiordnung* in **E**. In natürlicher Weise gibt die S-Quasiordnung in **E** Anlaß zu einer Äquivalenzrelation in **E**. Die Menge der Äquivalenzklassen kann sodann geordnet werden: Für je zwei Äquivalenzklassen $[\mathfrak{X}]$ und $[\mathfrak{Y}]$ wird $[\mathfrak{X}] \supset [\mathfrak{Y}]$ definiert durch die Forderung, daß für gewisses $\mathfrak{X} \in [\mathfrak{X}]$ und gewisses $\mathfrak{Y} \in [\mathfrak{Y}]$ gilt: $\mathfrak{X} \supset \mathfrak{Y}$. Für die Klassen $[\mathfrak{X}]$ und $[\mathfrak{Y}]$ ist das folgende Axiom erfüllt:

(3) $[\mathfrak{X}] \supset [\mathfrak{Y}]$ und $[\mathfrak{Y}] \supset [\mathfrak{X}] \Rightarrow [\mathfrak{X}] = [\mathfrak{Y}]$ (Antisymmetrie).

Offenbar ist die S-Quasiordnung nicht total, d.h. es gilt nicht nachstehendes Axiom:

(4) Wenn immer $\mathfrak{X}, \mathfrak{Y} \in \mathbf{E}$ ist, so gilt entweder $\mathfrak{X} \supset \mathfrak{Y}$ oder $\mathfrak{Y} \supset \mathfrak{X}$.

Dies folgt z.B. aus der Unvergleichbarkeit gewisser Elemente aus **E** (Bemerkung in 7.3.1).

Die in 5.1.2 eingeführte binäre Relation innerhalb der Klasse **K** aller klassischen Experimente ist für die Teilklasse **T** der topologischen Experimente, welche zudem dominiert sind, wegen 5.2.1 ebenfalls eine Quasiordnung. Wir nennen sie die *Blackwell-*(oder kurz *B-)Quasiordnung* in **T**.

Zwei binäre Relationen r bzw. s in Mengen A bzw. B, die einander eineindeutig zugeordnet sind, nennen wir *äquivalent*, wenn für je zwei Elemente $a_1, a_2 \in A$ und zugeordnete Elemente $b_1, b_2 \in B$ gilt: $a_1\, r\, a_2 \Leftrightarrow b_1\, s\, b_2$.

Mit **F** werde die Klasse aller endlichen klassischen Experimente mit gleicher (fester) Indexmenge bezeichnet, die ihr eineindeutig zugeordnete Klasse der Standardexperimente von Elementen in **F** mit **F***. Dann sind die S-Quasiordnungen von **F** und **F*** in obigem Sinne äquivalent. Dies folgt aus dem Satz in 7.2.

Nach [14], 227 weiß man ferner, daß die im Satz von 7.2 erklärte Relation von Bishop und de Leeuw eine Ordnungsrelation in der Menge $\mathcal{M}_+ := \mathcal{M}_+(K)$ aller positiven Radonmaße auf (dem Standardsimplex) K (zur Indexmenge I) ist. Diese werde die *Bishop-de Leeuw-*(oder kurz *BD-)Ordnung* in $\mathcal{M}_+$ genannt. Über Existenz und Eigenschaften maximaler Elemente im Sinne der BD-Ordnung erfährt man in [14], 227/228. (Man vergleiche auch unsere Bemerkung in 7.2.)

Des weiteren wurde von Loomis speziell in $\mathcal{M}_+$ die in Aussage (6) des Satzes von 7.2 enthaltene Definition einer binären Relation gegeben. Diese erweist sich wegen Äquivalenz mit der BD-Ordnung selbst als eine Ordnungsrelation in $\mathcal{M}_+$, welche wir *Loomis-*(oder kurz *L-)Ordnung* nennen wollen. In der Tat sind B-Quasiordnung von **F*** sowie BD- und L-Ordnung in der Menge der Standardmaße von Elementen in **F*** äquivalent.

Die Diskussion maximal informativer Experimente wird in 7.1 und 7.2 (Bemerkung) angesprochen: In der Klasse **P** der Placierungsexperimente von $\mathfrak{X} \in \mathbf{E}$ vom Grade n können maximal informative Elemente konstruiert werden. Maximal informative Experimente in **F** sind eineindeutig maximalen Maßen im Sinne der BD-Ordnung zugeordnet.

Schließlich sind die in 7.3 eingeführten auf Shannon und Fisher zurückgehenden Relationen offenbar Quasiordnungen in denjenigen Teilklassen von **E**, die durch die Definitionen von Shannon- bzw. Fisher-Information festgelegt sind. In der Tat sind dies Teilklassen von **K**. Äquivalenz dieser Quasiordnungen etwa mit der B-Quasiordnung liegt nicht vor (Bemerkung in 7.3.1).

Endlich bildet die Klasse **T** bezüglich der S-Quasiordnung und Produktbildung im Sinne von 2.1.4 wegen 5.3.1 ein quasigeordnetes Gruppoid.

Literatur

1. Bahadur, R. R.: Sufficiency and statistical decision functions. Ann. math. Statistics **25**, 423–462 (1954).
2. Blackwell, D.: Comparison of experiments. Proc. Second Berkeley Sympos. math. Statist. Probab., 93–102 (1951).
3. – Equivalent comparisons of experiments. Ann. math. Statistics **24**, 265–272 (1953).
4. Burkholder, D. L.: Sufficiency in the undominated case. Ann. math. Statistics **32**, 1191–1200 (1961).
5. Cartier, P., J. M. G. Fell, and P. A. Meyer: Comparaison des mesures portées par un ensemble convexe compact. Bull. Soc. math. France **29**, 435–445 (1964).
6. Csiszár, I.: Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizität von Markoffschen Ketten. Publ. math. Inst. Hungar. Acad. Sci., Ser. A, **8**, 85–108 (1963).
7. – Information-type measures of difference of probability distributions and indirect observations. Stud. Sc. Math. Hung. **2**, 299–318 (1967).
8. Day, M. M.: Fixed point theorems for compact convex sets. Illinois J. Math. **5**, 585–590 (1961) sowie **8**, 713 (1964). [Correction.]
9. Groot, M. H. De: Optimal allocation of observations. Ann. Inst. statist. Math. **18**, 13–28 (1966).
10. Halmos, P. R., and L. J. Savage: Application of the Radon-Nikodym theorem to the theory of sufficient statistics. Ann. math. Statistics **20**, 225–241 (1949).
11. LeCam, L.: Sufficiency and approximate sufficiency. Ann. math. Statistics **35**, 1419–1455 (1964).
12. Lehmann, E. L.: Testing Statistical Hypotheses. New York and London: John Wiley & Sons 1959.
13. Lindley, D. V.: On a measure of information provided by an experiment. Ann. math. Statistics **27**, 986–1005 (1956).
14. Meyer, P. A.: Probability and Potentials. Waltham, Mass.-Toronto-London: Blaisdell Publishing Company 1966.
15. Phelps, R. R.: Extreme positive operators and homomorphisms. Trans. Amer. math. Soc. **108**, 265–274 (1963).
16. Rudin, W.: Fourier Analysis on Groups. New York and London: Interscience 1962.
17. Sacksteder, R.: A note on statistical equivalence. Ann. math. Statistics **38**, 787–794 (1967).
18. Schaefer, H. H.: Topological Vector Spaces. New York and London: MacMillan Publishing Company 1966.
19. Schmetterer, L.: Einführung in die mathematische Statistik. Wien u. New York: Springer 1966.
20. Stone, M.: Non-equivalent comparisons of experiments and their use for experiments involving location parameters. Ann. math. Statistics **32**, 326–332 (1961).

Dr. H. Heyer
Mathematisches Institut
der Universität Erlangen-Nürnberg
852 Erlangen, Bismarckstraße $1^1/_2$

(Eingegangen am 26. Februar 1968)