JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 159, 550–594 (1991)

On the Scoring Approach to Admissibility of Uncertainty Measures in Expert Systems

IRWIN R. GOODMAN

Naval Ocean Systems Center, San Diego, California 92152

HUNG T. NGUYEN AND GERALD S. ROGERS

Department of Mathematical Sciences, New Mexico State University, Las Cruces, New Mexico 88003

Submitted by L. Zadeh

Received March 19, 1990

This paper arose from our need to rigorize, clarify, and address fully the results of Lindley's paper (Scoring rules and the inevitability of probability, *Internat. Statist. Rev.* 50, (1982), 1–26). Herein, we develop a calculus of admissibility in a game theoretic setting. In the case of an additive aggregation function, it is shown that decomposable measures, such as those used in fuzzy logics, are admissible. Also, the problem of the admissibility of the Dempster–Shafer belief functions is investigated via the concept of random sets. It is shown that the class of admissible measures in a scoring framework depends on the assumptions concerning the aggregation function in use. In particular, for nonadditive aggregation functions, an admissible measure may not be transformable to a finitely additive probability measure.

1. INTRODUCTION

With the advent of Artificial Intelligence and the development of expert systems, a number of schools of thought has arisen concerning how uncertainties in complex real-world situations are to be modeled. In addition to probability in all of its variants [34], approaches to uncertainty modeling now include the Dempster–Shafer theory of belief functions [26], Zadeh fuzzy logic [32], and nonmonotonic logics [20]; a survey of these approaches is in [28]. Roughly speaking, the choice of a set-function to model the uncertainty involved in a problem at hand is related to the pragmatic aspects or, at a deeper level, to the semantic nature of the type of uncertainty under consideration.

Lindley [16, 17] proposed a simple but novel approach, extending DeFinetti's original considerations on coherence of uncertainties to a more

0022-247X/91 $3.00

general setting, for judging the usefulness of different competing uncertainty measures. DeFinetti's original work [6] may be viewed as a two-person zero-sum game, played between player I, "nature" or "the master of ceremonies," and player II, "decision-maker" or "you" or "bookie" as denoted variously in the literature [8, 13, 22]. DeFinetti's uncertainty game has great appeal: it is determined by the cumulative amount of the bets. The concept of admissibility in DeFinetti's game is in fact a type of uniform local admissibility (see also [14]) which is commonly expressed as the coherence axiom. DeFinetti's chief result is that the only coherent uncertainty measures are finitely additive *conditional* probability measures.

Lindley's main contribution to the situation was to investigate DeFinetti's game by replacing the squared loss functions by a more general score function. For the most part, DeFinetti and Lindley assumed addition for the overall loss function (or aggregation function).

Lindley's chief results are as follows:

(i) If an uncertainty measure μ is admissible with respect to a score function f, then μ can be transformed into a finitely additive conditional probability measure via a known transform depending on f, say P_f;

(ii) Within the class of score functions f such that P_f is increasing, the necessary condition in (i) is also sufficient.

Roughly speaking, an admissible uncertainty measure has to be a function of a probability measure, i.e., one cannot avoid probability! However, note that such an admissible uncertainty measure need not be a probability measure! (See also the axiomatic work of Cox [5].)

(iii) As implications from (i), the Dempster–Shafer belief function, Zadeh's possibility measure, confidence values and significant statements are all inadmissible!

The purpose of our work is threefold:

(a) To analyze Lindley's results and implications, for which we recast Lindley's somewhat informal arguments and concepts totally within a game theoretic setting.

It is pointed out in this paper that DeFinetti in his earlier [6] more restricted work and later, Lindley [16] in his generalization of DeFinetti's efforts, both tacitly assumed:

(I) "measure-free" conditional events exist independent of any particular choice of probability measure, but are compatible with the usual evaluation of conditional probability.

(II) The usual (unconditional) event indicator function can be extended to be well defined upon conditional events.

(III) A natural conjunctive chaining relation holds between conditional events.

As a consequence of the above assumptions, in carrying out the analysis here, basically two cases are considered apropos to choosing an uncertainty measure in the DeFinetti–Lindley uncertainty game: (1) all finite sequences of conditional events, where each such sequence possesses a common antecedent—this includes as a special case all finite sequences of (unconditional) events, by identifying unconditional events as conditional ones having a universal antecedent—and (2) all finite sequences of conditional events with possibly differing antecedents—this case obviously includes the first as a special case.

The structure of uncertainty games is rigorously spelled out in Section 2. In Section 3, a calculus of admissibility, from an analytic viewpoint, is developed for games with arbitrary aggregation functions. In Section 4, uncertainty games with an additive aggregation function are considered in detail together with a Bayesian analysis.

(b) To show, contrary to Lindley's conclusions outlined in (iii) above, that there are rather large classes of nonadditive uncertainty measures, such as belief functions and decomposable measures in fuzzy logics, which are admissible. Also, Zadeh's max-possibility measures are shown to be uniform limits of admissible measures (Sections 5 and 6).

(c) To study the effects of the assumption of additive score functions, we present, in Section 7, various examples of non-additive aggregation functions. These illustrate the fact that the class of admissible measures in a scoring framework depends heavily on the nature of the aggregation functions. In particular, there exist aggregation functions such that admissible measures cannot be transformed into finitely additive probability measures (as opposed to the case of the additive aggregation function in Lindley's work).

In summary, by formalizing Lindley's work within a general and rigorous game theory framework, we develop a calculus of admissibility which can be used to compare competing uncertainty measures in Artificial Intelligence. We shed light on controversial conclusions concerning the inadmissibility of some well-known uncertainty measures and on the position of the "inevitability of probability."

2. Structure of Uncertainty Games

In this section, we will formalize Lindley's scoring approach in a game theory setting. Since the scoring approach involves concepts such as "conditional events," uncertainty measures, score functions, aggregation functions (implicitly), and admissibility, we need to define these terms rigorously.

2.1. *Conditional Events*

Let Ω be a set and $\mathscr{A}$ a Boolean (or σ^-) algebra of subsets of Ω. Elements of $\mathscr{A}$ are called events. The set complement of A in Ω is denoted by A^c; the intersection of A, B in Ω is AB; and their union is $A \cup B$.

For $A \in \mathscr{A}$, the indicator function has values

$$I_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \in A^c \end{cases}$$

In certain forms, it is simpler to identify A with I_A so that $A = 1$ if A "occurs" and $A = 0$ if A "does not occur."

For $A, B \in \mathscr{A}$, the "measure-free" conditional event $A \mid B$ is defined by DeFinetti [6] (see also [22]) as the restriction of I_A to B, i.e.,

$$(A \mid B)(\omega) = \begin{cases} 1 & \text{if } \ \omega \in AB \\ 0 & \text{if } \ \omega \in A^c B \\ (\text{undefined}) & \text{if } \ \omega \in B^c. \end{cases}$$

Except when $B = \Omega$ and $(A \mid \Omega)$ is identified with A, these conditional events are not elements of $\mathscr{A}$.

The above definition implies the invariant form

$$(A \mid B) = (AB \mid B).$$

Assume, also the fundamental homomorphic-like forms compatible with any fixed antecedent conditional probability

$$(A \mid B)^c = (A^c \mid B),$$

$$(A \cup C) \mid B) = (A \mid B) \cup (C \mid B),$$

$$(AC \mid B) = (A \mid B)(C \mid B).$$

Although DeFinetti recognized the potential use of measure-free conditional events, in obtaining his key results a formal calculus of relations was not developed. (Again, see [6, especially Vol. 1, Chap. 4, Vol. 2, pp. 266 et passim to 333].) However, DeFinetti and Lindley implicitly recognized the natural conjunctive chaining relation among conditional events mentioned earlier. (Specifically, see the remark at the end of Section 2.3 and Theorems

3.23 and 4.21′.) For a general treatment of conditional events, see [24] or [12].

Goodman and Nguyen have derived a full calculus of operators and relations extending the unconditional counterparts for boolean algebras of events to the conditional case. In addition, a wide variety of desirable mathematical properties of these entities have been proven to hold based upon a minimal set of elementary assumptions, including the tacitly assumed conjunctive chaining relation mentioned above.

In the context of uncertainty modeling, the uncertainty of an event A is, in general, assigned on the basis of additional information, another event B. But, a priori, conditional uncertainty measure μ need not be a probability so that an algebra of measure-free conditional events has to be investigated as a domain for μ.

Now let $\tilde{\mathscr{A}}$ be the class of all conditional events, i.e.,

$$\tilde{\mathscr{A}} = \{(A \mid B): A, B \in \mathscr{A}\}.$$

By the identification of $(A \mid \Omega)$ with A, we obtain $\mathscr{A} \subseteq \tilde{\mathscr{A}}$.

For any set X, and $n \geqslant 1$, X^n denotes the product space $X \times X \times \cdots \times X$ (n times). We will use the notation X_∞ to denote the space of all finite n-tuples $\{\hat{x}_n : \hat{x}_n = (x_1, ..., x_n), x_i \in X\}$. Unless otherwise indicated, "$x_i \in X$" and the like will always mean $x_1, ..., x_n$ for a generic n. For $\hat{A}_n = (A_1, ..., A_n) \in \mathscr{A}^n$, we identify $\hat{A}_n$ with its indicator function $\hat{A}_n(\omega)$ defined as $(A_1(\omega), ..., A_n(\omega)) \in \{0, 1\}^n$.

For example, $\hat{A}_6(\omega) = (1, 1, 0, 0, 0, 0)$ indicates that A_1, A_2 occurred but A_3, A_4, A_5, A_6 did not occur. Similarly, we identify

$$\hat{A}_n = ((E_1 \mid F_1), ..., (E_n \mid F_n)) \in \tilde{\mathscr{A}}^n$$

with the function $\hat{A}_n : \Omega \to \{0, 1, u\}^n$ having values

$$\hat{A}_n(\omega) = ((E_1 \mid F_1)(\omega), ..., (E_n \mid F_n)(\omega)) \in \{0, 1, u\}^n.$$

2.2. *Uncertainty Games*

We proceed now to formalize a special class of games called *uncertainty games*. Roughly speaking, these are triples $(\Lambda_1, \Lambda_2, L)$ in which Λ_1 is a set of configurations (or realizations) of finite collections of (conditional) events, Λ_2 is a collection of "set"-functions representing "uncertainty measures," and L is a real-valued loss (or penalty) function. Specifically, $\Lambda_1 = \tilde{\mathscr{A}}_\infty \times \Omega$. This set Λ_1 is regarded as the space of all possible "moves" or "pure strategies" of player I.

Next, fix, once for all, four real numbers, $a_2 < a_0 < a_1 < a_3$; let $\Lambda_2 = \{\mu: \mu: \tilde{\mathscr{A}} \to [a_2, a_3]\} = [a_2, a_3]^{\tilde{\mathscr{A}}}$. Each element of Λ_2 is a map, which assigns a number, describing its uncertainty, to each (conditional) event. A_2 is regarded as the space of "moves" of player II.

Consider now the choice of loss function. This is carried out in two stages as the composition (aggregation) of other (score) functions. As in Lindley's paper, we call any function f: $[a_2, a_3] \times \{0, 1, u\} \to \mathbb{R}$ (the real numbers), a *score function* if the following are satisfied:

(i) For each $j \in \{0, 1\}$, $f(\cdot, j)$: $[a_2, a_3] \to \mathbb{R}$ satisfies the following "regularity" conditions: $f(\cdot, j)$ is continuously differentiable, with a unique global minimum in $[a_2, a_3]$ at a_j, (decreasing over $[a_2, a_j]$ and increasing over $[a_j, a_3]$);

(ii) $f(x, u) = 0$, $\forall x \in [a_2, a_3]$.

For example, we may interpret the score function as follows: player I has selected $A \in \mathscr{A}$ and $\omega \in \Omega$ and player II has selected $\mu \in A_2$, then the score for player II is $f(\mu(A), 1)$ if A happens to occur, $f(\mu(A), 0)$ if A does not occur.

Now we need to extend f as a map from the space

$$\{(\hat{x}_n, \hat{t}_n): n \geqslant 1,\ \hat{x}_n \in [a_2, a_3]^n,\ \hat{t}_n \in \{0, 1, u\}^n\}$$

to the space $\mathbb{R}_\infty$: for each $n \geqslant 1$, $\hat{x}_n = (x_1, ..., x_n)$, $\hat{t}_n = (t_1, ..., t_n)$, $f(\hat{x}_n, \hat{t}_n) = (f(x_1, t_1), ..., f(x_n, t_n))$.

Similarly, each *uncertainty map* (or *measure*) μ: $\tilde{\mathscr{A}} \to [a_2, a_3]$ is extended to a map from $\tilde{\mathscr{A}}_\infty$ to $[a_2, a_3]_\infty$ as follows: for each $n \geqslant 1$,

$$\hat{A}_n = (A_1, ..., A_n) \in \tilde{\mathscr{A}}^n,\ \mu(\hat{A}_n) = (\mu(A_1), ..., \mu(A_n)) \in [a_2, a_3]^n.$$

An obvious way of combining individual scores $f(\mu(A_i), A_i(\omega))$ $(i = 1, 2, ..., n)$ to obtain the total score is using addition on $\mathbb{R}$, i.e., take $L_{f,+}(\hat{A}_n, \omega, \mu) = \sum_{i=1}^n f(\mu(A_i), A_i(\omega))$. Thus the loss function $L_{f,+}$ depends on two functions: f (score) and $+$ (aggregation). This special case will be referred to as the *additive aggregation* case. In general, by an *aggregation function*, we mean a mapping ψ: $\mathbb{R}_\infty \to \mathbb{R}$ such that

(a) ψ is continuously differentiable in all of its arguments;

(b) ψ is increasing in each of its arguments.

(c) $\psi(0_n) = 0$, $\forall n \geqslant 1$, where 0_n is the zero vector in $\mathbb{R}^n$.

Note that the additive aggregation function is generated by ordinary addition on $\mathbb{R}$: $\psi = +$ is equivalent to the sequence of functions $(g_n, n \geqslant 1)$, where g_n: $\mathbb{R}^n \to \mathbb{R}$, $g(x_1, ..., x_n) = \sum_{i=1}^n x_i$. Similarly, we identify an aggregation function ψ with the sequence $(\psi_n, n \geqslant 1)$, where ψ_n is the restriction of ψ to $\mathbb{R}^n$. Note also that, while the additive aggregation function is symmetric, there is no a priori reason to impose such a condition on arbitrary aggregation functions.

Now, given a score function f and an aggregation function ψ, we define the loss function $L_{f,\psi}$ as follows:

$$L_{f,\psi}: \Lambda_1 \times \Lambda_2 \to \mathbb{R}$$
$$L_{f,\psi}(\hat{A}, \omega, \mu) = \psi(f(\mu(\hat{A}), \hat{A}(\omega))).$$

(Note that f and μ are used in the extended sense.)

The triple $(\Lambda_1, \Lambda_2, L_{f,\psi})$ is called an *uncertainty game* and is denoted by $G_{f,\psi}$.

In DeFinetti's game [6], $a_0 = a_2 = 0$, $a_1 = a_3 = 1$, and

$$f(x, j) = \begin{cases} (x-j)^2 & \text{if } j = 0, 1 \\ 0 & \text{if } j = u(\text{undefined}); \end{cases}$$

here $\psi = +$. In Lindley's extension of DeFinetti's game, a_j, $j = 0, 1, 2, 3$, are not restricted to $[0, 1]$, $\psi = +$, and f is an arbitrary score function.

2.3. *Subgames*

To formalize various concepts of admissibility in an uncertainty game $G_{f,\psi}$, we first introduce the concept of subgame.

Suppose we are interested only in some given finite collection of conditional events, say $\hat{A}$, then we need to look only at the *subgame*

$$G_{f,\psi,\hat{A}} = (\Lambda_{1,\hat{A}}, \Lambda_{2,\hat{A}}, L_{f,\psi,\hat{A}}),$$

where

$$\Lambda_{1,\hat{A}} = \Omega, \; \Lambda_{2,\hat{A}} = [a_2, a_3]^{\hat{A}}, \; L_{f,\psi,\hat{A}}: \Omega \times [a_2, a_3]^{\hat{A}} \to \mathbb{R}$$

and

$$L_{f,\psi,\hat{A}}(\omega, \mu) = \psi(f(\mu(\hat{A}), \hat{A}(\omega))).$$

For example, we can view $A_n \in \tilde{\mathscr{A}}^n$ as a set $\hat{A}_n = \{A_1, ..., A_n\} \subseteq \mathscr{A}$. Then $[a_2, a_3]^{\hat{A}_n} = \{\mu: \{A_1, ..., A_n\} \to [a_2, a_3]\}$ so that each μ in $[a_2, a_3]^{\hat{A}_n}$ is the restriction of a μ in $[a_2, a_3]^{\tilde{\mathscr{A}}}$. In a subgame, the finite collection of conditional events in $\hat{A}$ is specified; if player II chooses μ to express an uncertainty about these conditional events, then the overall loss would be $L_{f,\psi,\hat{A}}(\omega, \mu)$. The subgame $G_{f,\psi,\hat{A}}$ is regarded as a game with partial information, namely player II does know that $\hat{A}$ is to be considered.

For example, if $\hat{A} = ((E|F), (E^c|F))$ and $\psi = +$, $\hat{A}(\omega) = ((E|F)(\omega), (E^c|F)(\omega))$, and the set of configurations of $\hat{A}$ giving rise to non-zero losses $(f(x, u) \equiv 0)$ is

$$\{((E|F), 1), ((E^c|F), 0), ((E|F), 0), ((E^c|F), 1)\}.$$

For $\hat{t} \in \{(1,0),(0,1)\}$, $\hat{x}=(x_1,x_2)$, where $x_1=\mu(E|F)$, $x_2=\mu(E^c|F)$, we have two overall scores: $f(x_1,1)+f(X_2,0)$ (if $E|F$ occurs) and $f(x_1,0)+f(x_2,1)$ (if $E^c|F$ occurs).

It should be noted that so far only those finite sequences of events have been considered which have a common antecedent. Relevant to this, denote for any $F \in \mathscr{A}$,

$$\mathscr{A}_F=\{(E|F):E\in\mathscr{A}\}.$$

The reason for this is that if one wished to determine the possible indicator evaluation combinations among any sequence such as $((E_1|F_1),(E_2|F_2),(E_3|F_3))$, until recently, no standard technique existed for dealing with this issue. However, Goodman, Nguyen, and Walker [12] and Schay [24] have proposed independent nonstandard syntactic approaches for treating this and related problems. Only one such situation will be considered in this paper, namely the fundamental natural conjunctive chaining relation,

$$(E|FG)(F|G)=(EF|G),$$

which is obviously true for all corresponding conditional probability evaluations

$$p(E|FG)\cdot p(F|G)=p(EF|G)$$

for $p(G)>0$. More details of this will be seen in Theorem 3.2.3 et passim.

2.4. *Equivalent Reduced Forms of Games and Subgames*

The space $A_1=\tilde{\mathscr{A}}_\infty\times\Omega$ in the game $G_{f,\psi}$ is infinite, in general, but for each $\hat{A}\in\tilde{\mathscr{A}}_\infty$, the space of configurations of $\hat{A}$, namely $\hat{A}(\Omega)$ is finite, a subset of $\{0,1,u\}^{|\hat{A}|}$, where $|\hat{A}|$ denotes the "dimension" of $\hat{A}$; e.g., if $\hat{A}\in\tilde{\mathscr{A}}^n$, then $|\hat{A}|=n$. On the other hand, the domain of $L_{f,\psi}$ involves Ω, but since $L_{f,\psi}(\hat{A},\cdot,\mu)$ is constant on each $(\hat{A})^{-1}(t)$, $t\in\hat{A}(\Omega)$, Ω can be replaced by the finite $\mathscr{A}$-partition (of Ω) generated by $\hat{A}$. Specifically, for each $\hat{A}\in\tilde{\mathscr{A}}_\infty$, say $\hat{A}=((E_1|F_1),...,(E_n|F_n))$, consider the finite collection of events $\mathscr{C}(\hat{A})=\{E_iF_i, E_i^cF_i, F_i^c, i=1,...,n\}$. Let $\pi(\hat{A})$ denote the *canonical partition* of Ω generated by $\mathscr{C}(\hat{A})$ (which reduces to $\hat{A}=\{E_i\}$ when all $F_i=\Omega$. Then, $\pi(\hat{A})=\{B_j, j=1,2,...,2^{3n}\}$, where each B_j is of the form $\prod_{k=1}^{3n}D_k^{\varepsilon_k}$, $D_k\in\mathscr{C}(\hat{A})$, $\varepsilon_k=1$ or c; $D_k^1=D_k$, D_k^c is the set complement of D_k. Also, each D_k is a union of the B_j's. (See [23, p. 12–15].) Note that because the E's and F's are not necessarily distinct, the cardinality of $\pi(\hat{A})$, say $m=|\pi(\hat{A})|$, is most often $<2^{3n}$, but at least 2.

The cardinality of $\pi(\hat{A})$ may be less than $n=|\hat{A}|$ as we now demonstrate.

Let

$$\hat{A} = (EF, E^cF, F^c, F).$$

Then

$$\mathscr{C}(\hat{A}) = \{D_1 = EF, D_2 = E^cF, D_3 = F^c, D_4 = F\}.$$

Only the following configurations of "occurrences" can arise so that

$$m = |\pi(\hat{A})| = 3 < 4 = |\hat{A}|:$$

$$\begin{array}{cccccl} 1 & 0 & 0 & 1 & \text{for} & B_1 = EF, \\ 0 & 1 & 0 & 1 & \text{for} & B_2 = E^cF, \\ 0 & 0 & 0 & 1 & \text{for} & B_3 = F^c. \end{array}$$

Thus,

$$B_1^1 B_2^c B_3^c = EF$$
$$B_1^c B_2^1 B_3^c = E^cF$$
$$B_1^c B_2^c B_3 = F^c$$
$$B_1 B_2^c B_3^c \cup B_1^c B_2 B_3^c = F.$$

Of course, $m = 3 > 2$ when $\hat{A} = ((E|F), (E^c|F))$ for $\mathscr{C}(\hat{A}) = \{EF, E^cF, F^c\} = \pi(\hat{A})$.

The (equivalent) reduced form of $G_{f,\psi}$ is

$$G^*_{f,\psi} = (A^*_1, A_2, L^*_{f,\psi}),$$

where

$$A^*_1 = \{(\hat{A}, B) : \hat{A} \in \tilde{\mathscr{A}}_\infty, B \in \pi(\hat{A})\},$$
$$L^*_{f,\psi}(\hat{A}, B, \mu) = \psi(f(\mu(\hat{A}), \hat{A}(B))),$$

where

$$\hat{A}(B) = t \in \{0, 1, u\}^{|\hat{A}|} \Leftrightarrow B = \{\omega : \hat{A}(\omega) = t\},$$

i.e.,

$$\hat{A}(B) = \hat{A}(\omega) \qquad \text{for any choice of } \omega \text{ in } B.$$

Similarly, the (equivalent) reduced form $G_{f,\psi,\hat{A}}$ is

$$G^*_{f,\psi,\hat{A}} = (A^*_{1,\hat{A}}, A_{2,\hat{A}}, L^*_{f,\psi,\hat{A}}),$$

where

$$A^*_{1,\hat{A}} = \pi(\hat{A})$$

and

$$L^*_{f,\psi,\hat{A}}(B, \mu) = \psi(f(\mu(\hat{A}), \hat{A}(B))).$$

3. Analytic Study of Admissibility

In this section, we will introduce various concepts of admissibility for $G_{f,\psi}$ and then develop analytic techniques for (weak) local admissibility with arbitrary aggregation functions ψ. This also extends Lindley's results in the case of an additive aggregation function, namely, giving sufficient conditions for (weak) local admissibility. For related works in Statistics, see [4, 14]. For the concept of Pareto optimality, see [2, 27].

3.1. *Concepts of Admissibility*

In this subsection, we spell out relevant forms of admissibility in the reduced form of the game

$$G^*_{f,\psi} = (\Lambda^*_1, \Lambda_2, L^*_{f,\psi}).$$

First, $\mu \in \Lambda_2$ is (ordinary) *admissible* with respect to $G^*_{f,\psi}$ if there is no $\nu \in \Lambda_2$ such that $L^*_{f,\psi}(\hat{A}, B, \nu) \leqslant L^*_{f,\psi}(\hat{A}, B, \mu)$ for all $(\hat{A}, B) \in \Lambda^*_1$ with the strict inequality for at least one $(\hat{A}, B)$.

Similarly with respect to the subgame $G^*_{f,\psi,\hat{A}}$, $\mu \in \Lambda_{2,\hat{A}} = [a_2, a_3]^{\hat{A}}$ is *$\hat{A}$-admissible* ($\hat{A}$-AD) if there is no $\nu \in \Lambda_{2,\hat{A}}$ such that

$$L^*_{f,\psi,\hat{A}}(B, \nu) \leqslant L^*_{f,\psi,\hat{A}}(B, \mu) \tag{1}$$

for all $B \in \pi(\hat{A})$ with strict inequality for at least one B.

More generally, let $\mathscr{E}$ belong to the power set $\mathscr{P}(\tilde{\mathscr{A}}_\infty)$. Then $\mu \in \Lambda_2$ is *$\mathscr{E}$-admissible* with respect to $G^*_{f,\psi}$ if μ is $\hat{A}$-admissible for all $\hat{A} \in \mathscr{E}$. (Note that $\Lambda_{2,\hat{A}} \subseteq \Lambda_2$.) When $\mathscr{E} = \tilde{\mathscr{A}}_\infty$, μ is *uniformly admissible*.

It is easy to see that uniform admissibility implies ordinary admissibility.

Continuing with the subgame where $\hat{A}$ is fixed, we note that $\mu(\hat{A}) \in \mathbb{R}^n$ for which there is the usual topology based on the Euclidean norm $\| \ \|$. Thus we may have a neighborhood of μ

$$N(\mu, r) = \{\nu \in \Lambda_{2,\hat{A}} \colon \|\nu(\hat{A}) - \mu(\hat{A})\| < r\}.$$

Then $\mu \in \Lambda_{2,\hat{A}}$ is *$\hat{A}$-local admissible* ($\hat{A}$-LAD) if there is some $N(\mu, r)$ such that (1) does not hold for all $\nu \in N(\mu, r)$.

For the last two concepts of admissibility, it will be convenient to introduce the following notation.

With $\hat{A}=(A_1, ..., A_n)$, let $\{B_1, ..., B_m\}$ be a listing of the elements of $\pi(\hat{A})$. The set of values

$$L^*_{f,\psi,\hat{A}}(B_1, \mu)=\psi(f(\mu(A_1), A_1(B_1)), ..., f(\mu(A_n), A_n(B_1))), ...,$$
$$L^*_{f,\psi,\hat{A}}(B_m, \mu)=\psi(f(\mu(A_1), A_1(B_m)), ..., f(\mu(A_n), A_n(B_m)))$$

can be thought of as the value of a transformation $L^*_{f,\psi,\hat{A}}\colon [a_2, a_3]^n \to \mathbb{R}^m$ at the point $\mu(\hat{A})=(\mu(A_1), ..., \mu(A_n))$. For simplicity, we drop the extra symbols and write, in general,

$$L(\hat{x})=\begin{bmatrix} L_1(\hat{x}) \\ \vdots \\ L_m(\hat{x}) \end{bmatrix}$$

for $\hat{x}=(x_1, ..., x_n)\in[a_2, a_3]^n=D$.

By the natures of ψ and f, L can be arranged in the partitioned form

$$\begin{bmatrix} L^{(1)} \\ L^{(2)} \end{bmatrix},$$

where $L^{(2)}$ is constant on D but $L^{(1)}$ (of length k) is not constant in any neighborhood in D. Of course, $L^{(2)}$ may not appear.

Then $\mu\in A_{2,\hat{A}}$ or equivalently $\hat{x}\in D$, is *$\hat{A}$-weakly admissible* ($\hat{A}$-WAD) if there is no $\hat{y}\in D$ such that $L^{(1)}(\hat{y})<_s L^{(1)}(\hat{x})$, where $<_s$ is the strong (Pareto) order in real Euclidean space R^q:

$$\hat{\mu}=(\mu_1, ..., \mu_q)<_s(\nu_1, ..., \nu_q)=\hat{\nu} \qquad \text{if} \quad \mu_j<\nu_j \text{ for all } j=1, 2, ..., q.$$

If $\mu_j\leqslant\nu_j$ for all $j=1, 2, ..., q$ we write simply $\hat{\mu}\leqslant\hat{\nu}$; when the inequality holds for only some j, we write $\hat{\mu}<_w\hat{\nu}$.

It is easy to see that $\hat{A}$-AD is stronger than $\hat{A}$-WAD; for, if there is a $\hat{y}\in D$ such that $L^{(1)}(\hat{y})<_s L^{(1)}(\hat{x})$, then

$$L(\hat{y})=\begin{bmatrix} L^{(1)}(\hat{y}) \\ L^{(2)}(\hat{y}) \end{bmatrix}<_w\begin{bmatrix} L^{(1)}(\hat{y}) \\ L^{(2)}(\hat{y}) \end{bmatrix}$$

since, if $L^{(2)}$ appears, it is constant on D. Moreover, the admissibility considered informally by Lindley turns out to be "weak-local" and will be considered further in Section 4.

Finally, $\hat{x}$ is *weak-local admissible* ($\hat{A}$-WLAD) if for each $\hat{y}$ in R^n with $\|y\|=1$, and each $\alpha>0$ there is an $r=r(\hat{x}, \hat{y}, \alpha)>0$ such that there is no $t\in(0, r)$ for which

$$L^{(1)}(\hat{x}+t\hat{y})-L^{(1)}(\hat{x})\leqslant-\alpha t 1_k. \tag{2}$$

Here 1_k is a k by 1 vector all of whose components are 1.

Note that $-\alpha t 1_k <_s 0$ in R^k. It is convenient to refer to such $\hat{y}$ as a direction. Then, locally, $\hat{x}$ cannot be "beat linearly in any direction."

Of course, $\hat{A}$-WAD implies $\hat{A}$-WLAD: if (2) holds for some $\hat{y}$, α, t, then $L^{(1)}(\hat{x}+t\hat{y}) <_s L^{(1)}(\hat{x})$.

We close this section with a theorem which gives global and local equivalences under the additional hypothesis that L is convex, that is, each component of L is convex.

THEOREM 3.1.1. *Let L be convex. Then*

(i) *$\hat{A}$ admissibility is equivalent to $\hat{A}$ local admissibility;*

(ii) *$\hat{A}$-WAD is equivalent to $\hat{A}$-WLAD.*

Proof. Since it is obvious that admissibility implies local admissibility, we consider only the converses.

(i) In terms of L, $\hat{x}$ is not $\hat{A}$-admissible if there is a $\hat{y}$ in D such that

$$L(\hat{y}) <_w L(\hat{x}). \tag{3}$$

Convexity of L means that for each t in $(0, 1)$,

$$L(\hat{x}+t(\hat{y}-\hat{x})) \leqslant (1-t)\, L(\hat{x}) + tL(\hat{y}).$$

Combining this with (3), we obtain

$$L(\hat{x}+t(\hat{y}-\hat{x})) <_w L(\hat{x}).$$

For any $r>0$, choosing $t = r/(r+\|\hat{y}-\hat{x}\|)$ makes $\|t(\hat{y}-\hat{x})\| < r$ whence $\hat{x}$ is not $\hat{A}$-LAD.

(ii) If $\hat{x}$ is not $\hat{A}$-WAD, there is a $\hat{y}$ such that $L^{(1)}(\hat{y}) <_s L^{(1)}(\hat{x})$. Then $\alpha = \min\{L_j^{(1)}(\hat{x}) - L_j^{(1)}(\hat{y}),\ j = 1(1)k\} > 0$ and $L^{(1)}(\hat{y}) \leqslant L^{(1)}(\hat{x}) - \alpha 1_k$.

Combining this with convexity, we obtain for all $t \in (0, 1)$,

$$\begin{aligned} L^{(1)}(\hat{x}+t(\hat{y}-\hat{x})) &= L^{(1)}((1-t)\hat{x}+t\hat{y}) \\ &\leqslant (1-t)\, L^{(1)}(\hat{x}) + tL^{(1)}(\hat{y}) \\ &\leqslant (1-t)\, L^{(1)}(\hat{x}) + t(L^{(1)} - \alpha 1_k) \\ &= L^{(1)}(\hat{x}) - \alpha t 1_k. \end{aligned}$$

Therefore, $L^{(1)}(\hat{x}+t(\hat{y}-\hat{x})) - L^{(1)}(\hat{x}) \leqslant -\alpha t 1_k <_s 0$ for all $t \in (0, 1)$, in particular, for $t = s/\|\hat{y}-\hat{x}\| < 1$, $t(\hat{y}-\hat{x}) = s\hat{z}$, where $\|\hat{z}\| = 1$ and $\hat{x}$ is not $\hat{A}$-WLAD. ∎

All of the above admissibility concepts—for unrestricted $\tilde{\mathscr{A}}$—can be modified in the obvious way for restrictions such as requiring only finite sequences of conditional events, each having a common antecedent.

3.2. *Some Characterizations of $\hat{A}$-WLAD*

Since WLAD involves only the nonconstant $L^{(1)}$, we simplify by writing this as L. By regularity conditions on ψ and f, L is differentiable. We denote the k by n matrix of partial derivatives at $\hat{x}$ as $J = (\partial L_i/\partial x_j)$; we also take $\hat{x}$, $\hat{y}$ as column vectors. Then we have the following.

THEOREM 3.2.1. *$\hat{x}$ is WLAD if and only if there is no $\hat{y}$ such that $J\hat{y} <_s 0$.*

Proof. If there is a $\hat{y}$ such that $J\hat{y} <_s 0$, then for the direction $\hat{z} = \hat{y}/\|\hat{y}\|$, $J\hat{z} <_s 0$. Differentiability yields

$$\lim_{t \to 0} \frac{L(\hat{x} + t\hat{z}) - L(\hat{x})}{t} = J\hat{z} <_s 0.$$

For each component L_j there is an α_j such that

$$L_j(\hat{x} + t_j\hat{z}) - L(\hat{x}) < -\alpha_j t_j$$

with t_j in a neighborhood of zero. Take t in the intersection of these neighborhoods and $\alpha = \min\{\alpha_1, ..., \alpha_k\}$. Then $L(\hat{x} + t\hat{z}) - L(\hat{x}) < -\alpha t 1_k$, which makes $\hat{x}$ not WLAD. Conversely, if $\hat{x}$ is not WLAD, then there is a direction $\hat{y}$ and an $\alpha > 0$ such that for all r so there is a $t_r < r$ for which $L(\hat{x} + t_r\hat{y}) - L(\hat{x}) \leqslant -\alpha t_r 1_k$. Since $t_r \to 0$ as $r \to 0$, $J\hat{y} <_s 0$. ∎

Recall that for $\hat{A} = (A_1, ..., A_n)$, $\mu \in [a_2, a_3]^{\hat{A}}$ is identified with $\hat{x} = (x_1, ..., x_n) \in D = [a_2, a_3]^n$. In view of Theorem 3.2.1, the analytic study of weak local admissibility of μ or $\hat{x}$ is reduced to finding conditions on $J(\hat{x})$ so that there is no $\hat{y} \in \mathbb{R}^n$ for which $J(\hat{x}) \cdot \hat{y} <_s 0$.

It is well known that the solutions of a system like $J\hat{y} = \hat{z}$ depend heavily on the rank ρ of J. Also, the columns of J and the rows of J, $\hat{y}$ and $\hat{z}$ can be permuted without changing the rank or character of the solutions; these can also be partitioned. In the following, we assume that this has been done so that when the rank is ρ, the system is

$$\begin{bmatrix} J_1 & J_1 \cdot C \\ R \cdot J_1 & R \cdot J_1 \cdot C \end{bmatrix} \cdot \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = \begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \end{bmatrix},$$

where J_1 is ρ by ρ and nonsingular, C is ρ by $n - \rho$ and R is $k - \rho$ by ρ. Of course, if $\rho = n$, the columns involving C do not appear and if $\rho = k$, the rows involving R do not appear.

The following theorem contains several results relevant to this work.

THEOREM 3.2.2. (a) *If* $\rho = n = k$, *then* $\hat{x}$ *is not WLAD.* (*Indeed, if* J *is non-singular, then the system* $J\hat{y} = \hat{z}$ *has a solution* $\hat{y}$ *for every* $\hat{z}$, *in particular* $\hat{z} <_s 0$.)

(b) *If* $n = k$ *and* $\hat{x}$ *is WLAD, then* $\det J = 0$.

(c) $\hat{x}$ *is WLAD if and only if there is no* $\hat{z} \in \mathbb{R}^\rho$ *such that*

$$\begin{bmatrix} J_1 \\ RJ_1 \end{bmatrix} \hat{z} <_s 0.$$

(d) *If* $\rho = 1$, *then* $\hat{x}$ *is WLAD if and only if the vector*

$$\begin{bmatrix} J_1 \\ RJ_1 \end{bmatrix}$$

has both positive and negative components or at least a zero component.

(e) *In case* $k = n = 3$, $\hat{x}$ *is WLAD if and only if* $\det J = 0$ *and either* (i) $\rho = 1$ *and*

$$\begin{bmatrix} J_1 \\ RJ_1 \end{bmatrix}$$

has a zero component or contains both positive and negative components or (ii) $\rho = 2$ *and* $R \leqslant 0$.

Partial Proof. Indeed, if $\det J = 0$, and $\rho = 1$ with the above specified structure of

$$\begin{bmatrix} J_1 \\ RJ_1 \end{bmatrix},$$

then $\hat{x}$ is WLAD as in (d); if $\det J = 0$ and $\rho = 2$, then if $\hat{x}$ is not WLAD, there will be $\hat{y} \in \mathbb{R}^2$ such that

$$\begin{bmatrix} J_1 \\ R \cdot J_1 \end{bmatrix} \hat{y} = \begin{bmatrix} J_1 \hat{y} \\ R \cdot J_1 \hat{y} \end{bmatrix} <_s 0,$$

i.e., $J_1 \hat{y} = \hat{z} <_s 0$ and $R\hat{z} <_s 0$ which is only possible if $R \leqslant 0$.

Conversely, suppose that $\hat{x}$ is WLAD. Then first, $\det J = 0$ as in (b); thus $\rho = 1$ or 2 since $J \equiv 0$. If $\rho = 1$, and $\hat{x}$ is WLAD, we have the above specified structure of

$$\begin{bmatrix} J_1 \\ RJ_1 \end{bmatrix}$$

by (d). If $\rho = 2$, and $\hat{x}$ is WLAD, we have by (c): there is no $\hat{y} \in \mathbb{R}^2$ such that

$$\begin{bmatrix} J_1 \\ RJ_1 \end{bmatrix} \hat{z} <_s 0,$$

i.e., there is no $\hat{w} <_s 0$ such that $R\hat{w} <_s 0$, and hence $R \leqslant 0$. ∎

COROLLARY 3.2.1 (Corresponds to [16, Lemma 1]). *Let $\hat{A} = \{(E \mid F)\}$. Then $\mu(E \mid F) = x$ is WLAD if and only if $x \in [a_0, a_1]$.*

Proof. Here

$$J = \begin{bmatrix} \psi'(f(x, 1))\, f'(x, 1) \\ \psi'(f(x, 0))\, f'(x, 0) \end{bmatrix}.$$

Obviously, for $x \in [a_2, a_3] - [a_0, a_1]$, the components of J are nonzero with the same sign, and hence x is not WLAD. For $x \in [a_0, a_1]$, J has at least a zero component, or two nonzero components with opposite signs, and thus x is WLAD.

Remark. In view of Corollary 3.2.1, from now on, the range of uncertainty measures is restricted to $[a_0, a_1]$.

COROLLARY 3.2.2 (Corresponds to [16, Lemma 2]). *Let $\hat{A} = \{(E \mid F), (E^c \mid F)\}$, then $\hat{x} = (x, y)$, $x = \mu(E \mid F)$, $y = \mu(E^c \mid F)$, is WLAD if and only if* $\det J = 0$.

Proof. Here

$$J = \begin{bmatrix} \psi'[f(x, 1), f(y, 0)]\, f'(x, 1) & \psi'[f(x, 1), f(y, 0)]\, f'(y, 0) \\ \psi'[f(x, 0), f(y, 1)]\, f'(x, 0) & \psi'[f(x, 0), f(y, 1)]\, f'(y, 1) \end{bmatrix}.$$

The condition is necessary by Theorem 3.2.2 (b). Suppose $\det J = 0$, then $\rho = 1$. The sufficiency follows by Theorem 3.2.2 (d). ∎

COROLLARY 3.2.3 (New Result). *For $\hat{A} = \{(E_1 \mid F), (E_2 \mid F), (E_1 \cup E_2 \mid F)\}$ with $E_1 E_2 = \varnothing$; set $x = \mu(E_1 \mid F)$, $y = \mu(E_2 \mid F)$, $z = \mu(E_1 \cup E_2 \mid F)$ with $x, y, z \in [a_0, a_1]$, $\hat{x} = (x, y, z)$. Take $\psi = +$. Then the nonzero losses are $f(x, 1) + f(y, 0) + f(z, 1)$, $f(x, 0) + f(y, 1) + f(z, 1)$, $f(x, 0) + f(y, 0) + f(z, 0)$, and hence*

$$J = \begin{bmatrix} f'(x, 1) & f'(y, 0) & f'(z, 1) \\ f'(x, 0) & f'(y, 1) & f'(z, 1) \\ f'(x, 0) & f'(y, 0) & f'(z, 0) \end{bmatrix}.$$

With respect to the game $G_{f,+,\hat{A}}$, the following are equivalent

(i) $\hat{x}$ *is* $\hat{A}$*-WLAD,*

(ii) $\det J = 0$.

Proof. (i) $\Rightarrow$ (ii) by Theorem 3.2.1.

(ii) $\Rightarrow$ (i). Since $\det J = 0$, the rank ρ of J is either 1 or 2.

(a) If $\rho = 1$, then the first column of J is

$$\begin{bmatrix} f'(a_0, 1) \\ 0 \\ 0 \end{bmatrix} \qquad \text{with} \quad f'(a_0, 1) < 0 \text{ when } x = a_0$$

and is

$$\begin{bmatrix} 0 \\ f'(a_1, 0) \\ f'(a_1, 0) \end{bmatrix} \qquad \text{with} \quad f'(a_1, 0) > 0 \text{ when } x = a_1;$$

for $x \in (a_0, a_1)$, this column has both positive and negative components. Hence $\hat{x}$ is $\hat{A}$-WLAD.

(b) If $\rho = 2$, permute the second and third columns of J to obtain the partition

$$\begin{bmatrix} J_1 & J_1 C \\ RJ_1 & RJ_1 C \end{bmatrix},$$

where

$$J_1 = \begin{bmatrix} f'(x, 1) & f'(z, 1) \\ f'(x, 0) & f'(z, 1) \end{bmatrix}$$

and $RJ_1 = (f'(x, 0), f'(z, 0))$. Then

$$\det J_1 = f'(x, 1)\, f'(z, 1) - f'(x, 0)\, f'(z, 1) > 0$$

when $z \neq a_1$. In this case, we have

$$R = \left[\frac{f'(x, 0)\, f'(z, 1) - f'(x, 0)\, f'(z, 0)}{\det J_1}, \frac{f'(x, 1)\, f'(z, 0) - f'(x, 0)\, f'(z, 1)}{\det J_1} \right]$$

with $f'(x, 0)[f'(z, 1) - f'(z, 0)] \leqslant 0$. Look at

$$f'(x, 1)\, f'(z, 0) - f'(x, 0)\, f'(z, 1). \tag{*}$$

By hypothesis, $\det J = 0$ and this is equivalent to $P_f(z) = P_f(x) + P_f(y)$, where $P_f\colon [a_0, a_1] \to [0, 1]$ is given by

$$P_f(x) = \frac{f'(x, 0)}{f'(x, 0) - f'(x, 1)}. \tag{**}$$

Thus $P_f(x) \leqslant P_f(z)$ which in turn implies that $(*) \leqslant 0$.

When $z = a_1$, the first and third columns of J become

$$\begin{bmatrix} f'(x, 1) \\ f'(x, 0) \\ f'(x, 0) \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ 0 \\ f'(z_1, 0) \end{bmatrix}.$$

Consider

$$\begin{bmatrix} J_1 \\ RJ_1 \end{bmatrix}, \qquad \text{where} \quad J_1 = \begin{bmatrix} f'(x, 1) & 0 \\ f'(x, 0) & f'(a_1, 0) \end{bmatrix}$$

and $RJ_1 = (f'(x, 0), 0)$. We have $\det J_1 = f'(x, 1)\, f'(a_1, 0) < 0$ for $x \neq a_1$. In this case

$$R = \left(\frac{f'(x, 1)\, f'(a_1, 0)}{\det J_1}, 0 \right)$$

with $f'(x, 1)\, f'(a_1, 0) < 0$.

Finally, when $x = z = a_1$, the first and third columns of J are

$$\begin{bmatrix} 0 \\ f'(a_1, 0) \\ f'(a_1, 0) \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ 0 \\ f'(a_1, 0) \end{bmatrix}.$$

Let

$$J_1 = \begin{bmatrix} f'(a_1, 0) & 0 \\ f'(a_1, 0) & f'(a_1, 0) \end{bmatrix} \qquad \text{and} \qquad RJ_1 = (0, 0).$$

Then, $\det J_1 = [f'(a_1, 0)]^2 > 0$, and $R = (0, 0)$. ∎

The following result is actually an application of Corollary 3.2.3. However, as it is basic for most of the rest of our work, we state it as a theorem.

From now on, $\mu \in [a_0, a_1]^{\mathscr{A}}$. We say that:

μ is $\mathscr{E}_2$-WLAD if μ is $\hat{A}$-WLAD for $\hat{A}$ of the form $\{(E \mid F), (E^c \mid F)\}$;

μ is $\mathscr{E}_3$-WLAD if μ is $\hat{A}$-WLAD for $\hat{A}$ of the form $\{(E_1 \mid F), (E_2 \mid F), (E_1 \cup E_2 \mid F)\}$ with $E_1 E_2 = \varnothing$.

Also, $P_f(x)$ is always given by (**) above.

THEOREM 3.2.2′ (Equal Antecedent Counterpart of [16, Lemma 5]). *Suppose $\psi = +$. Then the following are equivalent.*

(i) *μ is $\mathscr{E}_2$ and $\mathscr{E}_3$-WLAD,*

(ii) *$P_f \circ \mu\colon \mathscr{A}_F \to [0, 1]$ is a finitely additive probability measure, for all $F \in \mathscr{A}$, where $\mathscr{A}_F =^d \{(E \mid F) : E \in \mathscr{A}\}$.*

Proof. (i) $\Rightarrow$ (ii). Since μ is $\mathscr{E}_2$-WLAD, we have, for any $E, F \in \mathscr{A}$, $\mu(E \mid F) = x$, $\mu(E^c \mid F) = y$, $\det J = 0$, where

$$J = \begin{bmatrix} f'(x, 1) & f'(y, 0) \\ f'(x, 0) & f'(y, 1) \end{bmatrix} \tag{1}$$

then, $P_f(x) + P_f(y) = 1$.

In particular, for $F = \Omega$, we have

$$P_f \circ \mu(E) + P_f \circ \mu(E^c) = 1.$$

Next, since μ is $\mathscr{E}_3$-WLAD, for any $E_1, E_2, F \in \mathscr{A}$ with $E_1 E_2 = \varnothing$, we have, $\det J = 0$, where, $x = \mu(E_1 \mid F)$, $y = \mu(E_2 \mid F)$, $z = \mu(E_1 \cup E_2 \mid F)$ and

$$J = \begin{bmatrix} f'(x, 1) & f'(y, 0) & f'(z, 1) \\ f'(x, 0) & f'(y, 1) & f'(z, 1) \\ f'(x, 0) & f'(y, 0) & f'(x, 0) \end{bmatrix},$$

so that $P_f(z) = P_f(x) + P_f(y)$, by computation. In particular, for $F = \Omega$,

$$P_f \circ \mu(E_1 \cup E_2) = P_f \circ \mu(E_1) + P_f \circ \mu(E_2).$$

Thus $Q = P_f \circ \mu\colon \mathscr{A}_F \to [0, 1]$ is a finitely additive probability measure, for all $F \in \mathscr{A}$.

(ii) $\Rightarrow$ (i). First note that $\mu\colon \mathscr{A} \to [0, 1]$, (ii) means that the restriction of μ to $\mathscr{A}_F$ (still denoted by μ) is such that $P_f \circ \mu$ is a finitely additive probability on $\mathscr{A}_F$, say $Q_F = P_f \circ \mu$, for all $F \in \mathscr{A}$.

Thus

$$P_f \circ \mu(E \mid F) + P_f \circ \mu(E^c \mid F) = 1$$

which means $\det J = 0$, where J is as in (1). The rank ρ of J is therefore 1. Since $x \in [a_0, a_1]$, each column of J contains a zero component or has both positive and negative components, and hence (x, y) is $\hat{A}$-WLAD, $\forall \hat{A} = \{(E \mid F), (E^c \mid F)\}$, i.e., μ is $\mathscr{E}_2$-WLAD.

The fact that μ is $\mathscr{E}_3$-WLAD follows from Corollary 3.2.3, since for any $E_1, E_2, F \in \mathscr{A}$, with $E_1 E_2 = \varnothing$, the condition

$$P_f \circ \mu(E_1 \cup E_2 \mid F) = P_f \circ \mu(E_1 \mid F) + P_f \circ \mu(E_2 \mid F)$$

is equivalent to $\det J = 0$, where J is as in (2). ∎

In the proof of Theorem 3.2.2′, it might be tempting to conclude that the conditional probability is uniquely compatible with each Q_F. But this is not necessarily so. In fact Aczel [1, pp. 321–324] required additional properties before proving such as relation. These properties include continuity and monotonicity of functional forms in both antecedent and consequent probabilities. The appropriate strengthening of Theorem 3.2.2′ to account for possibly varying conditional event antecedents utilizes the following property:

For any $\mu \in [a_0, a_1]^{\mathscr{A}}$, call μ $\mathscr{E}_4$-WLAD, if μ is $\hat{A}$-WLAD for $\hat{A}$ of the form $((E \mid FG), (F \mid G), (EF \mid G))$, assuming the natural conjunctive chaining relation that both DeFinetti and Lindley implicitly assumed (see earlier discussions),

$$(E \mid FG)(F \mid G) = (EF \mid G); \qquad \text{all} \quad E, F, G \in \mathscr{A},$$

and interpreted via DeFinetti's conditional event indicator function.

THEOREM 3.2.3 (Corresponds to [16, Lemma 5]). *The following statements are equivalent under the above assumption and for* $\psi = +$:

(i) μ *is* $\mathscr{E}_2, \mathscr{E}_3, \mathscr{E}_4$-*WLAD*.

(ii) $P_f \circ \mu \colon \mathscr{A}_F \to [0, 1]$ *is finitely additive conditional probability measure for each* $F \in \mathscr{A}$, *i.e., assuming* $\mu(F) > 0$, *for all* $E \in \mathscr{A}$,

$$(P_f \circ \mu)((E \mid F)) = (P_f \circ \mu)(E \mid F) = P_f(\mu(EF))/P_f(\mu(F)).$$

Proof. Follows a similar format as for the proof of Theorem 3.2.2′, where now, in addition to Eqs. (1) and (2) holding when (i) is assumed, one has due to μ being $\mathscr{E}_4$-WLAD,

$$J = \begin{bmatrix} f'(u, 1) & f'(v, 1) & f'(w, 1) \\ f'(u, 0) & f'(v, 1) & f'(w, 0) \\ 0 & f'(v, 0) & f'(w, 0) \end{bmatrix}, \tag{3}$$

where $u = \mu(E \mid FG)$, $v = \mu(F \mid G)$, $w = \mu(EF \mid G)$. ∎

4. Games with an Additive Aggregation Function and Bayesian Analysis

This section is devoted to a detailed investigation of games with an additive aggregation function; to be complete, we consider also their mixed extensions. Under an additional assumption on the score functions, we establish the equivalences among different concepts of admissibility, including Bayesian decision functions and DeFinetti's coherence measures. We also show that nonatomic probability measures are not "coherent" in games with improper score functions.

4.1. *Mixed Extensions of Uncertainty Games*

Consider the game $G_{f,\psi}$ in its equivalent reduced form $(\Lambda_1^*, \Lambda_2, L_{f,\psi}^*)$ with

$$\Lambda_1^* = \{(\hat{A}, B) : \hat{A} \in \tilde{\mathscr{A}}_\infty, B \in \pi(\hat{A})\}.$$

Because of the nature of Λ_1^*, mixed strategies or prior probability measures on Λ_1^* will be defined as follows.

Player I will first pick an $\hat{A} \in \tilde{\mathscr{A}}_\infty$ according to some probability measure η on $(\tilde{\mathscr{A}}_\infty, \mathscr{B})$, where $\mathscr{B}$ is some σ-algebra of subsets of $\tilde{\mathscr{A}}_\infty$, and then depending upon $\hat{A}$, pick a configuration of occurrences of $\hat{A}$ (as a finite collection of conditional events, equivalently) an element of the partition $\pi(\hat{A})$, according to some probability measure $\tau_{\hat{A}}$ on the power class $\mathscr{P}(\pi(\hat{A}))$ of $\pi(\hat{A})$.

Now, since $\pi(\hat{A})$ is finite, each probability measure on $\mathscr{P}(\pi(\hat{A}))$ is identified by its probability density function on $\pi(\hat{A})$, i.e.,

$$\theta\colon \{B_j, j=1, ..., m\} \to [0, 1], \qquad \sum_{j=1}^{m} \theta(B_j) = 1,$$

where B_j's are some listing of the elements of $\pi(\hat{A})$, and $m = |\pi(\hat{A})|$.

Next, each such θ generates a probability measure P_θ on $(\Omega, \mathscr{A})$ such that

$$P_\theta(B_j) = \theta(B_j), \qquad \forall j = 1, ..., m.$$

Indeed, for each $j = 1, 2, ..., m$, let P_j be a probability measure on B_j, i.e., on the σ-algebra trace $\{AB_j : A \in \mathscr{A}\}$ with $P_j(B_j) = 1$. Define, for $A \in \mathscr{A}$, $P_\theta(A) = \sum_{j=1}^m \theta(B_j)\, P_j(AB_j)$. Note that $P_\theta(A_i)$ is completely determined by θ, $\forall i = 1, ..., n$, where $\hat{A} = (A_1, ..., A_n)$, $n = |\hat{A}|$. Then, since each $A_i \in \tilde{\mathscr{A}}$ in general, the probability measure P_θ on $\mathscr{A}$ is extended to $\tilde{\mathscr{A}}$ via the conditional probability operation. For a rigorous treatment of this extension, see [12].

If $(X, \mathscr{X})$ is a measurable space, we will denote by $\mathbb{P}(X, \mathscr{X})$ the collection of all probability measures on it. The collection of probability density functions on $\pi(\hat{A})$ is denoted by $\mathbb{P}(\pi(\hat{A}))$. Thus, we have, by identification,

$$\mathbb{P}(\pi(\hat{A})) \subseteq \mathbb{P}(\Omega, \mathscr{A}).$$

The space of prior probability measures on Λ_1^* is

$$\mathbb{P}(\Lambda_1^*) = \{(\eta(\cdot), \tau_\cdot(\cdot, \cdot)), \eta \in \mathbb{P}(\tilde{\mathscr{A}}_\infty, \mathscr{B}), \tau_{\hat{A}} \in \mathbb{P}(\pi(\hat{A})), \hat{A} \in \tilde{\mathscr{A}}_\infty\}.$$

The expected loss of $\mu \in \Lambda_2$, with respect to a prior $(\eta(\cdot), \tau_\cdot(\cdot, \cdot))$ is

$$\begin{aligned}
&\rho_{f,\psi}(\eta(\cdot), \tau_\cdot(\cdot, \cdot), \mu) \\
&\quad = \int_{\tilde{\mathscr{A}}_\infty} \int_\Omega L_{f,\psi}(\hat{A}, \hat{A}(\omega), \mu)\, d\tau_{\hat{A}}(\omega)\, d\eta(\hat{A}) \\
&\quad = \int_{\tilde{\mathscr{A}}_\infty} \sum_{j=1}^{|\pi(\hat{A})|} L_{f,\psi}(\hat{A}, B_j, \mu)\, \tau_{\hat{A}}(B_j)\, d\eta(\hat{A}),
\end{aligned}$$

so that the mixed extension of $G_{f,\psi}^*$ is

$$\bar{G}_{f,\psi} = (\mathbb{P}(\Lambda_1^*), \Lambda_2, \rho_{f,\psi}).$$

Similarly, the mixed extension of the subgame $G_{f,\psi,\hat{A}}^*$ is

$$\bar{G}_{f,\psi,\hat{A}} = (\mathbb{P}(\pi(\hat{A})), \Lambda_{2,\hat{A}}, \rho_{f,\psi,\hat{A}}),$$

where $\rho_{f,\psi,\hat{A}}(\theta, \mu) = \sum_{j=1}^{|\pi(\hat{A})|} L_{f,\psi,\hat{A}}(B_j, \mu)\, \theta(B_j)$.

Now, we view the subgame $G_{f,\psi,\hat{A}}^* = (\pi(\hat{A}), [a_2, a_3]^{\hat{A}}, L_{f,\psi,\hat{A}}^*)$ as a *statistical game* coupled with the random variable X whose distribution P_B depends on $B \in \pi(\hat{A})$; when X is constant, on each $B \in \pi(\hat{A})$, the risk is $L(B, \mu)$. On the other hand, since $\pi(\hat{A})$ is finite, standard results from decision theory (for finite games) hold for $G_{f,\psi,\hat{A}}^*$ (see, e.g., [7, 3]). Then the risk set $L(D)$ is closed and bounded since D is compact in $\mathbb{R}^n$ and L is continuous.

For convenience, we state below some definitions and basic results.

(i) $\mu_0 \in [a_0, a_1]^{\hat{A}}$ is Bayes with respect to a prior distribution τ on $\pi(\hat{A})$ if

$$E_\tau L(\cdot, \mu_0) = \inf_{\mu \in [a_0, a_1]^{\hat{A}}} E_\tau L(\cdot, \mu)$$

which is the minimum Bayes risk; E_τ denotes the expectation with respect to τ. We write μ_τ for the Bayes uncertainty measure with respect to τ.

(ii) τ_0 is a least favorable prior if

$$\inf_{\mu} E_{\tau_0} L(\cdot, \mu) = \sup_{\tau} \inf_{\mu} E_{\tau} L(\cdot, \mu)$$

which is the lower value of the game.

(iii) $\mathscr{C} \subseteq [a_0, a_1]^{\hat{A}}$ is complete if for each $\mu \in [a_0, a_1]^{\hat{A}} - \mathscr{C}$, there is $v \in \mathscr{C}$ which is "better" than μ, i.e.,

$$L(\cdot, v) <_{w} L(\cdot, \mu) \qquad (\hat{A}\text{-AD}).$$

$\mathscr{C}$ is minimal complete if $\mathscr{C}$ is complete but no proper subclass of $\mathscr{C}$ is complete.

(iv) For admissibility of Bayes rules in $G^*_{f,\psi,\hat{A}}$, we have

(a) If μ_τ exists uniquely up to equivalence ($\mu \sim v$ if and only if $L(\cdot, \mu) = L(\cdot, v)$), then μ_τ is admissible ($\hat{A}$-AD).

(b) If the prior τ is strictly positive (i.e., $\tau(B) > 0$, $\forall B \in \pi(\hat{A})$), then μ_τ exists, and is $\hat{A}$-AD.

(v) If μ is $\hat{A}$-AD, then $\mu = \mu_\tau$ for some τ.

(vi) The class of Bayes rules is complete, and the class of admissible Bayes rules is minimal complete.

(vii) Let $\mathscr{E} \subseteq \tilde{\mathscr{A}}_\infty$; then μ is said to be *$\mathscr{E}$-Bayes* if μ is Bayes with respect to $G^*_{f,\psi,\hat{A}}$ for all $\hat{A} \in \mathscr{E}$. In particular, μ is *uniform Bayes* if $\mathscr{E} = \tilde{\mathscr{A}}_\infty$. Note that μ is Bayes with respect to $G^*_{f,\psi}$ when the prior is in $\mathbb{P}(A_1)$, i.e., of the form $\xi = \eta(\cdot), \tau_{\cdot}(\cdot, \cdot)$.

4.2. *Equivalences of Various Concepts of Admissibility*

In the rest of this section, we consider $\psi = +$.

THEOREM 4.2.1 (Equal Antecedent Form of [16, Theorem 2]). *Consider the game $G^*_{f,+}$ with f such that P_f is increasing. Let $\mu \in [a_0, a_1]^{\tilde{\mathscr{A}}}$. The following are equivalent.*

(i) *μ is uniformly admissible, w.r.t. all finite equal antecedent conditional event sequences.*

(ii) *μ is $\mathscr{E}_2$ and $\mathscr{E}_3$-weak local admissible.*

(iii) *μ is uniform Bayes, w.r.t. all finite equal antecedent conditional event sequences.*

(iv) *$P_f \circ \mu$: $\mathscr{A}_F \to [0, 1]$ is a finitely additive probability measure, for all $F \in \mathscr{A}$.*

Proof. (i) $\Rightarrow$ (ii). Obvious by definition.

(i) ⇒ (iii). Follows by standard results of the game $G^*_{f,+,\hat{A}}$, namely if μ is $\hat{A}$-AD, then μ is Bayes (with respect to some prior τ on $\pi(\hat{A})$), and conversely, if μ is Bayes ($\mu = \mu_\tau$), then since μ_τ is unique, μ_τ is $\hat{A}$-AD. The uniqueness of μ_τ (up to equivalence) can be seen as follows.

Let $\tau \in \mathbb{P}(\pi(\hat{A}))$, $\hat{A} = ((E_i | F_i), i = 1, ..., n)$; we have

$$E_\tau L(\cdot, \mu) = \sum_{i=1}^{n} \sum_{B \in \pi(\hat{A})} f[\mu(A_i), A_i(B)]\, \tau(B)$$
$$= \sum_{i=1}^{n} \left\{ f[\mu(A_i), 1] \sum_{B \in Q_1(i)} \tau(B) + f[\mu(A_i), 0] \sum_{B \in Q_2(i)} \tau(B) \right\},$$

where

$$Q_1(i) = \{B \in \pi(\hat{A}) : B \subseteq E_i F_i\}$$
$$Q_2(i) = \{B \in \pi(\hat{A}) : BE_i F_i = \varnothing\}.$$

Then

$$\frac{\partial}{\partial x_i} E_\tau L(\cdot, \mu)$$
$$= f'[\mu(A_i), 1] \sum_{B \in Q_1(i)} \tau(B) + f'[\mu(A_i), 0] \sum_{B \in Q_2(i)} \tau(B) = 0$$
$$\Rightarrow \sum_{B \in Q_1(i)} \tau(B) = P_f \circ \mu(A_i)$$

since

$$\sum_{B \in Q_2(i)} \tau(B) = 1 - \sum_{B \in Q_1(i)} \tau(B).$$

Therefore, $\mu(A_i) = P_f^{-1}[\sum_{B \in Q_1(i)} \tau(B)]$ is uniquely determined by τ.

(ii) ⇒ (iv). By Theorem 3.2.2′.

(ii) ⇒ (iii). Assume (ii). Then (iv) holds. Since $\sum_{B \in \pi(\hat{A})} \tau(B) = 1$, $\tau = P_f \circ \mu$ can be taken as a prior for $\pi(\hat{A})$. Then $\mu = \mu_\tau$ and hence (iii) holds. Conversely, when (iii) holds, (i) also holds, and, a fortiori, (ii). ∎

THEOREM 4.2.1′ (Corresponds to [16, Theorem 2]). *Make the same assumptions as in Theorem* 4.2.1. *Then Theorem* 4.2.1 *holds with the following strengthening.*

(1) *Omit the constraint "w.r.t. all finite equal antecedent conditional event sequences" in both* (i) *and* (iii).

(2) *Add the property that μ is $\mathscr{E}_4$-weak local admissible to* (ii).

(3) *Replace $\mathscr{A}_F$, for each $F \in \mathscr{A}$ by simply $\mathscr{A}$ in* (iv).

Proof. Obvious by inspection of the proof of Theorem 4.2.1 and the role that $\mathscr{E}_4$-WLAD plays in making stronger Theorem 4.2.1 (iv). ▮

Remark. For any $\hat{A} \in \tilde{\mathscr{A}}_\infty$, relative to $G^*_{f,+,\hat{A}}$,

(i) A least favorable prior $\tau_0 \in \mathbb{P}(\pi(\hat{A}))$ can be obtained by minimizing the objective function

$$\rho_{f,+,\hat{A}}(\tau, \mu) = \sum_{i=1} \sum_{B \in \pi(\hat{A})} f[\mu(A_i), A_i(B)]\, \tau(B)$$

subject to the constraint $\sum_{B \in \pi(\hat{A})} \tau(B) = 1$.

(ii) Using the proof of the uniqueness of μ_τ in the above theorem, we can obtain a minimax uncertainty measure $\mu_0 \in A_{2,\hat{A}}$, where

$$\inf_{\mu \in A_{2,\hat{A}}} \sup_{\tau \in \mathbb{P}(\pi(\hat{A}))} \rho_{f,+,\hat{A}}(\tau, \mu) = \sup_{\tau \in \mathbb{P}(\pi(\hat{A}))} \rho_{f,+,\hat{A}}(\tau, \mu_0).$$

(iii) $G^*_{f,+,\hat{A}}$ has the value $V_0(f, \hat{A})$, where

$$V_0(f, \hat{A}) = \sup_{\tau \in \mathbb{P}(\pi(\hat{A}))} \rho_{f,+,\hat{A}}(\tau, \mu_0) = \rho_{f,+,\hat{A}}(\tau_0, \mu_{\tau_0}).$$

THEOREM 4.2.2. *Consider the game $G^*_{f,+}$ with P_f increasing. Suppose that f is not a proper score function (i.e., $P_f(x) \equiv x$) and f is twice differentiable. Then no nonatomic conditional probability measure μ on $\mathscr{A}$ can be $G^*_{f,+}$-uniformly admissible.*

Proof. Assume the contrary, i.e., the nonatomic probability μ on $\mathscr{A}$ is uniformly admissible with respect to $G^*_{f,+}$. By Theorem 4.2.1, we should have, by the nonatomicity of μ,

$$P_f(t) + P_f(1-t) = 1, \qquad \forall t \in [0, 1]. \tag{1}$$

By hypothesis here, it can be shown

$$P_f(xy) = P_f(x)\, P_f(y), \qquad \forall x, y \in [0, 1]. \tag{2}$$

Differentiate (2) with respect to x and then separately with respect to y, to obtain

$$y(P_f)'(xy) = (P_f)'(x)\, P_f(y),$$
$$x(P_f)'(xy) = P_f(x)(P_f)'(y).$$

Simple division yields $x[\log P_f(x)]' = y[\log P_f(y)]'$, so that

$$x[\log P_f(x)]' = c \qquad \text{(a constant)},$$

and hence $P_f(x) = x^c$. Then (1) forces c to be 1. ▮

Remarks. (i) To accomodate this situation, we will consider a concept of admissibility in the wide sense of Section 5.

(ii) The condition (iv) in Theorem 4.2.1 is the coherence principle of DeFinetti. The equivalences in Theorem 4.2.1 explain the concept of "reasonable" uncertainty measures, from a decision viewpoint, and lead to the discoveries of other "reasonable" measures which need not be probability measures. In view of Theorem 4.2.1, the coherence principle can be used as a definition of admissibility.

(iii) Relative to the invariance property of the probability transform P_f, we present the following. Consider two games $G_{f,+}$, $G_{g,+}$ (say, with the same a_j, $j = 0, 1, 2, 3$), with P_f, P_g increasing. Let μ (resp. ν) be uniformly admissible with respect to $G_{f,+}$ (resp. $G_{g,+}$). It is not true in general

$$P_f \circ \mu \equiv P_g \circ \nu. \tag{*}$$

Indeed, let Q_1, Q_2 be two probability measures on $\mathscr{A}$, with $Q_1 \neq Q_2$, take $\mu = P_f^{-1} \circ Q_1$ and $\nu = P_g^{-1} \circ Q_2$. For (*) to hold, one needs to consider the uniform admissibility of the (vector-valued) uncertainty measure (μ, ν) with respect to the joint uncertainty game $G_{(f,g),+} = (A_1, A_2^2, L_{(f,g),+})$, where $L_{(f,g),+}(\hat{A}, \omega, \mu, \nu) = L_{f,+}(\hat{A}, \omega, \mu) + L_{g,+}(\hat{A}, \omega, \nu)$.

5. Admissibility of Possibility Measures

This section is devoted to the study of admissibility of a class of uncertainty measures called decomposable measures and its implications for fuzzy logics.

5.1. *Decomposable Measures and Fuzzy Logics*

Since the concept of (Zadeh) possibility measures and the techniques in fuzzy logics might not be familiar to all, we first present some background (see, e.g., [33, 32, 11]).

Roughly speaking, fuzzy logics differ from ordinary two-valued logic by their semantic evaluations of logical connectives. For our purpose here, we will focus on the evaluations of the connective "or" which corresponds to the main properties of associated uncertainty measures.

A function (operator) $T\colon [0, 1] \times [0, 1] \to [0, 1]$ is called a *t-conorm* (see, e.g., [25]) if T is associative, commutative, and nondecreasing in each argument; also $T(x, 0) = x$ and $T(1, x) = 1$, $\forall x \in [0, 1]$.

A t-conorm T is said to be *Archimedean* if T is continuous and $T(x, x) > x$, $\forall x \in (0, 1)$.

Some examples of t-conorms are

$$T(x, y) = \begin{cases} x & \text{if } y = 0 \\ y & \text{if } x = 0 \\ 1 & \text{otherwise} \end{cases} \qquad \text{(not continuous)}$$

$$T(x, y) = \mathrm{Max}(x, y) \qquad \text{(not Archimedean)}$$

$$T(x, y) = \mathrm{Min}(x + y, 1) \qquad \text{(Archimedean).}$$

For the related concept of copulas in Statistics, see, e.g., [9, 10, 19].

An Archimedean t-conorm T has the following representation [18]: T is an Archimedean t-conorm if and only if there exists an increasing, continuous function g (called the *additive generator* or generator of T) which maps $[0, 1] \to [0, +\infty]$ with $g(0) = 0$, and such that

$$\forall x, y \in [0, 1], \qquad T(x, y) = g^*(g(x) + g(y)),$$

where the *pseudo-inverse* g^*: $[0, +\infty] \to [0, 1]$ is defined by

$$g^*(x) = \begin{cases} g^{-1}(x) & \text{if } x \in [0, g(1)] \\ 1 & \text{if } x > g(1). \end{cases}$$

For example,

(i) For $p > 0$, $T_p(x, y) = (x^p + y^p - x^p y^p)^{1/p} = g_p^*(g_p(x) + g_p(y))$, where $g_p(x) = -(1/p) \log(1 - x^p)$, $g_p^{-1}(x) = (1 - e^{-px})^{1/p} = g_p^*(x)$; note that $g_p(1) = +\infty$ here.

(ii) For $p \geqslant 1$, $T_p(x, y) = [\mathrm{Min}(x^p + y^p, 1)]^{1/p}$ has generator $g_p(x) = x^p$,

$$g^*(x) = \begin{cases} x^{1/p} & \text{if } x \in [0, 1] \\ 1 & \text{if } x > 1 \end{cases}$$

with $g_p(1) = 1$ here.

Since a t-conorm T is associative and commutative, we can extend T to

$$T\colon [0, 1]_\infty \to [0, 1], \qquad \text{where} \quad T(x) = x, \text{ by convention,}$$

$$T(x_1, ..., x_n) = T(x_1, T(x_2, ..., x_n)), \qquad n \geqslant 2.$$

The representation of an Archimedean t-conorm T becomes

$$T(x_1, x_2, ..., x_n) = g^*\left(\sum_{i=1}^{n} g(x_i)\right), \qquad n \geqslant 1.$$

Now, let $(\Omega, \mathscr{A})$ be a measurable space. A mapping μ from $\mathscr{A}$ to some interval of $\mathbb{R}$, say $[a_2, a_3]$, is called a *decomposable measure* if there exists $T: [a_2, a_3] \times [a_2, a_3] \to [a_2, a_3]$ such that, for $A, B \in \mathscr{A}$ with $AB = \varnothing$, $\mu(A \cup B) = T(\mu(A), \mu(B))$ (see, e.g., [31]). When such a T exists, it is called the *composition law* of μ.

As a generic example of a decomposable measure, begin with any fuzzy set membership function $\alpha: \Omega \to [0, 1]$ and any t-conorm T. Then, if Ω is finite, one can make the extension of α as $\mu_{\alpha, T}: P(\Omega) \to [0, 1]$, where for any $A \subseteq \Omega$,

$$\mu_{\alpha, T}(A) = T(\alpha(\omega) : \omega \in A). \tag{+}$$

$\mu_{\alpha, T}$ is clearly decomposable. Conversely, given any decomposable measure μ w.r.t. some t-conorm T, for any finite A as above, Eq. (+) holds with $\mu_{\alpha, T} = \mu$.

In fuzzy logics, composition laws are t-conorms (with $[a_2, a_3] = (0, 1])$. Note that probability measures are decomposable measures with $T(x, y) = \mathrm{Min}(x + y, 1)$; this is an Archimedean t-conorm with generator $g(x) = -\log(1 - x)$ and $g^*(x) = g^{-1}(x) = 1 - e^{-x}$, since $g(1) = +\infty$. Indeed, let P be a probability measure on $\mathscr{A}$, and $A, B \in \mathscr{A}$ with $AB = \varnothing$. Since $P(A) + P(B) = P(A \cup B) \leqslant 1$, we have $P(A \cup B) = P(A) + P(B) = \mathrm{Min}(P(A) + P(B), 1)$. Furthermore, in the case of probability measures, the generator g is such that $g(1) = +\infty$ (and hence $g^* = g^{-1}$); the corresponding t-conorm T is called a *strict* (Archimedean) t-conorm.

For a t-conorm T, a *T-possibility measure* is defined to be a map from $\mathscr{A}$ to $[0, 1]$ with T as composition law. For example, (Zadeh) *max-possibility measure* is a decomposable measure with $T(x, y) = \mathrm{Max}(x, y)$.

Let Ω be discrete and $\alpha: \Omega \to [0, 1]$. Let T be an Archimedean t-conorm with generator g. We denote by $\mu_{\alpha, T}$ the T-possibility measure defined as follows.

For $A \subseteq \Omega$, finite,

$$\mu_{\alpha, T}(A) = T(\alpha(\omega), \omega \in A) = g^*\left[\sum_A g(\alpha(\omega))\right].$$

For A countably infinite,

$$\mu_{\alpha, T}(A) = g^*\left[\sum_A g(\alpha(\omega))\right].$$

Note that $\sum_A g(\alpha(\omega)) \leqslant +\infty$.

5.2. *General Admissibility under Additive Aggregation*

In order to discuss Lindley's conclusions about the inadmissibility of uncertainty measures, we consider $G_{f, +}$. In view of the results of Sections

3 and 4, we have to consider the concept of admissibility of a given uncertainty measure in a wide sense. Specifically, an uncertainty measure $\mu \in [a_0, a_1]^{\tilde{\mathscr{A}}}$ is said to be *general admissible* if there is a game $G_{f,+}$ such that μ is uniformly admissible with respect to that game. In this sense, any probability measure is general admissible by taking the score function f to be a proper score function! But for a proper score function f, the associated probability transform $P_f(x) = x$, $\forall x \in [0, 1]$, is increasing, and hence P_f^{-1} exists. So we require, in general admissibility, the existence of a score function f such that P_f^{-1} exists. It is seen that with respect to a game $G_{f,+}$ with P_f increasing, μ is general admissible if and only if $P_f \circ \mu$ is a finitely additive probability. It is this equivalent property that we will use as a definition for general admissibility.

In discussing possibility measures on discrete spaces, we can even consider a stronger concept of admissibility, namely *general admissibility in the σ-additive sense*, i.e., $P_f \circ \mu$ is a σ-additive probability measure.

It will be shown in this subsection that operations in fuzzy logics are, in general, "admissible," and even (Zadeh)-max possibility measures are (uniform) limits of admissible decomposable measures. For related works in Statistics see, e.g., [15].

Throughout this subsection, Ω is a discrete space (finite or countably infinite), $\mathscr{A} = \mathscr{P}(\Omega)$ and restrict $\tilde{\mathscr{A}}$ to all $\mathscr{A}_F$, $F \in \mathscr{A}$. For $\mu: \mathscr{P}(\Omega) \to [0, 1]$, we write $\mu(\{\omega\}) = \mu(\omega)$, so that the restriction of μ to singletons is regarded as a function, still denoted as μ, from Ω to $[0, 1]$. We also omit, from now on, the qualification "w.r.t. all finite equal antecedent conditional event sequences."

THEOREM 5.2.1. *Let $\mu: \mathscr{P}(\Omega) \to [0, 1]$. Then the following are equivalent*:

(i) *μ is general admissible in the σ-additive sense.*

(ii) *μ is a decomposable measure with composition law T being an Archimedean t-conorm with generator g such that $g(1) = 1$, and*

$$\sum_{\Omega} g(\mu(\omega)) = 1. \tag{*}$$

Proof. (i) $\Rightarrow$ (ii). Let f be a score function such that P_f is increasing and such that $P_f \circ \mu$ is a σ-additive probability measure.

For $A \in \mathscr{P}(\Omega)$, we have

$$P_f(\mu(A)) = \sum_{A} P_f(\mu(\omega)),$$

and hence $\mu(A) = P_f^{-1}(\sum_A P_f(\mu(\omega)))$. By taking $g = P_f$, we have $g(1) = 1$, and we see that μ is decomposable with

$$T(x, y) = P_f^*(P_f(x) + P_f(y)).$$

Of course,

$$1 = P_f \circ \mu(\Omega) = \sum_{\Omega} P_f(\mu(\omega)).$$

(ii) $\Rightarrow$ (i). Since μ is T-decomposable on a discrete space, we have $\forall A \in \mathscr{P}(\Omega)$,

$$\mu(A) = T(\mu(\omega), \omega \in A).$$

Take $P_f = g$ (noting that we can solve for f), we have

$$\begin{aligned} P_f \circ \mu(A) &= g[T(\mu(\omega), \omega \in A)] \\ &= g\left[g^*\left(\sum_A g(\mu(\omega))\right)\right] = \sum_A g(\mu(\omega)) \end{aligned}$$

since, by hypothesis $\sum_A g(\mu(\omega)) \leqslant \sum_\Omega g(\mu(\omega)) = 1 = g(1)$. Thus $P_f \circ \mu$ is a σ-additive probability measure. ∎

Remark. In view of the above theorem, we see that if μ is T-composable (or equivalently, if μ is generated by its restriction to singletons and T) and if μ is general admissible in the σ-additive sense, then $\mu = \delta_{\omega_0}$, the Dirac (probability) measure at $\omega_0 \in \Omega$ when $\mu(\omega_0) = 1$.

The following result provides a necessary and sufficient condition for μ to be general admissible in the σ-additive sense when $\sup_\Omega \mu(\omega) < 1$.

THEOREM 5.2.2. *Let $\alpha: \Omega \to [0, 1]$ with Ω countably infinite, such that* $\sup_\Omega \alpha(\omega) < 1$ $(\alpha \not\equiv 0)$.

Then a necessary and sufficient condition for the existence of a T-possibility measure μ, generated by α and some T, which is general admissible in the σ-additive sense, is that

$$\forall x \in (0, 1], \qquad \alpha^{-1}[x, 1] = \{\omega : x \leqslant \alpha(\omega) \leqslant 1\} \tag{**}$$

is finite. In particular, if Ω is finite and $\sup_\Omega \alpha(\omega) < 1$, *then there exists T such that $\mu_{\alpha, T}$ is general admissible in the σ-additive sense.*

Proof. (a) *Necessity.* If there is $x_0 \in [0, 1]$ such that $\alpha^{-1}[x_0, 1]$ is infinite, then

$$\lim_{n \to +\infty} \sum_{A_n} g(\mu(\omega)) = +\infty,$$

where $A_n \subseteq \alpha^{-1}[x_0, 1]$. For, although A_n is finite, and $A_n \uparrow \alpha^{-1}[x_0, 1]$ for each $n \geqslant 1$, $\sum_{A_n} g(\mu(\omega)) \geqslant g(x_0)\,|A_n|$. Thus (*) of Theorem 5.2.1 will not hold.

(b) *Sufficiency.* In view of Theorem 5.2.1, we need to show only the existence of a generator g such that $g(1) = 1$ and $\sum_\Omega g(\alpha(\omega)) = 1$. Then we can take $\mu_{\alpha,T}(A) = g^*[\sum_A g(\alpha(\omega))]$, where T is the Archimedean t-conorm with generator g.

Let $0 < x_0 = \sup_\Omega \alpha(\omega) < 1$. Let $\{\alpha > 0\} = \{\omega : \alpha(\omega) > 0\} = \bigcup_{n=2}^{+\infty} K_n$, where $K_n = \{\omega : 1/n < \alpha(\omega) \leqslant 1/(n-1)\}$, $n \geqslant 2$. Let $n_0 \geqslant 2$ such that $1/n_0 < x_0 \leqslant 1/(n_0 - 1)$. By (**), $\alpha^{-1}[1/n_0, 1)$ is finite; thus $\{\alpha(\omega) : \omega \in K_{n_0}\} = \{x_1, x_2, ..., x_{n_1}\}$ for some n_1. We can assume

$$\frac{1}{n_0} < x_1 < x_2 < \cdots < x_{n_1} \leqslant \frac{1}{n_0 - 1}.$$

Note that (**) implies that $\sup_\Omega \alpha(\omega)$ is attained at ω', where $\alpha(\omega') = x_{n_1} = x_0$. (Indeed, if $\alpha(\omega'') > \alpha(\omega') = x_{n_1}$, then $\omega'' \in K_{n_0}$ and hence contradicts the definition

$$x_{n_1} = \operatorname*{Max}_{K_{n_0}} \alpha(\omega)).$$

Since $[0, 1] = [0, 1/n_0] \cup (1/n_0, 1]$, we first construct g on $[0, 1/n_0]$.

For $n \geqslant n_0$ and $r > 0$, define

$$g_r(0) = 0$$

and

$$g_r\left(\frac{1}{n}\right) = \left(\frac{1}{n}\right)^r \Bigg/ \sum_{j=2}^{n} (|K_j| + 1) < 1.$$

Then $1/m < 1/n \Rightarrow g_r(1/m) < g_r(1/n)$ and $\lim_{n \to \infty} g_r(1/n) = 0$. Thus we construct a continuous, increasing function on $[0, 1/n_0]$ by extending g_r continuously on each $[1/n, 1/(n-1)]$, for $n > n_0$, (say by joining $g_r(1/n)$ and $g_r(1/(n-1))$ by a straight line).

On $[1/n_0, 1]$, we proceed as follows. Let

$$\alpha(r) = \sum_{K_{n0}^c} g_r(\alpha(\omega)).$$

Since $x_0 \in K_{n_0}$, $K_{n_0}^c = \{\omega : \alpha(\omega) \leqslant 1/n_0\}$; i.e., for $\omega \in K_{n_0}^c$, $\alpha(\omega) \in [0, 1/n_0]$, which is the domain of g_r defined above.

We have

$$\begin{aligned}
0 \leqslant a(r) &= \sum_{n=n_0+1}^{+\infty} \sum_{K_n} g_r(\alpha(\omega)) \\
&\leqslant \sum_{n=n_0+1}^{+\infty} (|K_n|)\, g_r\left(\frac{1}{n-1}\right) \leqslant \sum_{n=n_0+1}^{+\infty} \frac{1}{(n-1)^r} \\
&= \sum_{n=n_0}^{+\infty} n^{-r}.
\end{aligned}$$

Thus $\lim_{r \to \infty} a(r) = 0$.

Let r be large enough so that

$$\max\left(a(r), g_r\left(\frac{1}{n_0}\right)\right) < \frac{1}{1+\sum_{j=1}^{n_1} |C_j|},$$

where $C_j = \alpha^{-1}(x_j) \subseteq K_{n_0}$, $\forall j = 1, ..., n_1$. There exist real numbers y_j, $j = 1, ..., n$ such that

$$\max\left(a(r), g_r\left(\frac{1}{n_0}\right)\right) < y_1 < y_2 < \cdots < y_n < 1$$

and

$$a(r) + \sum_{j=1}^{n_1} |C_j|\, y_j = 1$$

(see the lemma below). Thus, on $[1/n_0, 1]$, define $g_r(x_j) = y_j$, $j = 1, ..., n_1$, for r large enough. Extend g_r by continuity to $[1/n_0, x_1]$ and on each $[x_j, x_{j+1}]$, where $x_{n_1+1} = 1$, and $g_r(1) = 1$. The condition (*) of Theorem 5.2.1 *is satisfied. Indeed*

$$\sum_{K_{n_0}} g_r(\alpha(\omega)) = \sum_{j=1}^{n_1} |C_j|\, g_r(x_j) = \sum_{j=1}^{n_1} |C_j|\, y_j$$

so that

$$\sum_{\Omega} g_r(\alpha(\omega)) = \sum_{K_{n0}^c} g_r(\alpha(\omega)) = \sum_{K_{n_0}} g_r(\alpha(\omega))$$
$$= a(r) + \sum_{j=1}^{n_1} |C_j|\, y_j = 1.$$

Finally, take $\mu_{\alpha,T}(A) = G(\alpha(\omega))$, $\omega \in A = g^{-1}[\sum_A g(\alpha(\omega))]$, where $g = g_r$ for r large enough, $g^* = g^{-1}$, since $\forall A \subseteq \Omega$, $\sum_A g(\alpha(\omega)) \leqslant \sum_\Omega g(\alpha(\omega)) = 1 = g(1)$, and T is the Archimedean t-conorm with generator g. ∎

LEMMA. *Let n_j, $j = 1, ..., k$, be k positive integers. Let $q_1, q_2 \in \mathbb{R}^+$ such that*

$$0 \leqslant q_0 = \operatorname{Max}(q_1, q_2) < \frac{1}{1+\sum_{j=1}^{k} n_j}.$$

Then there exist real numbers $y_1, ..., y_k$ such that

(i) $q_0 < y_1 < y_2 < \cdots < y_k < 1,$

(ii) $q_1 = \sum_{j=1}^{k} n_j y_j = 1.$

Proof. Let $y_j = q_0 + j/c$, where

$$c = \left[\sum_{j=1}^{k} jn_j\right]\left[1 - q_1 - q_0 \sum_{j=1}^{k} n_j\right]^{-1} > 0.$$

Thus $q_0 < y_1 < y_2 < \cdots < y_k$.

Now $y_n < 1$ if and only if $q_0 \sum_{j=1}^{k} jn_j = k(1 - q_1 - q_0 \sum_{j=1}^{k} n_j) < \sum_{j=1}^{k} jn_j$, if and only if $k(1-q_1) < (1-q_0)\sum_{j=1}^{k} jn_j + kq_0 \sum_{j=1}^{k} n_j = d_k$. But $d_k \geqslant e_k \stackrel{d}{=} (1-q_0)\sum_{j=1}^{k} j + kq_0 \sum_{j=1}^{k} 1 = (1-q_0)(k(k+1)/2) + k^2 q_0$.

Now $(1-q_1)k \leqslant e_k$ if and only if

$$1 - q_1 \leqslant \frac{k+1}{2}(1-q_0) + kq_0$$

$$\text{if and only if} \quad 0 \leqslant \frac{k-1}{2}(1+q_0) + q_1,$$

which is true here since $q_0, q_1 \geqslant 0$, and $k \geqslant 1$. For (ii)

$$\begin{aligned} q_1 + \sum_{j=1}^{k} n_j \cdot y_j &= q_1 + \sum_{j=1}^{k} n_j (q_0 + j/c) \\ &= q_1 + q_0 \sum_{j=1}^{k} n_j + \left(1 - q_1 - q_0 \sum_{j=1}^{k} n_j\right) = 1. \quad \blacksquare \end{aligned}$$

If μ is T-decomposable with $T(x, y) = \text{Max}(x, y)$ then μ is not general admissible even in the additive sense. This is due essentially to the fact that $\text{Max}(x, y)$ is not an Archimedean t-conorm. However, $\text{Max}(x, y)$ can be approximated by Archimedean ones.

For example, let $T_p(x, y) = [\text{Min}(x^p + y^p, 1)]^{1/p}$, $p \geqslant 1$. For each $p \geqslant 1$, T_p is an Archimedean t-conorm with generator $g_p(x) = x^p$, $g_p(1) = 1$.

As usual, we extend T_p to n arguments, as $T_p(x_1, x_2, ..., x_n) = T_p(x_1, T_p(x_2, ..., x_n))$.

Then for each fixed n, $T_p(x_1, x_2, ..., x_n) \to \text{Max}(x_1, x_2, ..., x_n)$ as $p \to +\infty$, uniformly in $(x_1, x_2, ..., x_n)$. Indeed, since

$$\begin{aligned} \text{Max}(x_1, x_2, ..., x_n) &\leqslant T_p(x_1, x_2, ..., x_n) \\ &\leqslant \text{Min}[\text{Max}(x_1, ..., x_n) n^{1/p}, 1] \end{aligned}$$

we have

$$0 \leqslant T_p(x_1, x_2, ..., x_n) = \text{Max}(x_1, x_2, ..., x_n) \leqslant n^{1/p} - 1.$$

Thus, if Ω is finite, and $\mu: \mathscr{P}(\Omega) \to [0, 1]$ has $\text{Max}(x, y)$ as composition law, i.e., $\forall A \subseteq \Omega$,

$$\mu(A) = \text{Max}_{A}\, \mu(\omega),$$

then

$$\mu(A) = \lim_{P \to +\infty} T_p(\mu(\omega), \omega \in A),$$

uniformly in A.

Now, it is easy to see that, in Theorem 5.2.1, if we require only general admissibility in the additive sense (which is equivalent to uniform admissibility) the condition (*) can be weakened to

$$\sum_{\Omega} g(\mu(\omega)) \leqslant 1.$$

Thus, assuming in addition that $\sum_{\Omega} \mu(\omega) \leqslant 1$, for all $\forall p \geqslant 1$, $v_p(A) = T_p(\mu(\omega), \omega \in A)$ is general admissible in the additive sense, since

$$\sum_{\Omega} g_p(v_p(\omega)) = \sum_{\Omega} [\mu(\omega)]^p \leqslant 1.$$

Therefore, in this case, Max-possibility measures are uniform limits of general admissible T_p-possibility measures.

More generally, for Ω countably infinite, and $\alpha: \Omega \to [0, 1]$ such that $\sup_{\Omega} \alpha(\omega) < 1$, the max-possibility measure generated by α is

$$\mu_\alpha(A) = \sup_{A} \alpha(\omega)$$

and this can be approximated by admissible measures.

Specifically,

THEOREM 5.2.3. *Let Ω be countably infinite and $\alpha: \Omega \to [0, 1]$ such that $\sup_{\Omega} \alpha(\omega) < 1$. Suppose that there are non-negative reals a, b such that*

$$\left| \alpha^{-1} \left[\frac{1}{n}, \frac{1}{n-1} \right] \right| \leqslant an^b, \qquad \forall n \geqslant 2. \tag{*}$$

Then

(i) *For $A \in \mathscr{P}(\Omega)$, such that $A\alpha^{-1}(t_1) = \varnothing$, $\mu_\alpha(A) = \lim_{P \to +\infty} \mu_{\alpha, T_p}(A)$, the limit being uniform in all such A with $|A| \leqslant n_0$, for any fixed positive integer n_0; also $t_1 = \sup_{\Omega} \alpha(\omega)$.*

(ii) *For $A\alpha^{-1}(t_1) \neq \varnothing$,*

$$|\mu_{\alpha, T_p}(A) - \mu_\alpha(A)| \leqslant 1 - t_1.$$

Proof. (i) We are now going to construct generators g_p, for sufficiently large p, such that

$$\mu_{\alpha,p}(A) = g_p^{-1}\left[\sum_A g_p(\alpha(\omega))\right]$$

are general admissible in the σ-additive sense.

More specifically, not only is g_p such that $g_p(1) = 1$ and $\sum_\Omega g_p(\alpha(\omega)) = 1$, but precisely $g_p(x) = x^p$ on $[0, t_2]$, where

$$t_2 > t_1,\ t_2 = \sup_{C_1^c} \alpha(\omega) \qquad (>0 \text{ by assumption}) \qquad \text{and} \qquad C_1 = \alpha^{-1}(t_1).$$

First $\forall p > 0$, define $g_p(x) = x^p$ on $[0, t_2]$. We will extend g_p to $[0, 1]$ by defining a value for $g_p(t_1)$ such that $t_2 < g_p(t_1) < 1$, set $g(1) = 1$ (with extension by continuity as usual) and obtain

$$\sum_\Omega g_p(\alpha(\omega)) = 1.$$

Let $\alpha(\Omega) = \{t_j, j = 1, 2, ...\}$ and $C_j = \alpha^{-1}(t_j)$.

$$\Phi_{p,2} = \sum_{C_1^c} \alpha^p(\omega) = \sum_{j=2}^{+\infty} |C_j|\, t_j^p = \Phi_{p,3} + |C_2|\, t_2^p,$$

where

$$\begin{aligned}\Phi_{p,3} &= \sum_{C_1^c C_2^c} \alpha^p(\omega) = \sum_{j=3}^{+\infty} |C_j|\, t_j^p = \sum_{n=2}^{+\infty} \sum_{1/n < t_j \leqslant 1/(n-1)} |C_j|\, t_j^p \\ &= \left[\sum_{1/2 < t_j \leqslant 1} |C_j|\, t_j^p\right] + \sum_{n=3}^{+\infty} \sum_{1/n < t_j \leqslant 1/(n-1)} |C_j|\, t_j^p.\end{aligned}$$

By (*), the first term is at most, $a2^b$ times (a finite sum of t_j^p),

$$0 \leqslant t_j < 1; \text{ this sum tends to zero as } p \to +\infty.$$

In the second term

$$\sum_{n=3}^{+\infty} \sum_{1/n < t_j \leqslant 1/(n-1)} |C_j|\, t_j^p,$$

$$\sum_{1/n < t_j \leqslant 1/(n-1)} |C_j| \geqslant \left|\alpha^{-1}\left[\frac{1}{n}, \frac{1}{n-1}\right]\right| \leqslant an^b \qquad \text{by } (*).$$

Therefore, this second term is bounded by

$$a \sum_{n=3}^{+\infty} n^b (n-1)^{-p} = a \sum_{n=2}^{+\infty} \frac{(n+1)^b}{n^p}$$
$$\leqslant a \sum_{n=2}^{+\infty} \left(\frac{3n}{2}\right)^b \frac{1}{n^p} = a(3/2)^b \sum_{n=2}^{+\infty} n^{-(p-b)}.$$

Thus $\lim_{p \to +\infty} \Phi_{p,3} = 0$.

Let p be large enough so that $\Phi_{p,3} < \frac{1}{2}$ and $t_2^p < 1/2(|C_1| + |C_2|)$.

Define $g_p(t_1) = (1 - \Phi_{p,2})/|C_1|$. We have $t_2^p < g_p(t_1) < 1$ if and only if

$$\Phi_{p,2} + |C_1|\, t_2^p < 1 < |C_1| + \Phi_{p,2}$$

if and only if

$$\Phi_{p,3} + (|C_1| + |C_2|)t_2^p < 1 < |C_1| + \Phi_{p,2}$$

which is true by construction ($\Phi_{p,2} > 0$ since $t_2 > 0$). Also, by construction, we have

$$1 = \Phi_{p,2} + |C_1|\ g_p(t_1) = \sum_{\Omega} g_p(\alpha(\omega)).$$

Thus, if $A\alpha^{-1}(t_1) = \varnothing$,

$$\mu_{\alpha, T_p}(A) = g_p^{-1}\left[\sum_A g_p(\alpha(\omega))\right] = g_p^{-1}\left[\sum_A \alpha^p(\omega)\right]$$
$$= T_p(\alpha(\omega), \omega \in A) \text{ with } T_p(x, y) = [\mathrm{Min}(x^p + y^p, 1)]^{1/p}.$$

Hence $\lim_{p \to +\infty} \mu_{\alpha, T_p}(A) = \mathrm{Max}\{\alpha(\omega), \omega \in A\}$ for any finite A.

(ii) For $A\alpha^{-1}(t_1) \neq \varnothing$, we have

$$t_1 = \mathrm{Max}\{\alpha(\omega), \omega \in A\} \leqslant T_p(\alpha(\omega), \omega \in A) \leqslant 1$$

and hence

$$|\mu_{\alpha, T_p}(A) - \mu_\alpha(A)| \leqslant 1 - t_1. \quad \blacksquare$$

Remarks. (i) Let $(\Omega, \mathscr{A})$ be an arbitrary measurable space, and $\mu: \mathscr{A} \to [a_0, a_1]$ be general admissible in the σ-additive sense with $P_f \circ \mu = Q$, where Q is a discrete probability measure on $(\Omega, \mathscr{A})$. Then μ has the same representation as in the discrete case. Indeed, let $\Omega_0 \subseteq \Omega$, countable such that $Q(\Omega_0) = 1$. Then $\forall A \in \mathscr{A}$,

$$\mu(A) = P_f^{-1} \circ Q(A) = P_f^{-1}\left(\sum_{A\Omega_0} P_f(\mu(\omega))\right).$$

(ii) A continuous analog is as follows. Let $(\Omega, \mathscr{A}) = (\mathbb{R}, \mathscr{B})$, $F_Q: \mathbb{R} \to [0, 1]$, $F_Q(x) = Q((-\infty, x])$ and $F_\mu: \mathbb{R} \to [a_0, a_1]$, $F_\mu(x) = \mu((-\infty, x])$. Since $\mu = P_f^{-1} \circ Q$, and P_f^{-1} is increasing, μ is a nondecreasing set-function; hence F_μ is nondecreasing, and $A \in \mathscr{B}$, $\mu(A) = P_f^{-1}[\int_A d(P_f \circ F_\mu(x))]$. If, in addition, P_f is differentiable, then

$$\mu(A) = P_f^{-1}\left[\int_A P_f'(F_\mu(x))\, dF_\mu(x)\right].$$

6. Admissibility of Dempster–Shafer Belief Functions

Dempster–Shafer belief functions have become very popular in recent years for modeling aspects of expert systems and combination of evidence problems in *AI*. The purpose of this section is to respond to Lindley's comments about the inadmissibility of belief functions [16].

For simplicity let Ω be a finite set. A belief function Bel on $\mathscr{P}(\Omega)$, the power class of Ω, can be defined as follows. Let $m: \mathscr{P}(\Omega) \to [0, 1]$ such that $m(\varnothing) = 0$, $\sum_{\mathscr{P}(\Omega)} m(A) = 1$. Then,

$$\mathrm{Bel}(B) = \sum_{A \subseteq B} m(A).$$

m is the probability allocation for Bel. By the Möbius inversion formula, m can be recovered from Bel,

$$m(A) = \sum_{B \subset A} (-1)^{|A - B|}\, \mathrm{Bel}(B),$$

where $|A|$ denotes the cardinality of A. Note that if $\mu: \mathscr{P}(\Omega) \to [0, 1]$ is such that $\mu(\Omega) = 1$ and

$$\forall A \subseteq \Omega, \sum_{B \subseteq A} (-1)^{|A - B|}\, \mu(B) \geqslant 0$$

then μ is a belief function.

If we think of "sets" as "points," then m plays the role of a probability mass function, and Bel is a "cumulative distribution function." Since Ω is finite, we have

$$\mathrm{Bel}(A) = P(X \in \mathscr{P}(A)), \qquad \forall A \subseteq \Omega,$$

where X is a random set, defined on some probability space $(\Omega, \mathscr{F}, P)$, and taking values in $\mathscr{P}(\Omega)$ with "density" m, i.e.,

$$P(\theta : X(\theta) = A) = m(A).$$

Note that

$$\mathrm{Bel}(A) + \mathrm{Bel}(A^c) \leqslant 1.$$

We extend Bel to conditional events $\tilde{\mathscr{P}}(\Omega)$ as follows. For $E, F \in \mathscr{P}(\Omega)$, define

$$\mathrm{Bel}(E \mid F) = P(X \subseteq E \mid X \subseteq F)$$

when $P(X \subseteq F) > 0$.

For more information on belief functions, we refer the reader to [26, 21, 11, 30].

As in Section 5, we consider $G_{f,+}$ with P_f increasing. By the nature of belief functions, we consider $[a_0, a_1] = [0, 1]$. Note also that the existence of f such that P_f is increasing is equivalent to that of a (surjective) continuous, increasing function $h: [0, 1] \to [0, 1]$. Thus, in view of Theorem 4.2.1, $\mu \in [0, 1]^{\mathscr{A}}$ is general admissible if and only if there is such an h for which $h \circ \mu$ is a finitely additive probability measure.

In discussing the admissibility of belief functions on Ω finite, $\mathscr{A} = \mathscr{P}(\Omega)$, it should be noted that the range of a belief function is not the whole interval $[0, 1]$! As we will see, as in the case of fuzzy logics, some classes of belief functions are admissible while others are not. Thus, if DeFinetti's coherence principle is viewed as a rational way of choosing "reasonable" uncertainty measures, the following analysis will provide criteria for selecting "good" belief functions.

First, a simple condition of inadmissibility.

THEOREM 6.1. *Let $(\Omega, \mathscr{A})$ be a measurable space and $\mu \in [0, 1]^{\mathscr{A}}$. If there is $E, F \in \mathscr{A}$ such that*

$$\mu(E) = \mu(F) \qquad \text{and} \qquad \mu(E^c) \neq \mu(F^c),$$

then μ is not general admissible.

Proof. If μ were general admissible, then there would be an $h: [0, 1] \to [0, 1]$, (surjective) continuous, increasing, such that $h \circ \mu$ is a finitely additive probability measure. Thus

$$h \circ \mu(E) + h\mu(E^c) = h \circ \mu(F) + h \circ \mu(F^c) = 1$$

and hence

$$h \circ \mu(E^c) = h \circ \mu(F^c)$$

which contradicts the hypothesis, since h^{-1} exists. ∎

EXAMPLES. (a) $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$. Let $E = \{\omega_1, \omega_2\}$, $F = \{\omega_3, \omega_4\}$. Let $m: \mathscr{P}(\Omega) \to [0, 1]$ with

$$m(\omega_1) = m(\omega_3) = p_1 > 0$$
$$m(\omega_2) = m(\omega_4) = p_2 > 0$$
$$m(\{\omega_1, \omega_2\}) = m(\{\omega_3, \omega_4\}) = p_3 > 0$$
$$m(\{\omega_3, \omega_4, \omega_5\}) = p_4 > 0$$
$$m(\{\omega_3, \omega_6\}) = p_5 > 0$$
$$m(\{\omega_1, \omega_4\}) = p_6 > 0$$
$$m(A) = 0 \qquad \text{for all other subsets}$$

and

$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1.$$

We have $\mathrm{Bel}(E) = \mathrm{Bel}(F) = p_1 + p_2 + p_3$. $\mathrm{Bel}(E^c) = p_1 + p_2 + p_3 + p_4 + p_5$, but $\mathrm{Bel}(F^c) = p_1 + p_2 + p_3 \neq \mathrm{Bel}(E^c)$.

(b) Consider the degenerate belief function focused on A,

$$\mathrm{Bel}_A(B) = 1 \qquad \text{if} \quad A \subseteq B$$

and zero otherwise.

Let $B_1 A \neq \varnothing$ and $B_2 A = \varnothing$. Then $\mathrm{Bel}_A(B_1) = \mathrm{Bel}_A(B_2) = 0$, but

$$\mathrm{Bel}_A(B_1^c) = 0 \neq 1 = \mathrm{Bel}_A(B_2^c).$$

The hypothesis of Theorem 6.1 expresses the fact that $\mu(E^c)$ is not a function of $\mu(E)$. Thus, an uncertainty measure μ is not general admissible if there is no $\varphi: [0, 1] \to [0, 1]$ such that

$$\forall E, F \in \mathscr{A}, \qquad \mu(E^c \mid F) = \varphi(\mu(E \mid F)). \tag{*}$$

A typical example is the Max-possibility measure (which explains its inadmissibility mentioned in Section 5).

The relation (*) always holds for probability measures, but as we have just seen, (*) might fail in the case of belief functions.

The Theorem 6.1 provides a necessary condition for general admissibility: If μ is general admissible, then necessarily, (*) must hold.

The following result provides sufficient conditions for general admissibility of belief functions.

THEOREM 6.2. *Let Ω be finite and $\mathscr{A} = \mathscr{P}(\Omega)$.*

(i) *If* Bel: $\tilde{\mathscr{A}} \to [0, 1]$ (*extended to* $\tilde{\mathscr{A}}$ *as mentioned previously*) *is such that*

$$\forall E, F \in \mathscr{A}, \text{ with } \mathrm{Bel}(F) > 0, \qquad \mathrm{Bel}(E^c \mid F) = \varphi[\mathrm{Bel}(E \mid F)], \qquad (**)$$

where φ: $[0, 1] \to [0, 1]$ *is differentiable, then* Bel *is general admissible.*

(ii) *For each* $r \geqslant 1$ *and* Q: $\mathscr{A} \to [0, 1]$ *a finitely additive probability measure,* Q^r *is a belief function on* $\mathscr{A}$ *which is general admissible.*

Proof. (i) The proof of (i) follows from [5]: let $E_1, E_2, F \in \mathscr{A}$ with $\mathrm{Bel}(F) > 0$. Using the conditional extension of Bel, we have

$$\mathrm{Bel}(E_1 E_2 \mid F) = \mathrm{Bel}(E_1 \mid E_2 F)\, \mathrm{Bel}(E_2 \mid F) = \Lambda[\mathrm{Bel}(E_1 \mid E_2 F), \mathrm{Bel}(E_2 \mid F)],$$

where Λ: $[0, 1]^2 \to [0, 1]$, $\Lambda(x, y) = xy$.

We also have

$$\mathrm{Bel}(F \mid F) = 1.$$

Thus, together with (**), Bel^r is a finitely additive probability measure on $\mathscr{A}$ for some positive real r. Since $(\cdot)^r$: $[0, 1] \to [0, 1]$ is surjective, continuous and increasing, Bel is general admissible.

(ii) For $r \geqslant 1$ and Q a finitely additive probability measure in Ω (finite), Q^r is a belief function if

$$\forall A \subseteq \Omega, \qquad m_r(A) = \sum_{B \subseteq A} (-1)^{|A-B|} Q^r(B) \geqslant 0.$$

For $|A| = 0$, i.e., $A = \varnothing$, we have $m_r(\varnothing) = 0$.

For $|A| = 1$, say, $A = \{\omega_1\}$, we have

$$m_r(\{\omega_1\}) = Q^r(\{\omega_1\}) \geqslant 0.$$

For $|A| = 2$, say $A = \{\omega_1, \omega_2\}$,

$$\begin{aligned} m_r(\{\omega_1, \omega_2\}) &= Q^r(\{\omega_1, \omega_2\}) - Q^r(\{\omega_1\}) - Q^r(\{\omega_2\}) \\ &= [Q(\{\omega_1\}) + Q(\{\omega_2\})]^r - Q^r(\{\omega_1\}) \\ &\quad - Q^r(\{\omega_2\}) \geqslant 0. \end{aligned}$$

Since the function $u(\alpha) = \alpha^r + (1-\alpha)^r$ is such that $u(0) = u(1) = 1$ and $u(\cdot)$ is convex with minimum at $\alpha = \frac{1}{2}$, $u(\frac{1}{2}) \leqslant 1$, we have

$$u\left[\frac{Q(\{\omega_1\})}{Q(\{\omega_1\}) + Q(\{\omega_2\})}\right] \leqslant 1.$$

For $|A|=3$, say $A=\{\omega_1, \omega_2, \omega_2\}$, let $a=Q\{\omega_1\}$, $b=Q(\{\omega_2\})$, $c=Q(\{\omega_3\})$. Then $m_r(A)=(a+b+c)^r-(a+b)^r-(a+c)^r-(b+c)^r+a^r+b^r+c^r$. Thus $m_r(A)\geqslant 0$ if

$$(a+b+c)^r \geqslant (a+b)^r - a^r + (a+c)^r - c^r + (b+c)^r - b^r. \qquad (*)$$

Now, if $r=n\geqslant 1$ is integral, then the right hand side of $(*)$ is

$$\sum_{i=0}^{n-1}\binom{n}{i} a^i b^{n-i} + \sum_{j=0}^{n-1}\binom{n}{j} c^j a^{n-j} + \sum_{k=0}^{n-1}\binom{n}{k} b^k c^{n-k}$$
$$=\sum_S \binom{n}{i,j,k} a^i b^j c^k,$$

where $S=\{(i,j,k): i+j+k=n$ and not all the i, j, k are positive$\}$. Thus $(*)$ holds since $a, b, c\geqslant 0$ and

$$S\subseteq \{(i,j,k): i+j+k=n\}.$$

For $r\geqslant 1$, real, we rewrite $(*)$ as

$$(a+b+c)^r + a^r + b^r + c^r \geqslant (a+b)^r + (a+c)^r + (b+c)^r. \qquad (**)$$

Let $v(r)$, $w(r)$ be the left and right hand sides of $(**)$, respectively. Observe that $v(r)$ and $w(r)$ are convex functions on $[1, +\infty)$, since $0\leqslant a+b+c\leqslant 1$. Also $v(1)=w(1)=2(a+b+c)$ and

$$\lim_{r\to+\infty} v(r) = \lim_{r\to+\infty} w(r) = 0.$$

Thus $(**)$ holds since $(**)$ holds for any integer $r\geqslant 1$.

The above argument applies to the case $|A|\geqslant 4$ as well. ∎

Remarks. (i) Of course if $\mathrm{Bel}=Q^r$ where Q is a finitely additive probability measure and $r\geqslant 1$, then Bel is general admissible by Theorem 6.2 (i): $\forall E, F\in\mathscr{A}$, $\mathrm{Bel}(E^c|F)$ is a function of $\mathrm{Bel}(E|F)$, namely $\varphi(x)=(1-x)^r$.

(ii) We have seen that, in the scoring approach to admissibility of uncertainty measures, if the aggregation function is taken to be addition (as in Lindley's work), then well-known measures such as Max-possibility measures (which are consonant plausibility functions) and degenerate belief functions are not general admissible (i.e., they are incoherent in DeFinetti's sense). Although, we have shown that, in this case, there exist admissible T-possibility measures and admissible belief functions, it is useful to consider arbitrary aggregation functions in order to set up a general framework for studying the question of admissibility, i.e., a general concept of coherence. From a logical viewpoint, this is precisely the concern of *AI* researchers as well as statisticians dealing with applications of belief functions to statistical inference (e.g., [30, p. 106–107]). Section 7 will give some insights in this direction.

7. Uncertainty Games with Nonadditive Aggregation Functions

In this section, some specific nonadditive aggregation functions are considered. In Example 1 it is shown that a simple nonadditive form of aggregation leads to a corresponding uncertainty game for which uniformly admissible measures are not transformable to probabilities; this is in opposition to Lindley's games with an additive aggregation function. Example 2 presents a situation in which a nonadditive aggregation function has a general additive form. In this case, uniformly admissible measures have probability-like characterizations. In Example 3, we present some specialized cases of Example 2.

Example 1. Consider $\hat{A} = ((E_1 \mid F), (E_2 \mid F), (E_1 \cup E_2 \mid F))$, with $E_1 E_2 = \varnothing$.

Let $x = \mu(E_1 \mid F)$, $y = \mu(E_2 \mid F)$, $z = \mu(E_1 \cup E_2 \mid F)$.

We will discuss the weak local admissibility of $\hat{x} = (x, y, z)$ with respect to the game $G^*_{f,\psi,\hat{A}}$, where the game is specified as follows.

Recall that $\psi\colon \mathbb{R}_\infty \to \mathbb{R}$ is specified by the sequence $\psi_n\colon \mathbb{R}^n \to \mathbb{R}$, $n \geqslant 1$. For our purpose here, it suffices to consider $\psi_3\colon \mathbb{R}^3 \to \mathbb{R}$.

We take $\psi_3(u, v, w) = uv + w$.

For the score function f, we take $a_0 = 0$, $a_1 = 1$ and

$$f(x, 0) = x^2, \qquad f(x, 1) = (x-1)^2 \qquad \text{on } [0, 1],$$

We are going to show that there is no transform h: $[0, 1] \to [0, 1]$ such that

$$h(z) = h(x) + h(y) \qquad \text{when } (x, y, z) \text{ is WLAD.}$$

As a consequence, if μ is uniformly admissible, μ need not be transformable to a finitely additive probability measure.

With the notation of Section 3, we have

$$J = 2 \begin{bmatrix} (x-1)\,y^2 & (x-1)^2\, y & z-1 \\ x(y-1)^2 & x^2(y-1) & z-1 \\ xy^2 & x^2 y & z \end{bmatrix}.$$

To find $\hat{x} = (x, y, z) - \text{WLAD}$, we use the sufficient condition given in Theorem 3.2.2 (e).

First, $\det J = 0$ gives

$$z(x, y) = \frac{x^2 + y^2 - xy(x+y)}{2(x+y) - 3xy - 1} \qquad \text{for} \quad xy \neq 0, 1. \tag{*}$$

For $0 < x < 1$, $y = 1$, we have $z = 1$. For $\hat{x} = (x, 1, 1)$ with $0 < x < 1$, J has rank $\rho = 2$ with $R = (0, 0)$, and so $(x, 1, 1)$ is WLAD.

If $h\colon [0, 1] \to [0, 1]$ is such that $h(z) = h(x) + h(y)$ for any (x, y, z) – WLAD, then, in particular,

$$h(1) = h(x) + h(1), \qquad \forall x \in (0, 1),$$

implying that $h(x) = 0$, $\forall x \in (0, 1)$.

From (*), we see that when $x = y = \frac{1}{2}$, we have $z = 1$. For $\hat{x} = (\frac{1}{2}, \frac{1}{2}, 1)$,

$$J = 2\begin{bmatrix} -\frac{1}{8} & \frac{1}{8} & 0 \\ \frac{1}{8} & -\frac{1}{8} & 0 \\ \frac{1}{8} & \frac{1}{8} & 1 \end{bmatrix}$$

and $\rho = 2$. For

$$J = \begin{bmatrix} J_1 & J_1 C \\ RJ_1 & RJ_1 C \end{bmatrix}$$

with

$$J_1 = \begin{bmatrix} -\frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

and $RJ_1 = (\frac{1}{4}, -\frac{1}{4})$, we have $R = (-1, 0)$, so that $\hat{x} = (\frac{1}{2}, \frac{1}{2}, 1)$ is WLAD, and hence

$$h(1) = h(\tfrac{1}{2}) + h(\tfrac{1}{2}) = 0.$$

Now it can be seen that $\hat{x} = (0, 0, 0)$ is WLAD, since

$$J = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \rho = 1, \qquad \text{and} \qquad \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$$

contains one zero component. Thus $h(0) = h(0) + h(0)$, implying that $h(0) = 0$.

Therefore, $h \equiv 0$ on $[0, 1]$.

EXAMPLE 2. Consider $\hat{A}$ as in Example 1. Let f be an arbitrary score function and

$$\psi_3(u, v, w) = \varphi_0(\varphi_1(u) + \varphi_2(v) + \varphi_3(w)),$$

where $\varphi_j: \mathbb{R} \to \mathbb{R}$, $j = 0, 1, 2, 3$ are four surjective, increasing, continuously differentiable functions. This is a generalized form of the additive aggregation function. Note that $\psi_3(u, v, w) = uv + w$ in Example 1 is not of this form. Large classes of functions of several variables can be represented, or approximated, by general forms of addition of simple argument functions as above; see, e.g., the survey paper of Sprecher [29] on the work of Kolmogorov and others in considering the related 13th-problem of Hilbert.

If $\hat{x} = (x, y, z)$ is WLAD then $\det J = 0$, implying that

$$P_{\varphi_3 \circ f}(z) = P_{\varphi_1 \circ f}(x) + P_{\varphi_2 \circ f}(y),$$

where

$$P_{\varphi_i \circ f}(t) = \frac{\varphi_i'[f(t, 0)]\, f'(t, 0)}{\varphi_i'[f(t, 0)]\, f'(t, 0) - \varphi_i'[f(t, 1)]\, f'(t, 1)}.$$

EXAMPLE 3. As a special case of Example 2, one can consider uncertainty games with symmetric aggregation functions as follows.

Let $g: \mathbb{R} \to \mathbb{R}$ be a (surjective), increasing, continuously differentiable function.

For $n \geqslant 1$, define

$$\psi_{g,n}: \mathbb{R}^n \to \mathbb{R}$$

as

$$\psi_{g,n}(x_1, ..., x_n) = g^{-1}\left(\sum_{i=1}^{n} g(x_i)\right).$$

The game G_{f,ψ_g} can be identified with $G_{g \circ f, +}$, where ψ_g is specified by $\psi_{g,n}$, $n \geqslant 1$. For example, for a fixed $p \geqslant 1$, take

$$g_p(t) = t^p \qquad \text{for} \quad t \geqslant 0.$$

Note that $\mathrm{Max}(x_1, x_2, ..., x_n)$ is a limiting case of $\psi_{g_p,n}$ when $p \to +\infty$. However, $\psi = \mathrm{Max}$ yields essentially only trivial admissible uncertainty measures and hence is not a good candidate for being a viable nonadditive aggregation function.

ACKNOWLEDGMENTS

We thank D. V. Lindley and L. A. Zadeh for their kind discussion on an earlier draft of this work.

REFERENCES

1. J. ACZEL, "Lectures on Functional Equations and Their Applications," Academic Press, New York, 1966.
2. J. P. AUBIN, A Pareto minimum principle, *in* "Differential Games and Related Topics" (H. W. Kuhn and G. P. Szego, Eds.), pp. 147–175, North-Holland, Amsterdam, 1971.
3. D. BLACKWELL AND M. A. GIRSHICK, "Theory of Games and Statistical Decisions," Dover, New York, 1979.
4. L. D. BROWN, A heuristic method for determining admissibility of estimators—with applications, *Ann. Statist.* **7**, No. 5 (1979), 960–994.
5. R. T. COX, "The Algebra of Probable Inference," The Johns Hopkins Press, Baltimore, 1961.
6. B. DEFINETTI, "Theory of Probability," Vols. 1 and 2, Wiley, New York, 1974.
7. T. S. FERGUSON, "Mathematical Statistics," Academic Press, New York, 1967.
8. D. A. FREEDMAN AND R. A. PURVES, Bayes' method for bookies, *Ann. Math. Statist.* **40** (1969), 1177–1186.
9. C. GENEST AND R. J. MACKAY, Copules Archimédiennes et familles de lois bidimensionnelles dont les marges sont données, *Canad. J. Statist.* **14** (1986), 145–159.
10. C. GENEST AND R. J. MACKAY, The joy of copulas: bivariate distributions with uniform marginals, *Amer. Statist.* **40** (1986), 280–283.
11. I. R. GOODMAN AND H. T. NGUYEN, "Uncertainty Models for Knowledge-Based Systems," North-Holland, Amsterdam, 1985.
12. I. R. GOODMAN, H. T. NGUYEN, AND E. A. WALKER, "Conditional Inference and Logic for Intelligent Systems: A Theory of Measure-Free Conditioning," North-Holland, Amsterdam, 1991.
13. D. HEATH AND W. SUDDERTH, On finitely additive priors, coherence, and extended admissibility, *Ann. Statist.* **6** (1978), 333–345.
14. A. KOZEK, Towards a calculus for admissibility, *Ann. Statist.* **10**, No. 3 (1982), 825–837.
15. R. B. LATTA, Composition rules for probabilities from paired comparisons, *Ann. Statist.* **7**, No. 2 (1979), 349–371.
16. D. V. LINDLEY, Scoring rules and the inevitability of probability, *Intern. Statist. Rev.* **50** (1982), 1–26.
17. D. V. LINDLEY, The probability approach to the treatment of uncertainty in Artificial Intelligence and expert systems, *Statist. Sci.* **2**, No. 1 (1987), 17–24.
18. C. H. LING, Representation of associative functions, *Publ. Math. Debrecen* **12** (1965), 189–212.
19. A. W. MARSHALL AND I. OLKIN, Families of multivariate distributions, *J. Amer. Statist. Assoc.* **83**, No. 403 (1988), 834–841.
20. D. MCDERMOTT AND J. DOYLE, Non-monotonic logic, I, *Artificial Intelligence* **13**, No. 1 (1980), 41–72.
21. H. T. NGUYEN, On random sets and belief functions, *J. Math. Anal. Appl.* **65**, No. 3 (1978), 531–542.
22. E. REGAZZINI, DeFinetti's coherence and statistical inference, *Ann. Statist.* **15**, No. 2 (1987), 845–864.

23. A. Renyi, "Foundations of Probability," Holden–Day, San Francisco, 1970.
24. G. Schay, An algebra of conditional events, *J. Math. Anal. Appl.* **24**, No. 2 (1968), 334–344.
25. B. Schweizer and A. Sklar, "Probabilistic Metric Spaces," North-Holland, Amsterdam, 1983.
26. G. Shafer, "A Mathematical Theory of Evidence," Princeton Univ. Press, Princeton, NJ, 1976.
27. S. Smale, Global analysis and economics, I, Pareto optimum and a generalization of Morse theory, *in* "Dynamical Systems" (M. M. Peixoto, Ed.), pp. 531–544, Academic Press, New York, 1973.
28. D. J. Spiegelhalter, A statistical view of uncertainty in expert systems, *in* "Artificial Intelligence and Statistics" (W. A. Gale, Ed.), pp. 17–56, Addison–Wesley, Reading, MA, 1986.
29. D. A. Sprecher, A survey of solved and unsolved problems on super-positions of functions, *J. Approx. Theory* **6** (1972), 123–134.
30. L. A. Wasserman, "Some Applications of Belief Functions to Statistical Inference," Ph.D. thesis, Univ. of Toronto, Ontario, 1987.
31. S. Weber, $\perp$-decomposable measures and integrals for archimedean t-conorms $\perp$, *J. Math. Anal. Appl.* **101** (1984), 114–138.
32. L. A. Zadeh, The role of fuzzy logic in the management of uncertainty in expert systems, *Fuzzy Sets and Systems* **11** (1983), 199–227.
33. L. A. Zadeh, Possibility theory and soft data analysis, *in* "Mathematical Frontier of the Social and Policy Sciences" (L. Cobb and R. M. Thrall, Eds.), pp. 69–129, Westview, Boulder, CO, 1981.
34. T. L. Fine, "Theories of Probability," Academic Press, New York, 1973.