The Annals of Probability
1986, Vol. 14, No. 2, 710–723

A NOTE ON UNDOMINATED LOWER PROBABILITIES[1]

BY ADRIANOS PAPAMARCOU AND TERRENCE L. FINE

Cornell University

Interest in lower probability has largely focussed on lower envelopes and, more particularly, on belief functions. We consider those lower probabilities that do not admit of a dominating probability measure and hence are not lower envelopes. A simple and useful family of such undominated lower probabilities is constructed. We briefly explore the geometry of several important classes of lower probabilities and note that the class of undominated lower probabilities has the dimension of the set of all lower probabilities when these are modeled as vectors. While joint experiments can always be formed from given individual experiments characterized by probability measures, the existence of joint experiments is an open question as regards characterizations by lower probabilities. We constructively show the existence of joint experiments for a wide, but not exhaustive, range of characterizations of the marginal experiments. We also consider extensions of lower probabilities and show that a lower probability (including a measure) on a finite algebra can always be extended to an undominated lower probability on a larger, but still finite algebra. Finally we construct continuous undominated extensions of lower probabilities given on finite algebras.

1. Introduction. The theory of lower probability (LP) has a variety of origins and domains of application, encompassing both subjective and objective interpretations of chance, uncertainty, and indeterminacy. We find subjectivist accounts of LP in the works of Dempster (1967), Good (1962), Shafer (1976, 1982), Smith (1961), and Walley (1981). An objective basis for LP resembling the frequentist view of numerical probability has also been suggested in Walley and Fine (1982). All the above approaches introduce a pair of set functions $\underline{P}$ (lower probability) and $\overline{P}$ (upper probability) on a measurable space $(\Omega, \mathscr{A})$ that are related by

$$(\forall A \in \mathscr{A}) \qquad \underline{P}(A) + \overline{P}(A^c) = 1 \tag{1.1}$$

and satisfy the following axioms:

$$(\forall A \in \mathscr{A}) \qquad \underline{P}(A) \geq 0, \tag{1.2}$$

$$\underline{P}(\varnothing) = 0, \tag{1.3}$$

$$\underline{P}(\Omega) = 1, \tag{1.4}$$

$$(\forall A, B \in \mathscr{A} \text{ such that } A \cap B = \varnothing) \qquad \underline{P}(A) + \underline{P}(B) \leq \underline{P}(A \cup B) \quad (\textit{superadditivity} \text{ of } \underline{P}), \tag{1.5}$$

$$(\forall A, B \in \mathscr{A} \text{ such that } A \cap B = \varnothing) \qquad \overline{P}(A) + \overline{P}(B) \geq \overline{P}(A \cup B) \quad (\textit{subadditivity} \text{ of } \overline{P}). \tag{1.6}$$

Received May 1984; revised March 1985.

[1] Prepared with partial support from DOE Contract DE-AC 1-80RA5022 and NSF Grant SES-8420461.

AMS 1980 *subject classification*. 60A05.

Key words and phrases. Lower probability, interval-valued probability, belief function, lower envelope.

The intractability of the superadditivity and subadditivity axioms has favored the study of LP structures that are related to classes of probability measures; these are the so called *dominated* lower probabilities, for which there exists a probability measure μ such that

$$(\forall A \in \mathscr{A}) \qquad \underline{P}(A) \leq \mu(A).$$

Trivially, all probability measures are dominated lower probabilities. Other types of dominated LP, such as lower envelopes, 2-monotone lower probabilities, and belief functions are well suited to computation and have provided the setting for a variety of interesting results [see Levi (1980), Walley (1981) and Wolfenson and Fine (1982)]. It is clear, however, that the study of just these classes cannot bring out the full implications of the LP axioms, nor as we observe below can we hope to successfully model the full range of significant nondeterministic physical processes if we so restrict ourselves.

Our present aim is to investigate the relatively unexplored class of undominated lower probabilities. The need for this class of statistical models for certain noise-type processes has been argued by Kumar and Fine (1984) and Grize (1984). These works were motivated by the thoroughly studied example of so-called flicker or $1/f$ noise provided by the frequency fluctuations of high quality quartz crystal oscillators used in conjunction with atomic clocks [e.g., Kroupa (1983)]. The weight of experimental evidence and the beliefs of many experimentalists in that area support a mathematical model in which oscillator frequency fluctuations over time form a stationary process with bounded yet diverging time averages of frequency. The existence of such a mathematical model, however, is ruled out by conventional numerical probability theory based upon countably additive measures: in strengthening the well-known stationarity convergence theorems, Kalikow (1984) has shown that if the time averages of a stationary stochastic process have almost surely finite lim inf and lim sup, then these averages converge almost surely. If we add to that model the physically reasonable stipulation that the expectations be finite, then we come into conflict with the standard ergodic theorems.

In their search for lower probability-based models for this process, Kumar (1982) and Kumar and Fine (1984) showed that in the absence of some form of monotone continuity along sequences of observable events one could have lower probability models that agree on the cylinder sets yet assign different values to, say, divergence of time averages. Hence, one needs to postulate monotone continuity if one is to be able to infer the lower probability of such infinitary events as divergence of time averages from lower probabilities founded upon observations. They then proved that, for finite sample spaces, imposing this postulate implies that no stationary dominated lower probability that is monotonely continuous along the cylinder sets can assign positive probability to the event of bounded and diverging time averages. Hence models of the type needed to deal with the empirically significant class of flicker-like processes, if they exist, must be of the little-studied undominated type. Our interest in this class was further stimulated by the work of Grize (1984), Grize and Fine (1986), which established that the class of undominated, stationary, monotonely continuous along cylinder sets lower probabilities supporting divergence was nonempty.

In Section 2 we construct a class of undominated lower probabilities on finite spaces. Certain geometrical properties of LP are developed in Section 3, where it is seen that undominated lower probabilities are a sizable subclass of the existing LP structures. In Section 4 we examine the existence of bivariate lower probabilities having two given lower probabilities (dominated or undominated) as marginals. Although our account is not complete, we obtain results which allow us in Section 5 to view undominated lower probabilities as possible extensions of dominated lower probabilities or probability measures from a finite algebra $\mathscr{A}$ to the power set of an infinite sample space Ω.

REMARK. By virtue of (1.1), the subadditivity axiom (1.5) can be written in terms of $\underline{P}$ only:

$$(1.7)\quad (\forall A, B \in \mathscr{A}) \qquad \underline{P}(A) + \underline{P}(B) \leq 1 + \underline{P}(A \cap B) \quad (\textit{conjugacy} \text{ of } \underline{P}).$$

Thus we only need to consider $\underline{P}$ in order to verify the validity of a proposed LP structure; we shall also drop the lower bar in $\underline{P}$.

2. Examples of undominated lower probabilities. We begin by giving a criterion for a set function on a finite space to be undominated.

PROPOSITION 2.1. *A set function P on $\Omega = \{\omega_1, \dots, \omega_n\}$ is undominated if and only if there exists a collection of (not necessarily distinct) sets $A_1, \dots, A_m \subset \Omega$ such that*

$$(\forall\, \omega \in \Omega) \qquad \sum_{i=1}^{m} \left[I_{A_i}(\omega) - P(A_i) \right] < 0.$$

To prove this proposition, we represent $I_A(\omega) - P(A)$, $A \subset \Omega$, by a vector in R^n and consider the convex hull $\mathscr{L}$ of the resulting 2^n vectors. Application of the separating hyperplane theorem to $\mathscr{L}$ and the closed positive orthant yields the final result [for a detailed proof, see Walley (1981); note that similar geometrical arguments, dating back to Scott (1964), have been used to derive conditions under which comparative probability structures admit numerical probability representations, as in Walley and Fine (1979)].

The following construction is based on the above criterion and generalizes the only example of undominated LP in related literature [Walley (1981)]. Let $\mathscr{C}$ be a class of subsets of Ω such that no two sets in $\mathscr{C}$ are mutually disjoint. Take $\gamma \in [0,1]$ and define the set function Q_γ on Ω as follows:

$$(2.1)\qquad Q_\gamma(A) = \begin{cases} 0 & \text{if } A \text{ does not contain any set in } \mathscr{C}, \\ \gamma & \text{if } A \text{ contains a set in } \mathscr{C} \text{ and } A \neq \Omega, \\ 1 & \text{if } A = \Omega. \end{cases}$$

Clearly Q_γ is nonnegative, $Q_\gamma(\varnothing) = 0$ and $Q_\gamma(\Omega) = 1$. By definition $\mathscr{C}$ contains no two mutually disjoint sets, so that

$$A \cap B = \varnothing \Rightarrow Q_\gamma(A) \cdot Q_\gamma(B) = 0$$

and the superadditivity condition (1.5) always holds. To check the conjugacy condition (1.7), note first that if either A or B is Ω, then

$$Q_\gamma(A) + Q_\gamma(B) = 1 + Q_\gamma(A \cap B)$$

holds trivially. Otherwise,

$$Q_\gamma(A) + Q_\gamma(B) \leq 2\gamma$$

and thus Q_γ is a lower probability provided that $\gamma \leq \frac{1}{2}$. If $\mathscr{C}$ has the additional property that

$$(\forall\, \omega \in \Omega) \qquad \sum_{C \in \mathscr{C}} \left[I_C(\omega) - Q_\gamma(C)\right] < 0,$$

then the criterion of Proposition 2.1 is satisfied and Q_γ is an undominated lower probability. We shall now propose a candidate for $\mathscr{C}$.

We take an integer $k \leq n/2$ and let l be the least integer greater than $n/2k$. We define $\mathscr{C} = \{C_1, \ldots, C_n\}$ by

$$C_1 = \{\omega_1, \omega_2, \ldots, \omega_k, \omega_{2k}, \ldots, \omega_{lk}\}$$

and

$$C_i = \tau(C_{i-1}) \qquad (i = 2, \ldots, n),$$

where τ is the circular shift operator on the subsets of Ω [e.g., if $n = 6$ and $A = \{\omega_2, \omega_6\}$, then $\tau(A) = \{\omega_3, \omega_1\}$].

Proposition 2.2. *No two members of $\mathscr{C}$ are disjoint.*

Proof. By construction we have for $1 \leq i, j \leq n$

$$\tau(C_i \cap C_j) = \tau(C_i) \cap \tau(C_j)$$

and consequently

$$\tau^m(C_i \cap C_j) = \tau^m(C_i) \cap \tau^m(C_j) \quad (m \text{ arbitrary}).$$

Thus it suffices to show that C_1 intersects every set in $\mathscr{C}$. Observe that if we arrange the elements of Ω in a circle, then each C_i will consist of k consecutive elements (the *base*) and $l - 1$ elements spaced k positions apart. Hence, for $1 \leq i \leq lk$, the base of each C_i intersects C_1.

If $lk + 1 \leq i \leq n$, then $1 \leq n - i + 1 \leq n - lk$. By definition of l, $lk > n/2$ and therefore $1 \leq n - i + 1 < lk$. Since we can write C_1 as $\tau^{n-i+1}(C_i)$, the base of C_1 intersects each C_i for $lk + 1 \leq i \leq n$. This completes our proof. □

Proposition 2.3. $\sum_{C \in \mathscr{C}} I_C(\omega) = k + l - 1.$

Proof. Take $1 \leq i < n$. Then $(\forall\, A \subset \Omega)\omega_i \in A \Leftrightarrow \omega_{i+1} \in \tau(A)$. Therefore $I_A(\omega_i) = I_{\tau(A)}(\omega_{i+1})$ and

$$\sum_{C \in \mathscr{C}} I_C(\omega_i) = \sum_{C \in \mathscr{C}} I_{\tau(C)}(\omega_{i+1}) = \sum_{C \in \mathscr{C}} I_C(\omega_{i+1}).$$

Thus $\sum_{C \in \mathscr{C}} I_C(\omega)$ is a constant function on Ω, equal to α. Then

$$n\alpha = \sum_{i=1}^{n} \sum_{C \in \mathscr{C}} I_C(\omega_i) = \sum_{C \in \mathscr{C}} \sum_{i=1}^{n} I_C(\omega_i) = n\|C_1\|,$$

whence we conclude that $\alpha = \|C_1\| = k + l - 1$. □

We now return to Q_γ of (2.1). Q_γ is a lower probability for $0 \le \gamma \le \frac{1}{2}$. If $(k + l - 1)/n < \gamma \le \frac{1}{2}$, then

$$\sum_{C \in \mathscr{C}} \left[I_C(\omega) - Q_\gamma(c) \right] = k + l - 1 - n\gamma < 0$$

and consequently Q_γ is an undominated lower probability.

If $0 \le \gamma \le \min\{(k + l - 1)/n, \frac{1}{2}\}$, then Q_γ is dominated by the uniform measure on Ω. It is therefore clear that our construction can yield an undominated lower probability on a space of n elements only if

$$w(n) = \min_k \frac{k + l - 1}{n} < \frac{1}{2}.$$

We shall see for which n this is possible.

(i) *$n \le 3$.* For $n = 1$ every lower probability is a measure and for $n = 2$ every lower probability is dominated. Also, if P is a lower probability on $\{\omega_1, \omega_2, \omega_3\}$, then one can easily show that the measure μ defined by

$$\mu(\{\omega_1\}) = 1 - P(\{\omega_2, \omega_3\}),$$
$$\mu(\{\omega_2\}) = P(\{\omega_2\}),$$
$$\mu(\{\omega_3\}) = P(\{\omega_2, \omega_3\}) - P(\{\omega_2\})$$

dominates P. Hence undominated lower probabilities do not exist for $n \le 3$.

(ii) *$4 \le n \le 6$.* It can be easily verified that $w(n) \ge \frac{1}{2}$. Our construction is inapplicable. Grize (1984, page 28) has shown that lower probabilities on four elements are dominated, but the cases of five or six elements are as yet unresolved.

(iii) *$n = 7$.* $w(7) = \frac{3}{7}$ achieved at $k = 2$. Thus an undominated Q_γ can be constructed.

(iv) *$n = 8$.* $w(8) = \frac{1}{2}$ achieved at $k = 2$ or 3. Our construction is again unsuitable, but we can obtain an undominated example by regarding two elements as a single atom and applying the seven-atom construction.

(v) *$n \ge 9$.* By definition of l, $l - 1 \le n/2k$ and therefore

$$w(n) = \min_k \frac{k + l - 1}{n} \le \min_k \left[\frac{k}{n} + \frac{1}{2k} \right].$$

For $n \ge 9$ and $k = 2$

$$w(n) \le \tfrac{2}{9} + \tfrac{1}{4} < \tfrac{1}{2}$$

and our construction can give distinct examples of undominated lower probabilities. It is interesting to note that despite the absence of dominating measures, none of these examples assigns a lower probability greater than $\frac{1}{2}$ to proper subsets of Ω. The following proposition develops this fact and will be of use in Section 4.

PROPOSITION 2.4. *For any $\varepsilon > 0$ there exists on some finite space Ω an undominated lower probability P such that*

$$\max_{\substack{A \subset \Omega \\ A \neq \Omega}} P(A) \leq \varepsilon.$$

PROOF. It suffices to show that $w(n)$ can be made arbitrarily small. As we saw above,

$$w(n) \leq \frac{k}{n} + \frac{1}{2k}.$$

Take $k = \lceil \sqrt{n/2} \rceil$. Then

$$w(n) \leq \frac{\sqrt{n/2} + 1}{n} + \frac{1}{\sqrt{2n}} = \sqrt{\frac{2}{n}} + \frac{1}{n}$$

and $w(n)$ can be made arbitrarily small. □

3. The geometry of lower probability. We can gain further insight into the class of undominated lower probabilities by considering certain geometrical properties that follow from the isomorphism between the set functions on a space of n elements and the 2^n-dimensional Euclidean space. We can define the following sets in R^{2^n}:

$\mathscr{P}$: The class of all LP's on $\Omega = \{\omega_1, \dots, \omega_n\}$;
$\mathscr{M}$: the class of all probability measures on Ω;
$\mathscr{D}$: the class of all dominated LP's on Ω;
$\mathscr{U}$: the class of all undominated LP's on Ω.

The sets $\mathscr{P}$ and $\mathscr{M}$ are defined by finite systems of linear inequalities and equalities and are therefore closed and convex polyhedra. For each vertex μ of $\mathscr{M}$, let $\mathscr{F}_\mu$ be the set of vectors P such that $\mu \geq P$. Clearly $\mathscr{F}_\mu$ is closed and so is $\cup \mathscr{F}_\mu$, where the union is taken over the finite vertex set of $\mathscr{M}$. $\mathscr{D}$ is the intersection of $\mathscr{P}$ and the convex hull of $\cup \mathscr{F}_\mu$; it is thus both closed and convex.

The set function P_B defined by

$$P_B(A) = \begin{cases} 1 & \text{if } B \subset A,\ B \text{ fixed, nonempty} \\ 0 & \text{if } B \not\subset A \end{cases}$$

is a lower probability dominated by the measure that gives unit mass to a sample point in B. By varying B over the $2^n - 2$ nonempty proper subsets of Ω, we obtain $2^n - 2$ linearly independent vectors P_B. Therefore $\mathscr{D}$ and $\mathscr{P}$ are at least of dimension $2^n - 2$ and since $P(\varnothing)$ and $P(\Omega)$ are fixed, they are exactly of dimension $2^n - 2$. $\mathscr{M}$ is of dimension $n - 1$.

If $\mathscr{U} = \mathscr{P} - \mathscr{D}$ is nonempty (e.g., for $n \geq 7$) then it is also of dimension $2^n - 2$, since $\mathscr{D}$ and $\mathscr{P}$ are convex and of the same dimension. Thus in a sense $\mathscr{U}$ is as large as $\mathscr{D}$; it is not, however, always convex, as the following example shows.

EXAMPLE 3.1. Let $n = 7$ and consider the standard construction Q_γ based on $\mathscr{C} = \{\tau^i(\{\omega_1, \omega_2, \omega_4\}) | i = 0, \dots, n-1\}$. Now replace $\mathscr{C}$ by

$$\mathscr{C}' = \left\{\tau^i(\{\omega_4, \omega_6, \omega_7\}) \middle| i = 0, \dots, n-1\right\}$$

and define Q'_γ in a similar way. Then

$$P = \tfrac{1}{2}\left(Q_{1/2} + Q'_{1/2}\right)$$

is a lower probability by convexity of $\mathscr{P}$. Furthermore,

(i) if $\|A\| < 3$, then $Q_{1/2}(A) = Q'_{1/2}(A) = P(A) = 0$,
(ii) if $\|A\| = 3$, then since $\mathscr{C}$ and $\mathscr{C}'$ have no set in common, $P(A) \leq \frac{1}{4}$,
(iii) if $3 < \|A\| < 7$, then $Q_{1/2}(A) \leq \frac{1}{2}$, $Q'_{1/2}(A) \leq \frac{1}{2}$, and thus $P(A) \leq \frac{1}{2}$.

Therefore P is dominated by the uniform measure on Ω. □

Finally we should note that $\mathscr{M}$ is on the relative boundary (with respect to the affine hull) of $\mathscr{P}$, since every probability measure satisfies the superadditivity relationships (1.5) with equality. Furthermore, $\mathscr{M} \subset \mathscr{D}$ and thus $\mathscr{M}$ is on the relative boundary of $\mathscr{D}$. We now show that $\mathscr{M}$ is also on the relative boundary of $\mathscr{U}$.

PROPOSITION 3.2. *Let P be an undominated* LP, *μ a probability measure. Then $Q = \lambda P + (1 - \lambda)\mu$ is an undominated* LP *if* $\lambda \in (0,1)$.

PROOF. Q is a lower probability by convexity of $\mathscr{P}$. If Q is dominated by a probability measure ν, then

$$\lambda P + (1 - \lambda)\mu \leq \nu$$

and thus

$$P \leq \frac{\nu - (1 - \lambda)\mu}{\lambda} = R.$$

Since R is a linear combination of two measures, it is additive. Also $R(\phi) = 0$, $R(\Omega) = 1$, and $R \geq 0$ by nonnegativity of P. Therefore R is a probability measure and P is dominated. This contradicts our assumption and thus Q is undominated. □

The above result allows us to construct unlimited examples of undominated lower probabilities by taking convex combinations of probability measures with Q_γ.

4. Joint experiments. In numerical probability, joint experiments between two arbitrary experiments always exist. In lower probability, however, the absence of additivity significantly complicates the formation of bivariate joint lower probability. Take for example the simple case of two finite spaces $X = \{x_1, \dots, x_m\}$, $Y = \{y_1, \dots, y_n\}$ and two lower probabilities P (on X) and R (on Y). We seek a lower probability Q on $X \times Y$ which satisfies the marginal constraints

$$(\forall A \subset X) \qquad Q(A \times Y) = P(A),$$
$$(\forall B \subset Y) \qquad Q(X \times B) = R(B).$$

These constraints, together with the LP axioms for Q, form a system of linear equations and inequalities in $\mathbb{R}^{2^{nm}}$. The solvability of this system can be reduced, by Theorems of the Alternative, to the validity of relationships such as

$$\sum_{A \subset X} \alpha(A)P(A) + \sum_{B \subset Y} \beta(B)R(B) \leq \eta,$$

but such an approach has yielded little. We might also attempt the construction of a joint experiment Q in a fashion similar to the development of product measures:

(i) Specification of a suitable set function Q' on the class $\mathscr{R}$ of rectangles in $X \times Y$. As the cylinder sets $A \times Y$ and $X \times B$ are rectangles and $\mathscr{R}$ is closed under intersections, we expect Q' to be consistent with the marginal constraints and satisfy the conjugacy condition. Examples of such Q' include

$$Q_1'(A \times B) = P(A) \cdot R(B),$$
$$Q_2'(A \times B) = \max\{0, P(A) + R(B) - 1\}.$$

(ii) Extension of Q' to a lower probability Q on $X \times Y$. In light of the superadditivity constraint, the set function Q defined by

$$(\forall G \subset X \times Y) \qquad Q(G) = \max\left\{ \sum_{i=1}^{M} Q'(A_i \times B_i) \colon \{A_i \times B_i\} \text{ a partition of } G \right\}$$

becomes a natural candidate for the joint order.

Yet this extension is problematic. On the one hand it may not be a true extension of Q' [i.e., $Q(A \times B) > Q'(A \times B)$] and thus violate the marginal constraints; on the other, we have seen examples in which it violates the conjugacy condition when applied to Q_1' and Q_2' defined above [see Papamarcou (1983)]. In view of such difficulties we are unable to treat the general problem and proceed instead to show the existence of joint experiments in the following cases:

(a) When at least one of the marginals is a lower envelope (see below).
(b) When the sum of the maximum nonunity values taken by the marginals is less than or equal to unity.

DEFINITION 4.1. Let P be a dominated LP on $(\Omega, \mathscr{A})$ and $\mathscr{M}_P$ the collection of probability measures dominating P. Then P is a *lower envelope* if

$$(\forall A \in \mathscr{A}) \qquad P(A) = \inf_{\mu \in \mathscr{M}_P} \mu(A).$$

If $Z(\omega)$ is a bounded random variable on $(\Omega, \mathscr{A})$, then we define $\underline{E}_P Z$, the *lower expectation* of Z with respect to P, by

$$\underline{E}_P Z = \inf_{\mu \in \mathscr{M}_P} \int Z \, d\mu .$$

PROPOSITION 4.2. *Let P be a lower envelope on X and R be a lower probability on Y, where X and Y are finite. Then the set function Q on $X \times Y$ defined by*

$$Q(G) = \underline{E}_P R(G_x),$$

where G_x is the section of G at x, is a lower probability with P and R as marginals. Furthermore, Q is dominated if and only if R is dominated.

PROOF. Clearly $(\forall\, G \subset X \times Y)\ Q(G) \geq 0$

$$Q(\varnothing) = \underline{E}_P(0) = 0,$$
$$Q(X \times Y) = \underline{E}_P(1) = 1.$$

If $A \subset X$, $B \subset Y$, then

$$Q(A \times Y) = \underline{E}_P I_A = \inf_{\mu \in \mathscr{M}_P} \int I_A \, d\mu = \inf_{\mu \in \mathscr{M}_P} \mu(A) = P(A),$$
$$Q(X \times B) = \underline{E}_P R(B) = R(B),$$

and thus Q agrees with P and R on the cylinder sets. It remains to show that Q satisfies the superadditivity and conjugacy conditions. If G, $H \subset X \times Y$, then

$$\begin{aligned} Q(G) + Q(H) &= \underline{E}_P R(G_x) + \underline{E}_P R(H_x) \\ &= \inf_{\mu \in \mathscr{M}_P} \int R(G_x)\, d\mu + \inf_{\nu \in \mathscr{M}_P} \int R(H_x)\, d\nu \\ &\leq \inf_{\mu \in \mathscr{M}_P} \int [R(G_x) + R(H_x)]\, d\mu . \end{aligned}$$

If G and H are disjoint, then $(\forall\, x \in X) G_x \cap H_x = \varnothing$ and thus by superadditivity of R

$$(\forall\, x \in X) \qquad R(G_x) + R(H_x) \leq R(G_x \cup H_x).$$

Hence

$$Q(G) + Q(H) \leq \inf_{\mu \in \mathscr{M}_P} \int R(G_x \cup H_x)\, d\mu = Q(G \cup H).$$

If G and H are not disjoint, then by conjugacy of R

$$(\forall\, x \in X) \qquad R(G_x) + R(H_x) \leq 1 + R(G_x \cap H_x).$$

Hence

$$\begin{aligned} Q(G) + Q(H) &\leq \inf_{\mu \in \mathscr{M}_P} \int [1 + R(G_x \cap H_x)]\, d\mu \\ &= 1 + Q(G \cap H). \end{aligned}$$

Therefore Q is a lower probability on $X \times Y$ with P and R as marginals. Now assume that R is dominated. Then if $\mu \geq P$, $\nu \geq R$, we have

$$(\forall\, G \subset X \times Y) \qquad Q(G) = \underline{E}_P R(G_x) \leq \int R(G_x)\, d\mu$$

$$\leq \int \nu(G_x)\, d\mu = (\mu \times \nu)(G).$$

Therefore Q is dominated. If, on the other hand, R is undominated and Q were dominated by some measure π on $X \times Y$, then the restriction of π on the subsets of Y would dominate R, contradicting our assumption. □

The above result demonstrates the usefulness of measure-theoretic concepts in LP. Of different flavor is the proof of the following proposition, which applies to both dominated and undominated lower probabilities.

PROPOSITION 4.3. *Let P be a lower probability on X and R a lower probability on Y, where X and Y are finite. Then if*

$$\max_{\substack{A \subset X \\ P(A) \neq 1}} P(A) + \max_{\substack{B \subset Y \\ R(B) \neq 1}} R(B) \leq 1,$$

there exists a lower probability Q on $X \times Y$ with P and R as marginals.

PROOF. Let Γ be the intersection of all subsets A of X with $P(A) = 1$. Then by conjugacy $P(\Gamma) = 1$ and for $E \subset X$ we have $P(E) = P(\Gamma \cap E)$. Similarly let Δ be the intersection of all subsets B of Y with $Q(B) = 1$. Consider the product space $\Gamma \times \Delta$ and define T on the class $\mathscr{C}$ of cylinder sets ($\mathscr{C} = \{A \times \Delta, \Gamma \times B$: $A \subset \Gamma, B \subset \Delta\}$) by

$$T(A \times \Delta) = P(A),$$

$$T(\Gamma \times B) = R(B).$$

We now define the set function S on $\Gamma \times \Delta$ by

$$S(G) = \max_{\substack{C \in \mathscr{C} \\ C \subset G}} T(C).$$

Clearly S is nonnegative, $S(\varnothing) = 0$, $S(\Gamma \times \Delta) = 1$. For the sets G and H, let

$$S(G) = T(C_1),$$

$$S(H) = T(C_2),$$

where C_1 and C_2 are cylinder sets. We shall write $C_1 \| C_2$ if the bases of C_1 and C_2 lie on the same coordinate axis.

If G and H are disjoint, then necessarily $C_1 \| C_2$ and superadditivity of the marginals gives

$$S(G) + S(H) = T(C_1) + T(C_2) \leq T(C_1 \cup C_2) \leq S(G \cup H).$$

Thus S is superadditive. To establish conjugacy, it suffices to take G and H as proper subsets of $\Gamma \times \Delta$, so that $T(C_1) < 1, T(C_2) < 1$. If $C_1 \| C_2$, we have

$$S(G) + S(H) = T(C_1) + T(C_2) \le 1 + T(C_1 \cap C_2) \le 1 + S(G \cap H)$$

by conjugacy of the marginals. If not $C_1 \| C_2$, then

$$S(G) + S(H) = T(C_1) + T(C_2) \le \max_{\substack{A \subset X \\ P(A) \neq 1}} P(A) + \max_{\substack{B \subset Y \\ R(B) \neq 1}} R(B) \le 1$$

by hypothesis. Therefore S is a lower probability on $\Gamma \times \Delta$. We propose to extend S to a lower probability Q on $X \times Y$ according to the specification

$$Q(G) = S(G \cap (\Gamma \times \Delta)).$$

Clearly

$$Q(X \times Y) = S(\Gamma \times \Delta) = 1,$$
$$Q(\varnothing) = S(\varnothing) = 0$$

and Q agrees with the marginals:

$$Q(A \times Y) = S((A \cap \Gamma) \times \Delta) = P(A \cap \Gamma) = P(A),$$
$$Q(X \times B) = S(\Gamma \times (B \cap \Delta)) = R(B \cap \Delta) = R(B).$$

Superadditivity and conjugacy of Q follow from that of S, since

$$(G \cup H) \cap (\Gamma \times \Delta) = [G \cap (\Gamma \times \Delta)] \cup [H \cap (\Gamma \times \Delta)],$$
$$(G \cap H) \cap (\Gamma \times \Delta) = [G \cap (\Gamma \times \Delta)] \cap [H \cap (\Gamma \times \Delta)].$$

Hence Q is a bivariate LP with marginals P and R. □

It is therefore possible to combine an undominated LP with a dominated one when the condition of Proposition 4.3 is satisfied. Joint experiments also exist for certain pairs of undominated marginals; take for example the constructions of Section 2, where

$$\max_{\substack{A \subset \Omega \\ Q_\gamma \neq 1}} Q_\gamma(A) \le \gamma \quad \text{and} \quad \gamma \le \tfrac{1}{2}.$$

5. Extensions and refinements. Consider now a lower probability P on a finite space X. By Proposition 2.4, there exists on a finite space Y of size $n = n(P)$ an undominated lower probability R such that

$$\max_{\substack{B \subset Y \\ B \neq Y}} R(B) \le 1 - \max_{\substack{A \subset X \\ P(A) \neq 1}} P(A).$$

Thus an undominated joint experiment Q exists between P and R. Similarly, if Ω is an infinite space and P is defined on a finite algebra $\mathscr{A}$ generated by a partition of Ω into infinite sets $E_1, \dots, E_m$, we can

(i) partition each E_i into nonempty sets $E_{i1}, \dots, E_{in}$;
(ii) define, on the algebra $\mathscr{B}$ generated by the partition $\{\bigcup_{1 \le i \le m} E_{ij}, 1 \le j \le n\}$, an undominated lower probability R such that

$$\max_{\substack{B \in \mathscr{B} \\ B \neq \Omega}} R(B) \le 1 - \max_{\substack{A \in \mathscr{A} \\ P(A) \neq 1}} P(A);$$

(iii) construct an undominated joint experiment Q on $\mathscr{A}' = \mathscr{A} \times \mathscr{B}$ with P and R as marginals.

We thus obtain

PROPOSITION 5.1. *Consider an infinite space Ω and a finite partition of Ω into infinite sets: $\mathscr{E} = \{E_1, \dots, E_m\}$. Let P be a lower probability defined on $\mathscr{A}$, the algebra generated by $\mathscr{E}$. Then there exists on some finite algebra $\mathscr{A}' \supset \mathscr{A}$ of subsets of Ω an undominated lower probability which agrees with P on $\mathscr{A}$.*

In statistical terms, $\mathscr{A}$ may represent a finite experiment and P its empirically suggested model (numerical or lower probabilistic). It is then possible that further data leading to a refinement of $\mathscr{A}$ will point to an undominated lower probability as the correct model. As an example, let $\mathscr{A}$ be the σ-algebra generated by the simple random variables $Z_1, \dots, Z_n$ and P be the suggested model for these N observations. If further observations are made, it may be possible to attribute $Z_1, \dots, Z_M (M > N)$ to an undominated lower probability Q which agrees with P on the first N observations. Assuming that the entire sequence $\{Z_i\}_{i=1}^{\infty}$ admits a LP model, we must then conclude that this overall model is undominated.

In practice, it may be quite difficult to discriminate between a dominated and an undominated LP model. However, recent works by Kumar (1982), Kumar and Fine (1984), and Grize (1984) have pointed out the following.

(i) The two types of model can have different implications for the limiting behavior of certain sequences.
(ii) Estimation based on limiting behavior is meaningful only with LP models that are monotonely continuous along suitable sequences of sets (such as cylinder sets).

It thus seems desirable to carry the extension in Proposition 5.1 further to a lower probability R that is monotonely continuous on the σ-algebra generated by the cylinder sets. In fact, this is possible on the entire power set of Ω.

LEMMA 5.2. *Let P be a lower probability on a finite algebra $\mathscr{A}$ of subsets of Ω. Then there exists on 2^{Ω} a monotonely continuous lower probability R that agrees with P on $\mathscr{A}$. Furthermore, R is dominated if and only if P is dominated.*

PROOF. Let $\mathscr{A}$ be generated by the partition $\mathscr{E} = \{E_1, \dots, E_n\}$. We choose a point x_i in each E_i and define the function ψ: $2^{\Omega} \to \mathscr{A}$ by

$$\psi(A) = \bigcup_{\{i:\, x_i \in A\}} E_i.$$

We also define the set function R on 2^{Ω} by

$$(\forall A \subset \Omega) \qquad R(A) = P(\psi(A)).$$

It is easy to show that R is a lower probability on 2^{Ω} by observing that

$\psi(\varnothing) = \varnothing$, $\psi(\Omega) = \Omega$, $\psi(A \cup B) = \psi(A) \cup \psi(B)$, and $\psi(A \cap B) = \psi(A) \cap \psi(B)$. To show monotone continuity, consider first an increasing sequence $\{A_j\}$ of subsets of Ω converging to A. Let $\psi(A) = E_{\iota_1} \cup \cdots \cup E_{\iota_k}$. Then clearly

$$x_{\iota_r} \in A \Rightarrow (\exists\, M_r)(\forall\, j \geq M_r) x_{\iota_r} \in A_j$$

and

$$(\forall\, j \geq j_0 = \max\{M_1, \ldots, M_k\}) A_j \supset \{x_{\iota_1}, \ldots, x_{\iota_k}\}.$$

Consequently $(\forall\, j \geq j_0)\psi(A_j) = \psi(A)$ and $R(A_j) = R(A)$. Hence R is continuous from below on 2^Ω. Continuity from above is shown similarly.

Finally note that if P is undominated, so is R. If, on the other hand, P is dominated by a measure μ on $\mathscr{A}$, then for $(A \subset \Omega)$

$$\pi(A) = \mu(\psi(A))$$

defines a discrete measure on 2^Ω, concentrated at $x_1, \ldots, x_n$ and dominating R. Therefore R is dominated if and only if P is. □

PROPOSITION 5.3. *Let Ω be an infinite space and $\mathscr{A}$ a finite algebra of subsets of Ω in which every nonempty set is infinite. Then any lower probability P defined on $\mathscr{A}$ has an undominated, monotonely continuous extension to the power set of Ω.*

This follows immediately from Proposition 5.1 and Lemma 5.2. We conclude that empirical models based on probability measures or dominated lower probabilities always admit refinements in terms of undominated lower probabilities with desirable continuity properties.

REFERENCES

DEMPSTER, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Ann. Math. Statist.* **38** 325–339.

GOOD, I. J. (1962). Subjective probability as the measure of a non-measurable set. In *Logic, Methodology and Philosophy of Science* (E. Nagel, P. Suppes and A. Tarski, eds.) 319–329. Stanford Univ. Press.

GRIZE, Y. L. (1984). Towards a stationary continuous lower probability-based model for flicker noise. Ph.D. thesis, Cornell Univ., Ithaca, New York.

GRIZE, Y. L. and FINE, T. L. (1986). Continuous lower probability-based models for stationary processes with bounded and divergent time averages. *Ann. Prob.* To appear.

KALIKOW, S. (1984). Private communication.

KROUPA, V. (1983). *Frequency Stability: Fundamentals and Measurement.* IEEE, New York.

KUMAR, A. and FINE, T. (1984). Stationary lower probabilities and unstable averages. *Z. Wahrsch. verw Gebiete.* **69** 1–17.

KUMAR, A. (1982). Lower probabilities on infinite spaces and instability of stationary sequences. Ph.D. thesis, Cornell Univ., Ithaca, New York.

LEVI, I. (1980). *The Enterprise of Knowledge.* MIT Press.

PAPAMARCOU, A. (1983). Some results on undominated lower probabilities. M.S. thesis, Cornell Univ., Ithaca, New York.

SCOTT, D. (1964). Measurement structures and linear inequalities. *J. Math. Psych.* **1** 233–247.

SHAFER, G. (1976). *A Mathematical Theory of Evidence.* Princeton Univ. Press.

SHAFER, G. (1982). Belief functions and parametric models. *J. Roy. Statist. Soc. Ser. B* **44** 322–352.

SMITH, C. A. B. (1961). Consistency in statistical inference and decision. *J. Roy. Statist. Soc. Ser. B* **23** 1–25.

WALLEY, P. (1981). Coherent lower (and upper) probabilities. Technical Report, Dept. Statist., University of Warwick, England.

WALLEY, P. and FINE, T. L. (1979). Varieties of modal (classificatory) and comparative probability. *Synthese* **41** 321–374.

WALLEY, P. and FINE, T. L. (1982). Towards a frequentist theory of upper and lower probability. *Ann. Statist.* **10** 741–761.

WOLFENSON, M. and FINE, T. L. (1982). Bayes-like decision making with upper and lower probabilities. *J. Amer. Statist. Assoc.* **77** 80–88.

SCHOOL OF ELECTRICAL ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853