

SYMMETRIC UPPER PROBABILITIES

BY LARRY WASSERMAN¹ AND JOSEPH B. KADANE²

Carnegie Mellon University

As a first step toward developing statistical models based on upper and lower probabilities, we study upper probabilities and upper expectations on the unit interval that are symmetric, by which we mean invariant with respect to equimeasurability. These upper probabilities are generalizations of uniform probability measures. We give some characterizations of these upper probabilities. Specifically, we show that symmetry of the upper expectation functional is equivalent to the underlying set of densities being closed under majorization. We also show that a function is the upper distribution for a symmetric upper probability if and only if its lower graph is star-shaped with respect to the origin and to the point (1, 1). We derive inner and outer approximations to symmetric classes of probabilities based on the upper probability. The class of symmetric upper expectations that are completely determined by their values on the indicator functions is characterized. We provide a geometric characterization of a hierarchy of upper probabilities including Fine's generalized upper probabilities and 2-alternating Choquet capacities. In particular, we establish a 1–1 correspondence between symmetric, 2-alternating capacities and nonincreasing density functions. We prove that undominated generalized upper probabilities do not exist in the symmetric case. Examples from robust statistics are considered. An example is given that shows that symmetry of upper probabilities does not imply symmetry of upper expectations. A corollary is that symmetry of the Choquet integral does not imply symmetry of the upper expectation functional.

1. Introduction. Walley (1981, 1991) developed a theory of statistical inference based on upper and lower probabilities. If M is a set of probability measures on a σ -algebra \mathcal{B} , then $\bar{P}(A) = \sup_{P \in M} P(A)$ is called an upper probability and $\bar{E}X = \sup_{P \in M} \int X dP$ is called an upper expectation. [Fine (1988) has a more general definition; see Section 6.] Upper and lower probabilities arise in robust Bayesian inference [Berger (1990), DeRobertis and Hartigan (1981), Lavine (1988), Wasserman (1990), Wasserman and Kadane (1990b) and Wolfeson and Fine (1982)], robust classical inference [Huber (1973), Huber and Strassen (1973), Buja (1984, 1985, 1986), Bednarski (1981, 1982) and Rieder (1977)] and economic theory [Schmeidler (1989)], as well as in other work on coherent inference [Smith (1961) and Williams (1976)] and in the foundations of probability [Koopman (1940a, b), Good (1962), Fine (1988),

Received January 1991; revised February 1992.

¹Research supported by NSF Grant DMS-90-05858.

²Research supported by NSF Grants DMS-90-05858, DMS-87-05646, DMS-87-01770 and SES-89-00025 and by ONR contract N00014-89-J-1851.

AMS 1980 subject classifications. Primary 60A05; secondary 62A99.

Key words and phrases. Choquet capacities, doubly stochastic operators, equimeasurability, lower probabilities, majorization, 2-alternating capacities.

Grize and Fine (1987), Kumar and Fine (1985), Papamarcou and Fine (1986), Walley and Fine (1982), Puri and Ralescu (1983) and Suppes and Zanotti (1989)]. They also arise in the theory of belief functions [Dempster (1967, 1968) and Shafer (1976, 1979)].

For further development of statistical inference based on upper and lower probabilities, it is necessary to construct useful parametric models. The simplest probability measure is the uniform probability measure. The question then arises whether there exists a natural generalization of a uniform probability in the theory of upper probabilities. If so, this can be used as a basic building block for a theory of modeling using upper and lower probabilities. More specifically, once the uniform case is well understood, other upper probabilities can be defined by way of the inverse integral transform. Thus, we see the problem of studying a generalization of uniform probabilities as a first step toward developing more general upper probability models.

The purpose of this paper is to define one such generalization and to study the properties of these upper probabilities. Specifically, we study upper expectations that are symmetric in the sense that they are invariant with respect to equimeasurability. In terms of upper probabilities, this means that if A and B have the same Lebesgue measure, then $\bar{P}(A) = \bar{P}(B)$. This is a natural generalization of upper probabilities that are invariant with respect to permutations on a finite set. The finite case was studied in Wasserman and Kadane (1990a). Here, we take the underlying space to be the unit interval. Our hope is that by studying this special case in detail, insights will be gained that will help the study of more general cases in the future. Furthermore, we feel that the results are of interest in their own right.

The mathematical preliminaries for this investigation are set out in Section 2 and rely heavily on results from Ryff (1963, 1965, 1967, 1970). The relationship between symmetric upper expectations and closure properties of the set M are developed in Section 3. Upper distribution functions are studied in Section 4. We obtain an intriguing, geometric characterization of symmetric upper probabilities in terms of the upper distribution function. This result is analogous to the result that links probability measures to their distribution functions. Section 5 addresses the following question: Given a symmetric upper probability, what sets M could give rise to \bar{P} ? Section 6 characterizes Fine's (1988) generalized version of upper probabilities and the set of 2-alternating capacities [Choquet (1953)]. In particular, we show that in the symmetric case, the class of undominated generalized upper probabilities [Papamarcou and Fine (1986)] is empty. Some examples are considered in Section 7. Here, the difference between Choquet integrals and upper expectations is brought to light. This paper concludes with some discussion in Section 8.

2. Mathematical preliminaries. Let $\Omega = [0, 1]$ and let \mathcal{B} be the Borel subsets of Ω . Let \mathcal{P} be the set of all probability measures on \mathcal{B} and let M be a nonempty, convex, weakly closed set of probability measures on \mathcal{B} . It follows that M is tight and hence weakly compact. The set M generates an upper probability \bar{P} and an upper expectation \bar{E} by $\bar{P}(A) = \sup_{P \in M} P(A)$ and

$\bar{E}X = \sup_{P \in M} PX$, where $PX = \int X(w)P(dw)$. The lower probability and lower expectation \underline{P} and \underline{E} are defined by replacing the supremum with an infimum. Since $\underline{P}(A) = 1 - \bar{P}(A^c)$ and $\underline{E}X = -\bar{E}(-X)$, we will concentrate on \bar{P} and \bar{E} . For bounded, measurable X , the *Choquet integral* of X is defined by $\bar{C}X = \int_0^\infty \bar{P}(\bar{X} > t) dt$, where $\bar{X} = X + b$ and b is a constant chosen to make $X + b \geq 0$. $\bar{C}X$ does not depend on the choice of b [see Choquet (1953) and Huber and Strassen (1973)].

Let μ be Lebesgue measure and let L^1 be the space of Lebesgue integrable functions. The dual space of essentially bounded functions is denoted L^∞ ; \mathcal{C} is the set of bounded, continuous functions. If X and Y are measurable functions, we say that X and Y are *equimeasurable* and we write $X \sim Y$, if $\mu(\{X > t\}) = \mu(\{Y > t\})$ for every real number t . Here, $\{X > t\} \equiv \{w; X(w) > t\}$. Equimeasurable functions are studied in Ryff (1963, 1965, 1967, 1970) and most of the statements in this section are proved in those papers. Also, see Hardy, Littlewood and Pólya [(1952), Chapter 10] and Lorentz [(1953), Section 3.4]. If we replace Ω with a finite set and μ with counting measure, then X and Y are equimeasurable if Y can be obtained as a permutation of X . Thus, equimeasurability is a continuous version of permutation.

Corresponding to each measurable function X , there exists a unique right-continuous, nonincreasing function X^* such that $X \sim X^*$. We call X^* the *decreasing rearrangement* of X . If $X, Y \in L^1$, we say that X *majorizes* Y and write $Y < X$ if $\int_0^1 Y = \int_0^1 X$ and $\int_0^t Y^* \leq \int_0^t X^*$ for all $0 \leq t \leq 1$. The integrals are with respect to Lebesgue measure. Generally, we will write $\int Y$ for $\int_0^1 Y(\omega)\mu(d\omega)$. A linear transformation $T: L^1 \rightarrow L^1$ is *doubly stochastic* if $TX < X$ for every $X \in L^1$. The set \mathcal{D} of doubly stochastic operators is a convex, self-adjoint semigroup. (If $T \in \mathcal{D}$, its adjoint T^* acts on L^∞ , but T^* may be extended uniquely to act on L^1 .) An important result in Ryff [(1965), Theorem 3] is that if $X, Y \in L^1$, then $X < Y$ if and only if $X = TY$ for some $T \in \mathcal{D}$. For $X \in L^1$, let $\Lambda(X) = \{TX; T \in \mathcal{D}\}$.

A measurable function $\sigma: \Omega \rightarrow \Omega$ is *measure-preserving* if $\mu(E) = \mu(\sigma^{-1}(E))$ for all $E \in \mathcal{B}$. For each $X \in L^1$, there exists a measure-preserving function σ such that $X = X^* \circ \sigma$. Define T to be the operator induced by σ so that $TX^* = X$. Then the adjoint T^* satisfies $T^*X = X^*$. A useful fact is that for every $g \in L^1$ and every $A \in \mathcal{B}$, there is a doubly stochastic transformation T_A , say, such that $\int_A g = \int_0^{\mu(A)} T_A g$. We shall let χ_A denote the indicator function for A .

For $X: \Omega \rightarrow \Omega$ define $\underline{gr}(X) = \{(w, y) \in \Omega \times \Omega; X(w) \geq y\}$. Recall that a set A is *star-shaped* with respect to a point p if for each point $q \in A$, the line segment joining p and q is contained in A . We say that the function X is *star-shaped* if $(w, g) \in \underline{gr}(X)$ implies $(\alpha w, \alpha y) \in \underline{gr}(X)$ for $\alpha \in [0, 1]$. This is equivalent to $\underline{gr}(X)$ being star-shaped with respect to the origin. We shall call X *doubly star-shaped* if $\underline{gr}(X)$ is star-shaped with respect to the origin and to the point $(1, 1)$.

3. Symmetry and closure. From now on, assume that each $P \in M$ is absolutely continuous with respect to Lebesgue measure. This implies that \bar{P}

is nonatomic in the sense that $\mu(A) = 0$ implies that $\bar{P}(A) = 0$. Let $m = \{f = dP/d\mu; P \in M\}$. We will further assume that each $f \in m$ is essentially bounded. We say that M is *symmetric* if $\bar{E}X = \bar{E}Y$ whenever $X, Y \in L^\infty$ and $X \sim Y$. If M is symmetric, then we shall also say that m and \bar{E} are symmetric. We say M is *weakly symmetric* if $\mu(A) = \mu(B)$ implies that $\bar{P}(A) = \bar{P}(B)$. In this case, we shall also say that \bar{P} is weakly symmetric. In Example 4 (Section 7) we show that weak symmetry does not imply symmetry. We view symmetric sets M as a natural generalization of the concept of a uniform probability measure. We say that m is *closed with respect to majorization* if $f \in m$ and $g \prec f$ implies that $g \in m$. The rest of this section is devoted to proving the following theorem.

THEOREM 1. *\bar{E} is symmetric if and only if m is closed with respect to majorization.*

To prove the theorem, we first present some lemmas.

LEMMA 1. *$P \in M$ if and only if $PX \leq \bar{E}X$ for every $X \in C$.*

PROOF. $P \in M$ obviously implies that $PX \leq \bar{E}X$ for every $X \in \mathcal{C}$. Now suppose that $P \notin M$. Then M and $\{P\}$ can be strictly separated by a continuous linear functional l on \mathcal{P} . By Lemma 2.1 of Huber (1981), l is represented by some $X \in C$ so that $l(Q) = QX$ for every $Q \in \mathcal{P}$. Thus, $PX > \bar{E}X$ for some $X \in \mathcal{C}$ which establishes the other direction of the proof. \square

LEMMA 2. *Suppose \bar{E} is symmetric. Let $X, Y \in L^\infty$ and suppose that $Y \prec X$. Then, $\bar{E}Y \leq \bar{E}X$.*

PROOF. First suppose that $X \geq 0$. Then, $Y \geq 0$. By Ryff [(1965), Lemma 4], $\int fY \leq \int f^*Y^*$ for every $f \in m$. Also, by the same lemma, $\int f^*Y^* \leq \int f^*X^*$ since $Y \prec X$. Let σ be the measure-preserving transformation for which $f = f^* \circ \sigma$. Then, for every $f \in m$,

$$\begin{aligned} \int fY &\leq \int f^*Y^* \leq \int f^*X^* = \int (f^* \circ \sigma)(X^* \circ \sigma) \\ &= \int f(X^* \circ \sigma) \leq \bar{E}(X^* \circ \sigma) = \bar{E}(X^*) = \bar{E}(X). \end{aligned}$$

The penultimate equality holds since $X^* \circ \sigma \sim X^*$. Thus, $\bar{E}Y \leq \bar{E}X$. The extension to nonpositive X is obtained by adding a sufficiently large constant to X and then using the fact that doubly stochastic transformations are linear. \square

PROOF OF THEOREM 1. First we show that symmetry implies closure with respect to majorization. Assume \bar{E} is symmetric. Let $f \in m$ and suppose that $g \prec f$ so that $g = Tf$ for some $T \in \mathcal{D}$. Let T^* be the adjoint of T . Then, for

any $X \in L^\infty$, $\int Xg = \int XTf = \int T^*Xf \leq \overline{E}(T^*X) \leq \overline{E}X$. The last inequality is from Lemma 2. It now follows from Lemma 1 that $g \in m$ so that m is closed with respect to majorization.

Now suppose m is closed with respect to majorization. Let $X, Y \in L^\infty$ be such that $X \sim Y$. By Ryff [(1965), Theorem 1], there exists $T \in \mathcal{D}$ such that $Y = TX$. For every $f \in m$, $\int fY = \int fTX = \int T^*fX \leq \overline{E}X$. Hence, $\overline{E}Y \leq \overline{E}X$. A similar argument establishes that $\overline{E}X \leq \overline{E}Y$ so that $\overline{E}X = \overline{E}Y$ and hence, \overline{E} is symmetric. \square

4. Upper distribution functions. Let M be weakly symmetric and define the *upper and lower distribution functions* \overline{F} and \underline{F} by $\overline{F}(\omega) = \overline{P}([0, \omega])$ and $\underline{F}(\omega) = \underline{P}([0, \omega])$. Let \mathcal{H} be the set of concave distribution functions F on Ω such that $\underline{F}(\omega) \leq F(\omega) \leq \overline{F}(\omega)$ for every $\omega \in \Omega$.

LEMMA 3. *If M is weakly symmetric with upper distribution function \overline{F} then $\overline{F}(\omega) = \sup_{F \in \mathcal{H}} F(\omega)$.*

PROOF. By definition, $\overline{F}(\omega) \geq \sup_{F \in \mathcal{H}} F(\omega)$. For every $\omega \in \Omega$, there exists $f \in m$ such that $\overline{F}(\omega) = \int_0^\omega f$. Let $F(\omega) = \int_0^\omega f^*$. We claim that $F \in \mathcal{H}$ and that $F(\omega) = \overline{F}(\omega)$. That F is concave follows from the fact that f^* is nonincreasing. There exists a measure-preserving transformation σ so that $f^* \circ \sigma = f$. Hence, for every $t \in \Omega$,

$$\begin{aligned} F(t) &= \int_0^t f^* = \int_0^t f^* \chi_{[0,t]} = \int_0^1 (f^* \circ \sigma)(\chi_{[0,t]} \circ \sigma) = \int_0^1 f(\chi_{[0,t]} \circ \sigma) \\ &= \int_0^1 f(\chi_{\sigma^{-1}[0,t]}) \leq \overline{P}(\sigma^{-1}[0,t]) = \overline{P}([0,t]) = \overline{F}(t). \end{aligned}$$

Thus, $F \in \mathcal{H}$. And $F(\omega) = \int_0^\omega f^* \geq \int_0^\omega f = \overline{F}(\omega)$ implies that $F(\omega) = \overline{F}(\omega)$, which establishes the equality. \square

LEMMA 4. *Let \overline{F} be the upper distribution function for a weakly symmetric set M . Then, for every $\alpha, \omega \in \Omega$, $\overline{F}(\alpha\omega) \geq \alpha\overline{F}(\omega)$ and $\overline{F}(\alpha\omega + 1 - \alpha) \geq \alpha\overline{F}(\omega) + 1 - \alpha$.*

PROOF. By Lemma 3, $\overline{F}(\omega) = \sup_{F \in \mathcal{H}} F(\omega)$, where each $F \in \mathcal{H}$ is concave and satisfies $F(0) = 0$ and $F(1) = 1$. Hence,

$$\begin{aligned} \overline{F}(\alpha\omega) &= \sup_{F \in \mathcal{H}} F(\alpha\omega) = \sup_{F \in \mathcal{H}} F(\alpha\omega + (1 - \alpha)0) \\ &\geq \sup_{F \in \mathcal{H}} \alpha F(\omega) + (1 - \alpha)F(0) = \alpha \sup_{F \in \mathcal{H}} F(\omega) = \alpha\overline{F}(\omega). \end{aligned}$$

Similarly,

$$\overline{F}(\alpha\omega + 1 - \alpha) \geq \sup_{F \in \mathcal{H}} (\alpha F(\omega) + 1 - \alpha) = \alpha\overline{F}(\omega) + 1 - \alpha. \quad \square$$

For every $t \in [0, 1]$, define a density function f_t by

$$f_t(\omega) = \begin{cases} \bar{F}(t)/t, & 0 \leq \omega < t, \\ (1 - \bar{F}(t))/(1 - t), & t \leq \omega \leq 1. \end{cases}$$

Note that $f_t = f_t^*$. Let $F_t(\omega) = \int_0^\omega f_t$.

THEOREM 2. *Let $\bar{F}: \Omega \rightarrow \Omega$. Then the following three statements are equivalent:*

- (i) \bar{F} is the upper distribution function for some weakly symmetric M .
- (ii) $gr(\bar{F})$ is doubly star-shaped.
- (iii) For every $t \in \Omega$, $F_t(\omega) \leq \bar{F}(\omega)$ for all $\omega \in \Omega$.

PROOF. We begin by showing that (i) implies (ii). Suppose $(\omega, y) \in gr(\bar{F})$. Then $\alpha y \leq \alpha \bar{F}(\omega)$. By Lemma 4, $\alpha \bar{F}(\omega) \leq \bar{F}(\alpha\omega)$ so that $(\alpha\omega, \alpha y) \in gr(\bar{F})$. Hence, $gr(\bar{F})$ is star-shaped with respect to the origin. A similar argument shows that $gr(\bar{F})$ is star-shaped with respect to $(1, 1)$.

To show that (ii) implies (iii), note that $F_t(t) = \bar{F}(t)$ by definition. Since $gr(\bar{F})$ is star-shaped with respect to the origin and since F_t is linear on $[0, t]$, it follows that $F_t(\omega) \leq \bar{F}(\omega)$ on $[0, t]$. Since $gr(\bar{F})$ is star-shaped with respect to $(1, 1)$ and since F_t is linear on $[t, 1]$, it follows that $F_t(\omega) \leq \bar{F}(\omega)$ on $[t, 1]$.

Finally, we show that (iii) implies (i). Let m be the closed, convex hull of $\cup_{t \in \Omega} \Lambda(f_t)$ and let $M = \{P = f d\mu; f \in m\}$. By Theorem 1, M is symmetric. By construction \bar{F} is the upper distribution function for M . \square

As a result of Theorem 2, it is simple to check, graphically, whether a set function is an upper probability. Although characterizations for upper probabilities exist, they are usually cumbersome. The symmetry assumption thus brings about considerable conceptual simplification.

We remark that weak symmetry implies that the Choquet integral \bar{C} is symmetric and $\bar{C}X = \int X^* d\bar{F}$. See Example 4, Section 7, for more on this point.

5. Inner and outer approximations. Let \bar{P} be weakly symmetric. In this section we investigate the following question: Given only \bar{P} , what can be said about the M that generated \bar{P} ? Huber [(1981), Section 10.2] has some results about this question. The weakly symmetric case provides interesting and precise results.

Let $\Gamma(\bar{P})$ be the class of all symmetric m that generate \bar{P} . Let m_0 be the closed, convex hull of $\cup_{t \in \Omega} \Lambda(f_t)$ and let $\tilde{m} = \{f; \int_A f \leq \bar{P}(A) \text{ for all } A \in \mathcal{B}\}$. Also, let M_0 and \tilde{M} be the sets of probability measures corresponding to m_0 and \tilde{m} .

LEMMA 5. *If \bar{P} is weakly symmetric, then \tilde{m} is symmetric.*

PROOF. Suppose $f \in \tilde{m}$ and that $g < f$. We must show that $g \in \tilde{m}$. There exists a measure-preserving transformation σ so that $f^* \circ \sigma = f$. Hence,

$$\begin{aligned} \int_0^\omega g &\leq \int_0^\omega g^* \leq \int_0^\omega f^* = \int_0^1 f^* \chi_{[0, \omega]} = \int_0^1 (f^* \circ \sigma)(\chi_{[0, \omega]} \circ \sigma) \\ &= \int_0^1 f \chi_{\sigma^{-1}[0, \omega]} \leq \bar{P}(\sigma^{-1}[0, \omega]) = \bar{P}([0, \omega]). \end{aligned}$$

Thus, $\int_0^\omega g \leq \bar{P}([0, \omega])$ for every ω . Now, $\int_A g = \int_0^{\mu(A)} T_A g \leq \bar{P}[0, \mu(A)] = \bar{P}(A)$ since $T_A g < g < f$. Therefore, $g \in \tilde{m}$. \square

LEMMA 6. Let $m \in \Gamma(\bar{P})$. Then, for every $t \in [0, 1]$, $f_t \in m$.

PROOF. There exists $g \in m$ such that $\int_0^t g = \bar{F}(t)$. Let G^* be the distribution function for g^* . Then $G^*(t) = \bar{F}(t)$ and $g^* \in m$. Since $F_t(0) = G^*(0)$, $F_t(t) = G^*(t)$, F_t is linear on $[0, t]$ and G^* is concave, it follows that $F_t(\omega) \leq G^*(\omega)$ for $\omega \in [0, t]$. A similar argument shows that that $F_t(\omega) \leq G^*(\omega)$ for $\omega \in [t, 1]$. Hence, $f_t < g^*$. By Theorem 1, $f_t \in m$. \square

THEOREM 3. Let \bar{P} be weakly symmetric and let $\Gamma(\bar{P})$ be the set of all symmetric sets m that generate the upper probability \bar{P} . Then, $m_0 = \bigcap_{m \in \Gamma(\bar{P})} m$ and $\tilde{m} = \bigcup_{m \in \Gamma(\bar{P})} m$.

PROOF. By Lemma 6, it follows that $m_0 \subset m$, for every $m \in \Gamma(\bar{P})$. Since $m_0 \in \Gamma(\bar{P})$, it follows that $\bigcap_{m \in \Gamma(\bar{P})} m = m_0$. It is easy to show that $m \subset \tilde{m}$, for every $m \in \Gamma(\bar{P})$. By Lemma 5, $\tilde{m} \in \Gamma(\bar{P})$. Hence, $\bigcup_{m \in \Gamma(\bar{P})} m = \tilde{m}$. \square

Suppose now that a weakly symmetric \bar{P} is given and we need to compute $\bar{E}X$ for some X . Clearly the best we can do is to bound $\bar{E}X$. This is a generalization of the general moment problem [Kemperman (1968)] where the expectation of a finite set of random variables is given and one must deduce bounds on the expectation of another random variable. In our case, we are given bounds on the expectation of indicator functions, namely, the upper probabilities. Theorem 3 shows that $\bar{E}X$ is bounded above and below by the upper expectation over the sets M_0 and \tilde{M} . It is easy to see that

$$E_0 X \equiv \sup_{P \in M_0} PX = \max_{\omega \in \Omega} \left(\bar{F}(\omega) \underline{av}_{X^*}(\omega) + (1 - \bar{F}(\omega)) \overline{av}_{X^*}(\omega) \right),$$

where $\underline{av}_Y(\omega) = \int_0^\omega Y/\omega$ and $\overline{av}_Y(\omega) = \int_\omega^1 Y/(1 - \omega)$. Thus, the lower bound is easily computed. Regarding the upper bound, it is straightforward that $\bar{E}X \leq \tilde{E}X \leq \bar{C}X$, where $\tilde{E}X = \sup_{P \in \tilde{M}} PX$. Equality of $\tilde{E}X$ and \bar{C} holds if \bar{P} is 2-alternating. [See Section 6 for the definition of 2-alternating; the equality is proved in Choquet (1953).] Little seems to be known about $\tilde{E}X$ if \bar{P} is not 2-alternating, but see Example 3 in Section 7. Also, little attention has been paid to the difference between the Choquet integral and the upper expectation in case \bar{P} is 2-alternating but M is strictly contained in \tilde{M} . Nonetheless, we always have $E_0 X \leq \bar{E}X \leq \tilde{E}X \leq \bar{C}X$.

Now we characterize those upper probabilities for which E_0 and \tilde{E} coincide.

THEOREM 4. *Let \bar{F} be weakly symmetric with upper distribution \bar{F} . Then the following three statements are equivalent:*

- (i) $\bigcup_{t \in \Omega} \Lambda(f_t) = \tilde{m}$.
- (ii) $F \in \mathcal{H}$ implies that $F \leq F_t$ for some $t \in \Omega$.
- (iii) $\bar{F} \equiv F_t$ for some $t \in \Omega$.

Before proving Theorem 4, we prove the following lemma.

LEMMA 7. *Let f be a probability density function and let $F(\omega) = \int_0^\omega f^*$. Then, $f \in \tilde{m}$ if and only if $F \in \mathcal{H}$.*

PROOF. Let $f \in \tilde{m}$. Write $f = f^* \circ \sigma$, where σ is measure-preserving. Then,

$$\begin{aligned} F(\omega) &= \int f^* \chi_{[0, \omega]} = \int (f^* \circ \sigma)(\chi_{[0, \omega]} \circ \sigma) \\ &= \int f(\chi_{[0, \omega]} \circ \sigma) = \int f \chi_{\sigma^{-1}[0, \omega]} \leq \bar{P}(\sigma^{-1}[0, \omega]) = \bar{P}([0, \omega]) = \bar{F}(\omega), \end{aligned}$$

and F is concave since f^* is nonincreasing. Hence, $F \in \mathcal{H}$.

Conversely, suppose that $F \in \mathcal{H}$ so that $\int_0^\omega f^* \leq \bar{F}(\omega)$. Then,

$$\begin{aligned} \int_A f &= \int_0^{\mu(A)} T_A f \leq \int_0^{\mu(A)} (T_A f)^* \leq \int_0^{\mu(A)} f^* \\ &\leq \bar{F}(\mu(A)) = \bar{P}([0, \mu(A)]) = \bar{P}(A). \end{aligned}$$

Therefore, $f \in \tilde{m}$. \square

PROOF OF THEOREM 4. (i) implies (ii). Let $F \in \mathcal{H}$. By the previous lemma, $f \in \tilde{m}$, where $f = dF/d\mu$ and $\int_0^\omega f = \int_0^\omega f^*$ since F is concave. By (i), $f < f_t$ for some t so that $F \leq F_t$.

(ii) implies (iii). We will show that “not (iii)” implies “not (ii).” Suppose that, for every $t \in \Omega$, \bar{F} is not identically equivalent to F_t .

CASE 1. Assume that \bar{F} is concave. Now, $F_t \leq \bar{F}$ for every t . It cannot be that $\bar{F} \leq F_t$, for this would imply that $\bar{F} \equiv F_t$, contradicting the assumption. Therefore we have found an element of \mathcal{H} , namely, \bar{F} itself, that is not dominated by any F_t so the implication in (ii) fails.

CASE 2. Assume that \bar{F} is not concave. We can find t and u such that $0 < t < u < 1$ and $\bar{F}(\omega) < c(\omega)$ for $\omega \in (t, u)$, where $c(\omega)$ is the chord joining the points $(t, \bar{F}(t))$ and $(u, \bar{F}(u))$. Let $G = \alpha F_t + (1 - \alpha)F_u$, where $\alpha \in (0, 1)$. It is easy to see that $G \in \mathcal{H}$ and we claim that G is not dominated by any F_z . To see this, consider F_z . For $z \leq t$, $F_z(u) \leq F_t(u) < G(u)$ so F_z fails to dominate G . For $z \geq u$, $F_z(t) \leq F_u(t) < G(t)$ so F_z fails to dominate G . Now

suppose that $t < z < u$. For F_z to dominate G on $[0, t]$, z must be less than or equal to the point a_1 , where $\omega(\alpha\bar{F}(t)/t + (1 - \alpha)\bar{F}(u)/u)$ intersects \bar{F} . Since \bar{F} is strictly less than $c(\omega)$ on (t, u) , a_1 is strictly less than the point b_1 , where $\omega(\alpha\bar{F}(t)/t + (1 - \alpha)\bar{F}(u)/u)$ intersects $c(\omega)$. Some algebra shows that $b_1 = tu/(\alpha u - \alpha t + t)$. For F_z to dominate G on $[u, 1]$, z must be greater than or equal to the point a_2 , where $k(\omega) \equiv \alpha\bar{F}(t) + \alpha(\omega - t)(1 - \bar{F}(t))/(1 - t) + (1 - \alpha)\bar{F}(u) + (1 - \alpha)(\omega - u)(1 - \bar{F}(u))/(1 - u)$ intersects \bar{F} . Since \bar{F} is strictly less than $c(\omega)$ on (t, u) , a_2 is strictly greater than the point b_2 , where $k(\omega)$ intersects $c(\omega)$, and $b_2 = (u - \alpha u - tu + \alpha t)/(1 - \alpha u + \alpha t - t)$. So far we have that $z \leq a_1 < b_1$ and $z \geq a_2 > b_2$. However, from the fact that $\alpha \in (0, 1)$ and $\alpha(t - u)^2(1 - \alpha) > 0$, we conclude that $b_1 < b_2$ so there can be no z simultaneously satisfying $z < b_1$ and $z > b_2$. Hence, there is no F_z that dominates G . Thus, the implication in (ii) fails.

(iii) implies (i). Let $f \in \tilde{m}$ and let F be the distribution function for f^* . Then, by Lemma 7, $F(\omega) \leq \bar{F}(\omega) = F_t(\omega)$. Therefore, $f < f_t$ so that $f \in \bigcup_{t \in \Omega} \Lambda(f_t)$. \square

Thus, we have shown that the inner and outer approximations agree, just when the upper distribution is of the form F_t for some t . Upper probabilities of this form arise in Bayesian robustness and are discussed in more detail in Example 2 in Section 7.

6. Hierarchy of upper probabilities. The upper probabilities considered thus far can both be generalized and specialized. That is, there is a hierarchy of upper probabilities. In this section we show that some of these upper probabilities may be given simple geometric characterizations based on their upper distribution functions. Fine (1988) considers upper probabilities that are more general than those defined in this article so far. Formally, we shall call $\bar{P}: \mathcal{B} \rightarrow [0, 1]$ a *generalized upper probability* if it satisfies the following:

1. $\bar{P}(\emptyset) = 0, \bar{P}(\Omega) = 1;$
2. $A \cap B = \emptyset$ implies that $\bar{P}(A \cup B) \leq \bar{P}(A) + \bar{P}(B)$ and $\bar{P}(A^c \cap B^c) \leq \bar{P}(A^c) + \bar{P}(B^c) - 1.$

Fine (1988), Grize and Fine (1987), Kumar and Fine (1985), Papamarcou and Fine (1986) and Walley and Fine (1982) give arguments to suggest that generalized upper probabilities are of great importance and can be used to model real-world phenomena.

A generalized upper probability is *dominated* if there exists a probability measure P on \mathcal{B} such that $P(A) \leq \bar{P}(A)$ for every $A \in \mathcal{B}$. As before, \bar{P} is weakly symmetric if $\mu(A) = \mu(B)$ implies that $\bar{P}(A) = \bar{P}(B)$. Throughout this section we assume that \bar{P} is nonatomic in that if $\mu(A) = 0$, then $\bar{P}(A) = 0$.

Another important class of upper probabilities is 2-alternating upper probabilities, also known as 2-alternating Choquet capacities. We say that \bar{P} is 2-alternating if for every $A, B \in \mathcal{B}$, $\bar{P}(A \cup B) \leq \bar{P}(A) + \bar{P}(B) - \bar{P}(A \cap B)$.

These upper probabilities arise in analysis [Choquet (1953)], classical robustness [Huber (1973), Huber and Strassen (1973), Buja (1984, 1985, 1986), Bednarski (1982) and Rieder (1977)], Bayesian robustness [Wasserman and Kadane (1990b)] and in other contexts [Walley (1981, 1991)]. Recall that the Choquet integral is defined by $\bar{C}X = \int_0^\infty \bar{P}(X > t) dt$ for bounded $X \geq 0$. It is known that if \bar{P} is 2-alternating, then $\bar{E}X = \bar{C}X$ [see Choquet (1953) and Huber and Strassen (1973)].

Let \mathcal{U}_1 be the set of weakly symmetric generalized upper probabilities, \mathcal{U}_2 the set of weakly symmetric dominated generalized upper probabilities, \mathcal{U}_3 the set of weakly symmetric upper probabilities and \mathcal{U}_4 the set of weakly symmetric 2-alternating upper probabilities. The class \mathcal{U}_3 is the class we have been using in the previous sections. It is well known that $\mathcal{U}_4 \subset \mathcal{U}_3 \subset \mathcal{U}_2 \subset \mathcal{U}_1$. See Walley and Fine (1982), for example. We now investigate some properties of these classes. The proof of the next lemma is straightforward and is omitted.

LEMMA 8. *If $\bar{P} \in \mathcal{U}_1$, then its upper distribution function \bar{F} satisfies the following:*

- (i) $\bar{F}(0) = 0$ and $\bar{F}(1) = 1$;
- (ii) $\bar{F}(x + y) \leq \bar{F}(x) + \bar{F}(y)$;
- (iii) $\bar{F}(1 - x - y) \leq \bar{F}(1 - x) + \bar{F}(1 - y) - 1$.

Conversely, if $\bar{F}: \Omega \rightarrow \Omega$ satisfies (i)–(iii), then $\bar{P} \in \mathcal{U}_1$, where $\bar{P}(A) = F(\mu(A))$.

LEMMA 9. *If \bar{F} is the upper distribution function for $\bar{P} \in \mathcal{U}_1$, then $\bar{F}(\omega) \geq \omega$ for every $\omega \in \Omega$.*

PROOF. If $\omega_1, \dots, \omega_n \in \Omega$ are such that $0 \leq \sum \omega_i \leq 1$, then from Lemma 8 it follows that $\bar{F}(\sum \omega_i) \leq \sum \bar{F}(\omega_i)$ and $\bar{F}(1 - \sum \omega_i) \leq \sum \bar{F}(1 - \omega_i) + (n - 1)$. Let $k \geq 1$ be an integer and let $\omega = 1/k$. If $\bar{F}(\omega) < \omega$, then $1 = \bar{F}(1) = \bar{F}(k\omega) = \bar{F}(\omega + \dots + \omega) \leq k\bar{F}(\omega) < k\omega = 1$. This is a contradiction, so $\bar{F}(\omega) \geq \omega$.

If $\bar{F}(1 - \omega) < 1 - \omega$, then

$$\begin{aligned} \bar{F}(\omega) &= \bar{F}(1 - (k - 1)/k) = \bar{F}(1 - 1/k - \dots - 1/k) \\ &\leq \bar{F}(1 - 1/k) + \dots + \bar{F}(1 - 1/k) + (k - 2) \\ &= (k - 1)\bar{F}(1 - 1/k) + (k - 2) \\ &< (k - 1)(1 - 1/k) + (k - 2) = 1/k = \omega. \end{aligned}$$

Again, this is a contradiction so $\bar{F}(1 - \omega) \geq 1 - \omega$.

Now we proceed inductively. Suppose that $\bar{F}(r/k) \geq r/k$ and

$$\bar{F}(1 - r/k) \geq 1 - r/k \quad \text{for } r \in \{1, \dots, h\}.$$

Consider $r = h + 1$. There exists an integer s such that

$$1 - s(h + 1)/k = j/k, \quad 0 \leq j/k < h/k.$$

Suppose that $\bar{F}((h + 1)/k) < (h + 1)/k$. Then,

$$\bar{F}(1 - j/k) = \bar{F}(s(h + 1)/k) \leq s\bar{F}((h + 1)/k) < s(h + 1)/k.$$

But, $j \in \{1, \dots, h\}$, so by assumption, $\bar{F}(1 - j/k) \geq 1 - j/k = s(h + 1)/k$. Thus we have a contradiction, so it must be that $\bar{F}((h + 1)/k) \geq (h + 1)/k$.

From this we conclude that $\bar{F}(\omega) \geq \omega$ whenever ω is rational. For any $\omega \in \Omega$ and any $\varepsilon > 0$, there exists a rational number $r \in \Omega$ such that $r \leq \omega < r + \varepsilon$. Then, $\bar{F}(\omega) \geq \bar{F}(r) \geq r > \omega - \varepsilon$. Thus, $\bar{F}(\omega) \geq \omega$. \square

By Lemma 9, if $\bar{P} \in \mathcal{U}_1$, then $\bar{P}(A) = \bar{F}(\mu(A)) \geq \mu(A)$. Thus we have the following corollary.

COROLLARY 1. $\mathcal{U}_1 = \mathcal{U}_2$. *That is, there do not exist nonatomic, symmetric, undominated generalized probabilities.*

The following example shows that \mathcal{U}_2 is proper subset of \mathcal{U}_3 .

EXAMPLE 1. Define \bar{F} by $\bar{F}(0) = 0$, $\bar{F}(1/4) = 1/2$, $\bar{F}(1/2) = x$, $\bar{F}(3/4) = \bar{F}(1) = 1$ and \bar{F} is piecewise linear otherwise. Let $\bar{P}(A) = \bar{F}(\mu(A))$. Then, $\bar{P} \in \mathcal{U}_2$ if and only if $1/2 \leq x \leq 1$ and $\bar{P} \in \mathcal{U}_3$ if and only if $2/3 \leq x \leq 1$.

Now we characterize the class \mathcal{U}_4 of 2-alternating capacities.

THEOREM 5. *Let $\bar{F}: \Omega \rightarrow \Omega$ satisfy $\bar{F}(0) = 0$ and $\bar{F}(1) = 1$. Define \bar{P} by $\bar{P}(A) = \bar{F}(\mu(A))$ and let $\tilde{m} = \{f \in L^1; \int_A f \leq \bar{P}(A) \text{ for all } A \in \mathcal{B}\}$. Then the following three statements are equivalent:*

- (i) $\bar{P} \in \mathcal{U}_4$.
- (ii) \bar{F} is concave.
- (iii) *There exists a density function f such that $\tilde{m} = \Lambda(f)$.*

PROOF. (i) if and only if (ii). To show that (i) implies (ii), we argue as in Bednarski [(1981), Lemma 3.1]. Let $x, y \in \omega$ with $y < x$. Set $A = [0, (x + y)/2]$ and $B = [0, y] \cup [(x + y)/2, x]$. Apply the 2-alternating condition to A and B to deduce that $\bar{F}(x)/2 + \bar{F}(y)/2 \leq \bar{F}((x + y)/2)$ so that \bar{F} is midconcave. The boundedness of \bar{F} implies that \bar{F} is concave [Roberts and Varberg (1973), Section 72].

To show that (ii) implies (i) we follow Buja [(1986), page 151]. Set $x = \mu(A)$, $y = \mu(B)$, $u = \mu(A \cap B)$ and $v = \mu(A \cup B)$, so that $u + v = x + y$, $u \leq x \leq v$ and $u \leq y \leq v$. There exists $\alpha \in \Omega$ such that $x = \alpha u + (1 - \alpha)v$ and $y = (1 - \alpha)u + \alpha v$. Applying concavity, we have $\bar{P}(A \cap B) + \bar{P}(A \cup B) = \bar{F}(u) + \bar{F}(v) \leq \bar{F}(x) + \bar{F}(y) = \bar{P}(A) + \bar{P}(B)$, so that \bar{P} is 2-alternating.

(ii) if and only if (iii). To see that (ii) implies (iii), note that \bar{F} is a distribution function. Let $f = d\bar{F}/d\mu$. Since \bar{F} is concave, $f = f^*$ almost surely so take $f = f^*$. We need to show that $g \in \tilde{m}$ if and only if $g < f$. Suppose that $g \in \tilde{m}$. By Lemma 5, $g^* \in \tilde{m}$ so that $\int_0^\omega g^* \leq \int_0^\omega f$. Hence, $g < f$. Now suppose that $g < f$. Then, $\int_A g = \int_0^{\mu(A)} T_A g \leq \int_0^{\mu(A)} f = \bar{F}(\mu(A)) = \bar{P}(A)$. Therefore, $g \in \tilde{m}$.

Finally we show that (iii) implies (ii). For this it suffices to show that $\bar{F}(\omega) = \int_0^\omega f^*$. Clearly, $\bar{F}(\omega) \geq \int_0^\omega f^*$ since $f^* \in \tilde{m}$. For any $g \in \tilde{m}$, $g < f$ implies that $\int_0^\omega g \leq \int_0^\omega g^* \leq \int_0^\omega f^*$ so that $\sup_{g \in \tilde{m}} \int_0^\omega g = \bar{F}(\omega) \leq \int_0^\omega f^*$. \square

The equivalence between conditions (i) and (iii) is intriguing. We may paraphrase the equivalence by saying that \bar{P} is weakly symmetric and 2-alternating if and only if \tilde{m} is the orbit of a single density function. In this sense, 2-alternating upper probabilities are very simple symmetric upper probabilities. This characterization establishes, then, a 1–1 relationship between 2-alternating, weakly symmetric Choquet capacities and nonincreasing density functions. As such, it generates a rich source of examples of 2-alternating capacities. Furthermore, if f and g are two density functions, neither of which is majorized by the other, then the closed convex hull of the union of $\{h; h < f\}$ and $\{h; h < g\}$ generates a symmetric set m that by virtue of Theorem 5, cannot be 2-alternating. Thus, we have a source of non-2-alternating upper probabilities.

EXAMPLE 1 (Continued). It is easy to see that $\bar{P} \in \mathcal{U}_4$ if and only if $3/4 \leq x \leq 1$. Thus, \mathcal{U}_3 is strictly contained in \mathcal{U}_2 . We have shown that $\mathcal{U}_4 \subset \mathcal{U}_3 \subset \mathcal{U}_2 = \mathcal{U}_1$, where the containments are proper.

In Section 5, we established that E_0 and \tilde{E} coincide, just when $\bar{F} \equiv F_t$ for some t . Since each F_t is concave, we have the following corollary to Theorem 5.

COROLLARY 2. *If \bar{P} is not 2-alternating, then E_0 and \tilde{E} cannot coincide.*

In general, when \bar{P} is not 2-alternating, \bar{C} will overestimate \bar{E} for two reasons. First, \bar{C} will overestimate \tilde{E} since \bar{P} is not 2-alternating. Second, \tilde{E} will overestimate \bar{E} since m may be a proper subset of \tilde{m} . When \bar{P} is 2-alternating, we know that $\bar{C} = \tilde{E}$. However, \tilde{E} may still overestimate \bar{E} . Only in the special case where $\bar{F} \equiv F_t$ can we deduce from the upper probability that \bar{C} , \tilde{E} and \bar{E} all coincide.

7. Examples. In this section we consider three examples in detail. The first two examples are special cases of classes of probability measures that arise in Bayesian robustness. The sets of probabilities in those examples may be thought of as probabilities that are approximately uniform. The second of these (Example 3) emphasizes that even if an upper probability is 2-alternating, the Choquet integral may overestimate the upper expectation since m

may be a proper subset of \tilde{m} . The third example illustrates the difference between symmetry and weak symmetry.

EXAMPLE 2. Let l and u be real numbers such that $0 < l < 1 < u$. Let $m = \{f; l \leq f(\omega) \leq u, \text{ for } \mu\text{-almost all } \omega\}$. This is a special case of the class considered by Lavine (1988) in Bayesian robustness. Let $L(A) = l\mu(A)$ and $U(A) = u\mu(A)$. It is easy to show that $\bar{P}(A) = \min(U(A), 1 - L(A^c))$. Suppose $f \in m$ and $g < f$. If $g(\omega) > u$ on a set A of positive Lebesgue measure, then $\int_0^{\mu(A)} g^* > u\mu(A) \geq \int_0^{\mu(A)} f^*$, contradicting the fact that $g < f$. So $g \leq u$ almost surely. Similarly, if $g(\omega) < l$ on a set A of positive Lebesgue measure, then $\int_{1-\mu(A)}^1 g^* < l(1 - \mu(A)) \leq \int_{1-\mu(A)}^1 f^*$ so that $\int_0^{1-\mu(A)} g^* > \int_0^{1-\mu(A)} f^*$, which again contradicts $g < f$. Thus, g is almost surely between l and u so that $g \in m$. This establishes that m is symmetric. From the formula for \bar{P} we get

$$\bar{F}(\omega) = \begin{cases} u\omega, & \text{if } \omega \leq \Delta, \\ 1 - l + l\omega, & \text{if } \omega \geq \Delta, \end{cases}$$

where $\Delta = (1 - l)/(u - l)$. Thus, $\bar{F} \equiv F_\Delta$. Recall from Theorem 4 of Section 5, that it is precisely upper distribution functions of this form for which the inner approximation m_0 and the outer approximation \tilde{m} coincide. Therefore, the examples of this form obtained by varying l and u generate the class of all such weakly symmetric upper probabilities. From direct calculations we get $E_0 X = \bar{E}X = \underline{E}X = \bar{C}X = \underline{C}X^* = u\int_0^\Delta X^* + l\int_\Delta^1 X^*$ for $X \geq 0, X \in L^\infty$. In conclusion, this class is unusual in that it is 2-alternating, the inner and outer approximation agree, and a simple formula exists for computing upper expectations.

EXAMPLE 3. A similar example is obtained from another class of probabilities used in Bayesian robustness by DeRobertis and Hartigan (1981). Let

$$m = \{f \in L^1; \text{ess sup } f(\omega) / \text{ess inf } f(\omega) \leq k\},$$

where $k > 1$. To see that m is symmetric, note that if T is doubly stochastic, then $\text{ess sup } Tf \leq \text{ess sup } f$ and $\text{ess inf } Tf \geq \text{ess inf } f$. Thus, if $f \in m$ and $g < f$ so that $g = Tf$, say, then $g \in m$.

For $t \in [0, 1]$ define the density f_t by

$$f_t(\omega) = \begin{cases} k/(t\Delta + 1), & \text{if } \omega < t, \\ 1/(t\Delta + 1), & \text{if } \omega \geq t, \end{cases}$$

where $\Delta = k - 1$. Clearly $f_t \in m$. It is easy to see that $\int_0^\omega f_\omega = \bar{P}([0, \omega])$ for if there is a g such that $f_\omega^\omega g > \int_0^\omega f_t$, then we can find a set of positive Lebesgue measure over which $g(\omega)/g(\omega') > k$. By direct calculation, $\bar{F}(\omega) = k\omega/(\omega\Delta + 1)$. This is concave so that, by Theorem 5, \bar{P} is 2-alternating. However, we claim that m is strictly contained in \tilde{m} . To see this, let $f = d\bar{F}/d\omega = k/(\Delta\omega + 1)^2$. Then, $f(0)/f(1) = k^2$ so f is not in m . However, $f = f^*$, and for any $A, \int_A f = \int_0^{\mu(A)} T_A f \leq \int_0^{\mu(A)} f^* = \bar{F}(\mu(A)) = \bar{P}(A)$. Hence $f \in \tilde{m}$. Thus, despite the fact that \bar{P} is 2-alternating, the Choquet integral will typically

overestimate \bar{E} . This is a point that is often overlooked. Interest usually centers on establishing that \bar{P} is 2-alternating. The Choquet integral may be calculated explicitly and, in fact, $\bar{C}X = \int X^* d\bar{F}$.

Now we show that $m_0 = m$. To do this, we show that for every nonnegative $X \in L^\infty$ and every $f \in m$, there is an f_t such that $\int f_t X \geq \int f X$ so that $E_0 X = \bar{E}X$. First suppose that X and f are nonincreasing. Let $t = (1 - f(1))/(f(0) - f(1))$. Then $\int f_t X \geq \int f X$ if and only if $\int_0^t X(f_t - f) \geq \int_t^1 X(f - f_t)$. Now,

$$\int_0^t X(f_t - f) \geq X(t) \int_0^t (f_t - f) = X(t) \int_t^1 (f - f_t) \geq \int_t^1 X(f - f_t),$$

as required. Now remove the restriction that X and f be nonincreasing. Then, $\int Xf \leq \int X^* f^* \leq \int X^* f_t^*$ for $t = (1 - f^*(1))/(f^*(0) - f^*(1))$. We conclude that

$$\bar{E}X = E_0 X^* = \max_{\omega \in \Omega} \left[\{\omega\Delta + 1\}^{-1} \left\{ \Delta \int_0^\omega X^* + \bar{X} \right\} \right],$$

where $\bar{X} = \int_0^1 X$. Thus, $m_0 = m \subset \tilde{m}$.

It is interesting that even though the Choquet integral fails to produce the correct upper bound, we still have a simple formula for computing the upper expectation. To see the difference between \bar{E} and \bar{C} , consider $X(\omega) = \omega$. Then

$$\bar{C}X = \bar{C}X^* = k \int_0^1 \frac{1 - v}{(\Delta v + 1)^2} dv = \frac{k(k - 1 - \log(k))}{(k - 1)^2}.$$

On the other hand,

$$\begin{aligned} \bar{E}X = E_0 X^* &= \max_{\omega \in [0, 1]} \left[\{\omega\Delta + 1\}^{-1} \left\{ \Delta \int_0^\omega X^* + \bar{X} \right\} \right] \\ &= \max_{\omega \in [0, 1]} \frac{\Delta\omega(1 - \omega/2) + 1/2}{\omega\Delta + 1} = \frac{k^{3/2} - k}{k^{1/2}(k - 1)}. \end{aligned}$$

It seems that \bar{C} does worst at estimating \bar{E} around $k = 17$.

EXAMPLE 4. Here we construct a set that is weakly symmetric but is not symmetric. Let $a = 1 + \delta$, $b = 1$ and $c = 1 - \delta$, where $0 < \delta < 1$. Let $A_1 = [0, 1/3]$, $A_2 = [1/3, 2/3]$ and $A_3 = [1/3, 1]$. Let g equal c on A_1 , b on A_2 and a on A_3 . For every measurable set A we will define a density function f_A . To do so, we first define a function γ_A by

$$\gamma_A(\omega) = \begin{cases} \mu([\omega, 1] \cap A), & \text{if } \omega \in A, \\ \mu([\omega, 1] \cap A^c) + \mu(A), & \text{if } \omega \notin A. \end{cases}$$

Now set $f_A(\omega) = g(1 - \gamma(\omega))$. Loosely speaking, f_A may be described in the following way. Over the set A , f_A is identical to that portion of g over $[1 - \mu(A), 1]$, and over A^c , f_A is identical to that portion of g over $[0, \mu(A)]$. Note that each f_A is equimeasurable with g and takes on the values a , b and

c equally often. Also note that within A , f_A is increasing and similarly, within A^c , f_A is increasing. If $A = [0, t]$, then f_A is that density obtained by shifting g to the right by t units, mod 1.

Let m consist of all convex combinations of such densities. It follows immediately that $\int_A f_A = \int_{1-\mu(A)} g$ and

$$\bar{P}(A) = \int_A f_A = \begin{cases} \mu(A)a, & \text{if } \mu(A) \leq 1/3, \\ a/3 + b(\mu(A) - 1/3), & \text{if } 1/3 < \mu(A) \leq 2/3, \\ a/3 + b/3 + c(\mu(A) - 2/3), & \text{if } 2/3 < \mu(A). \end{cases}$$

Let f be the decreasing rearrangement of g . We claim that the L^1 distance between f_A and f is at least $\delta/3$ for every A . To see this, first consider the case where $\mu(A) \leq 1/3$. Then $f_A^{-1}(\{1 - \delta\}) \subset A_1 \cup A_2$ and $\mu(f_A^{-1}(\{1 - \delta\})) = 1/3$. However, $f \geq b$ on $A_1 \cup A_2$. Hence, $\int |f - f_A| \geq \delta/3$. Now suppose $1/3 < \mu(A) < 2/3$. Then f_A takes value b or c for some subset of A_1 of measure at least $\mu(A) - 1/3$, and f_A takes value a or b for some subset of A_3 of measure at least $2/3 - \mu(A)$. Thus, $\int |f - f_A| \geq \delta(\mu(A) - 1/3) + \delta(2/3 - \mu(A)) = \delta/3$. Finally, suppose that $\mu(A) \geq 2/3$. Then f_A equals b or c on A_1 . Since f is equal to a on this set, we again have that $\int |f - f_A| \geq \delta/3$.

We therefore conclude that f is not in m , that is, the rearrangement of g is omitted from m . Hence, m is not symmetric. That the upper probability can be symmetric without the upper expectation being symmetric emphasizes, in an explicit way, that upper expectations are not determined by their upper probabilities in contrast to the relationship that holds between probability and expectation. This issue has nothing to do with regularity conditions since similar examples may be constructed on finite sets. It is easy to see that weak symmetry implies that the Choquet integral is symmetric. Thus, we have shown that symmetry of the Choquet integral does not imply symmetry of the upper expectation functional.

8. Discussion. By studying the symmetric case we feel we have shed some light on the structure of upper probabilities. More work will be needed to see how these results can be carried over to the nonsymmetric case. Also, it would be interesting to investigate other types of invariance.

We restricted ourselves to the nonatomic case in this articles, but many examples in robustness have measures with singular components. The ϵ -contaminated neighborhoods studied by Huber and Strassen (1973) and Berger (1984) are of this type. It should be possible to extend our results to that setting.

We have emphasized upper expectations because they are quantities of direct interest in robust statistics and they are fundamental in generalizations of the betting approach to probability [see Walley (1981, 1991) and Williams (1976)]. Choquet integrals seem to have attracted more attention than upper expectations. Armstrong [(1990), Section 9] contains some results on symmetric Choquet integrals. Also, see Talagrand (1978). As we pointed out in Section 7, symmetry of the Choquet integral does not imply symmetry of the upper expectation functional.

Acknowledgment. The authors thank the referee for helpful comments.

REFERENCES

- ARMSTRONG, T. (1990). Comonotonicity, simplicial subdivision of cubes and non-linear expected utility via Choquet integrals. Technical report, Dept. Mathematics and Statistics, Univ. Maryland.
- BEDNARSKI, T. (1981). On solutions of minimax test problems for special capacities. *Z. Wahrsch. Verw. Gebiete* **58** 397–405.
- BEDNARSKI, T. (1982). Binary experiments, minimax tests and 2-alternating capacities. *Ann. Statist.* **10** 226–232.
- BERGER, J. O. (1984). The robust Bayesian viewpoint (with discussion). In *Robustness of Bayesian Analysis* (J. B. Kadane, ed.) 63–144. North-Holland, Amsterdam.
- BERGER, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer, New York.
- BERGER, J. O. (1990). Robust Bayesian analysis: Sensitivity to the prior. *J. Statist. Plann. Inference* **25** 303–328.
- BUJA, A. (1984). Simultaneously least favourable experiments. I: Upper standard functionals and sufficiency. *Z. Wahrsch. Verw. Gebiete* **65** 367–384.
- BUJA, A. (1985). Simultaneously least favourable experiments. II: Upper standard loss functions and their applications. *Z. Wahrsch. Verw. Gebiete* **69** 387–420.
- BUJA, A. (1986). On the Huber–Strassen theorem. *Probab. Theory Related Fields* **73** 367–384.
- CHOQUET, G. (1953). Theory of capacities. *Ann. Inst. Fourier (Grenoble)* **5** 131–295.
- DEMPSTER, A. P. (1967). Upper and lower probabilities induced from a multivalued mapping. *Ann. Math. Statist.* **38** 325–339.
- DEMPSTER, A. P. (1968). A generalization of Bayesian inference (with discussion). *J. Roy. Statist. Soc. Ser. B* **30** 205–247.
- DEROBERTIS, L. and HARTIGAN, J. A. (1981). Bayesian inference using intervals of measures. *Ann. Statist.* **9** 235–244.
- FINE, T. (1988). Lower probability models for uncertainty and nondeterministic processes. *J. Statist. Plann. Inference* **20** 389–411.
- 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. and FINE, T. (1987). Continuous lower probability based models for stationary processes with bounded and divergent time averages. *Ann. Probab.* **15** 783–803.
- HARDY, G. H., LITTLEWOOD, J. E. and PÓLYA, G. (1952). *Inequalities*. Cambridge Univ. Press.
- HUBER, P. J. (1973). The use of Choquet capacities in statistics. *Bull. Inst. Internat. Statist.* **45** 181–191.
- HUBER, P. J. (1981). *Robust Statistics*. Wiley, New York.
- HUBER, P. J. and STRASSEN, V. (1973). Minimax tests and the Neyman–Pearson lemma for capacities. *Ann. Statist.* **1** 251–263.
- KEMPERMAN, J. H. B. (1968). The general moment problem, a geometric approach. *Ann. Math. Statist.* **39** 93–122.
- KOOPMAN, B. O. (1940a). The bases of probability. *Bull. Amer. Math. Soc.* **46** 763–774.
- KOOPMAN, B. O. (1940b). The axioms and algebra of intuitive probability. *Ann. of Math.* **42** 169–187.
- KUMAR, A. and FINE, T. (1985). Stationary lower probabilities and unstable averages. *Z. Wahrsch. Verw. Gebiete* **69** 1–17.
- LAVINE, M. (1988). Prior influence in Bayesian statistics. Discussion paper 88-06, Inst. Statistics and Decision Sciences, Duke Univ.
- LORENTZ, G. G. (1953). *Bernstein Polynomials*. Univ. Toronto Press.

- MARSHALL, A. W. and OLKIN, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. Academic, New York.
- PAPAMARCOU, A. and FINE, T. (1986). A note on undominated lower probability. *Ann. Probab.* **14** 710–723.
- PURI, M. L. and RALESCU, D. A. (1983). Strong law of large numbers with respect to a set-valued probability measure. *Ann. Probab.* **11** 1051–1054.
- RIEDER, H. (1977). Least favourable pairs for special capacities. *Ann. Statist.* **5** 909–921.
- ROBERTS, A. W. and VARBERG, D. E. (1973). *Convex Functions*. Academic, New York.
- RYFF, J. (1963). On the representation of doubly stochastic operators. *Pacific J. Math.* **13** 1379–1386.
- RYFF, J. (1965). Orbits of L^1 -functions under doubly stochastic transformations. *Trans. Amer. Math. Soc.* **117** 92–100.
- RYFF, J. (1967). On Muirhead's theorem. *Pacific J. Math.* **21** 567–576.
- RYFF, J. (1970). Measure preserving transformations and rearrangements. *J. Math. Anal. Appl.* **31** 449–458.
- SCHMEIDLER, D. (1989). Subjective probability and expected utility without additivity. *Econometrica* **57** 571–587.
- SHAFER, G. (1976). *A Mathematical Theory of Evidence*. Princeton Univ. Press.
- SHAFER, G. (1979). Allocations of probability. *Ann. Probab.* **7** 827–839.
- SMITH, C. A. B. (1961). Consistency in statistical inference and decision (with discussion). *J. Roy. Statist. Soc. Ser. B* **23** 1–37.
- SUPPES, P. and ZANOTTI, M. (1989). Conditions on upper and lower probabilities to imply probabilities. *Erkenntnis* **31** 323–345.
- TALAGRAND, M. (1978). Capacités invariantes extrémales. *Ann. Inst. Fourier (Grenoble)* **28** 79–146.
- WALLEY, P. (1981). Coherent lower (and upper) probabilities. Statistics research report, Univ. Warwick, Coventry.
- WALLEY, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London.
- WALLEY, P. and FINE, T. (1982). Towards a frequentist theory of upper and lower probability. *Ann. Statist.* **10** 741–761.
- WASSERMAN, L. (1990). Prior envelopes based on belief functions. *Ann. Statist.* **18** 454–464.
- WASSERMAN, L. and KADANE, J. B. (1990a). Permutation invariant upper and lower probabilities. Technical Report 466, Dept. Statistics, Carnegie Mellon Univ.
- WASSERMAN, L. A. and KADANE, J. (1990b). Bayes' theorem for Choquet capacities. *Ann. Statist.* **18** 1328–1339.
- WILLIAMS, P. (1976). Indeterminate probabilities. In *Formal Methods in the Methodology of Empirical Sciences* (M. Przelecki, K. Szaniawski and R. Wojcicki, eds.) 229–246. Reidel, Dordrecht.
- WOLFESON, M. and FINE, T. (1982). Bayes-like decision making with upper and lower probabilities. *J. Amer. Statist. Assoc.* **77** 80–88.

DEPARTMENT OF STATISTICS
232 BAKER HALL
CARNEGIE MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213