PAYOFFS OF PROBABILITY FORECASTERS AND A THEOREM OF MCCARTHY*

by

Arlo D. Hendrickson

Technical Report No. 141

University of Minnesota
Minneapolis, Minnesota

June 1970

*This research was supported by National Science Foundation Grant GP-9556.

1. Introduction.

I. J. Good (1952), in a section entitled "fair fees," raised the question of "how a firm can encourage its experts to give fair estimates of probabilities." If E was an event whose probability was estimated to be p_1 by an expert, then Good suggested paying the expert $k \log (2p_1)$ if E occurred and $k \log (2 - 2p_1)$ if E did not occur. This payoff function had the "desirable property that its expectation is maximized when $p_1 = p$, the true probability of E."

McCarthy (1956) generalized Good's problem from the case of two possible outcomes to that of n outcomes whose probabilities were to be estimated by a forecaster. A forecaster was to be paid a payoff $f_k(q)$ if the k^{th} event occurred where $q = (q_1, q_2, \ldots, q_n)$ was the forecaster's estimate of the true probability vector $p = (p_1, p_2, \ldots, p_n)$. McCarthy defined a payoff rule which "keeps the forecaster honest" to be a rule such that "regardless of the value of p, the forecaster's expectation is maximized if and only if he puts $q = p$." In this paper we will call such functions "payoff functions which encourage honesty." We will also consider payoff functions which satisfy the less restrictive condition that the forecaster's expectation is maximized at $q = p$ and possibly at other values of q. At least this condition does not discourage honesty; although it may not provide incentive for collecting more evidence.

Several authors have studied the properties of functions which encourage honesty. The most general theorem seems to be McCarthy's (1956) in which he gives necessary and sufficient conditions in the finite discrete case. He omitted the proof. McCarthy's theorem can be interpreted as follows: A random variable $f(p) = (f_1(p), f_2(p), \dots, f_n(p))$ keeps the forecaster honest iff $f_k(p) = \frac{\partial H}{\partial p_k}(p)$ where H is a convex function which is homogeneous of the first degree. The function H is the maximum expectation function

$$H(p) = \sum_{k=1}^{n} p_k f_k(p).$$

McCarthy's theorem needs a slight modification to be correct.

Aczel and Pfanzagl (1965) obtained the interesting result that in the finite discrete case with $n > 2$ outcomes, the logarithmic payoff, $A \log p_k + B$, suggested by Good, is the only function which both encourages honesty and has the property that the payoff for the occurrence of the k^{th} event depends only on the estimated probability p_k of that event. They give more general solutions for $n = 2$.

Marschak (1960) in his comments on McCarthy's paper noted the distinction between functions which are "expected costs to the client" to provide incentive for honest appraisals by the expert and functions which are "a good measure of worth to the client to be given these probabilities." However, although the two measures may be different, according to McCarthy's theorem the former is restricted only to be convex (because any function which is convex on $\{(p_1, p_2, \dots, p_n): \sum_{k=1}^{n} p_k = 1, p_k \geq 0 \text{ for all } k\}$ can be extended uniquely

to a homogeneous and convex function on the Euclidean space R^n).
Thus if the second measure is restricted to be convex, then by an appropriate choice of the payoff function, the two measures of worth and cost could be made equal.

We quote from McCarthy (1956) the intuitive concept of the convexity restriction on the maximum expected payoff function H:

> The intuitive concept of the convexity restriction is that it is always a good idea to look at the outcome of an experiment if it is free. For suppose that the experiment has two outcomes, A_* and A^*, which would give one probabilities p and p^* for the event in question. Let t be the probability that A is the outcome. If we decide not to look, our expectation is $H(tp + (1-t)p^*)$, while if we decide to look, our expectation is $tH(p) + (1-t)H(p^*)$.

The present paper generalizes McCarthy's theorem by using convex analysis.

2. Payoff Functions.

Let $(\mathfrak{X}, \mathfrak{A}, \mu)$ be a measure space and let $\mathcal{P}$ be a set of densities on $(\mathfrak{X}, \mathfrak{A})$ with respect to μ. Let $\mathfrak{L}$ be the set of real-valued random variables on $(\mathfrak{X}, \mathfrak{A})$. We will define a "payoff function" to mean any function f which maps $\mathcal{P}$ into $\mathfrak{L}$. Hence the value of $f(p)$ will be dependent upon the value of p and the outcome of the experiment, $\omega \in \mathfrak{X}$.

If a payoff is defined to be a real-valued random variable on $(\mathfrak{X}, \mathfrak{A})$, then the function $f(p_1)$ can be considered as the payoff given to an assessor for estimating p_1 as the true probability density. The assessor's choice $p_1 \in \mathcal{P}$ may actually be presented as an infinite sequence of choices and f may be an infinite series of increments in payoffs depending only on previous choices, as in Hendrickson and Buehler (1970). These increments in payoffs may arise from bets, they may be awards, or penalties, and may have known distributions or unknown distributions depending on p.

Let us assume that a person, when given the choice among the random payoffs in the range of f, will select a random payoff which he believes to have maximum expectation. Then a payoff function will keep that person honest if when offered $f(p_1)$ for telling us p_1, he must tell us the true density p in order that he maximizes his expectation, $E(f(p_1)|p)$. Hence, we will say a payoff function encourages honesty if for $p, p_1 \in \mathcal{P}$ such that $p_1 \neq p$ a.e. μ,

$$(1) \qquad E(f(p)|p) > E(f(p_1)|p).$$

We will also consider characterizations of f which satisfy the less restrictive inequality

$$(2) \qquad E(f(p)|p) \geq E(f(p_1)|p) \quad \text{for all } p_1, p \in \mathcal{P}.$$

If (2) is satisfied, then in order to maximize his expectation with respect to p, an assessor must choose his assessment of p from the class

$$(3) \quad B(p) = \{p_1 \colon E(f(p)|p) = E(f(p_1)|p)\}.$$

Thus, if an assessor chooses p_1 for the questioned density and is allowed to be dishonest, then (2) implies the assessed value of p is a member of

(4) $$A(p_1) = \{p\colon E(f(p)|p) = E(f(p_1)|p)\}.$$

If (2) holds, it will be shown in Section 6 that $A(p_1)$ is a convex set.

Example 1.

Let $P(E) = p$ be unknown and let $P(E_1) = \frac{1}{2}$. Define the payoff function f by

$$\text{if } p \geq \frac{1}{2} \text{ then } f(p) = \begin{cases} 1 & \text{if } E \text{ occurs} \\ 0 & \text{if } E^c \text{ occurs} \end{cases}$$

$$\text{if } p < \frac{1}{2} \text{ then } f(p) = \begin{cases} 1 & \text{if } E_1 \text{ occurs} \\ 0 & \text{if } E_1^c \text{ occurs.} \end{cases}$$

Then f satisfies condition (2). If $\mathcal{P} = [0, 1]$ then $A(\frac{1}{4}) = [0, \frac{1}{2})$ and $A(\frac{3}{4}) = [\frac{1}{2}, 1]$. See Hendrickson and Buehler (1970).

Example 2.

Let $\|p\|^2 = \int p^2(x)d\mu(x) < \infty$ for all $p \in \mathcal{P}$. Define the random variable $f(p)$ by

$$f(p)(x) = \frac{p(x)}{\|p\|} \text{ for all } x \in \mathcal{X},\ p \in \mathcal{P}.$$

Then condition (1) is satisfied by the Cauchy-Schwarz inequality.

3. $\mathcal{P}$ as a Convex Subset of a Hilbert Space.

The space $\mathcal{L}$ of random variables is a vector space. Define the norm of p by

(5) $$\|p\|^2 = \int p^2(x)d\mu(x).$$

Let $\mathcal{L}_2(\mu)$ be the space of all p where $\|p\| < \infty$. Define an inner product on $\mathcal{L}_2(\mu)$ by

(6) $$< p, q > = \int p(x)q(x)d\mu(x).$$

The space $\mathcal{L}_2(\mu)$ together with the inner product given by (6) is a Hilbert space (Halmos (1957), §9). The following relation is crucial in applying convex analysis to conditions (1) and (2):

(7) $$< p, q > = E(q|p) \quad \text{whenever} \quad p \in \mathcal{P}.$$

To avoid difficulties in (1), (2), and (7) we will assume that the range of f is contained in the set

(8) $$\mathcal{L}_1 = \mathcal{L}_1(\mathcal{P}) = \{q \in \mathcal{L}\colon E(q|p) \text{ exists and is finite for all } p \in \mathcal{P}\}.$$

An important function related to the payoff function $f(p) \in \mathcal{L}_1$ is the expected payoff function defined on $\mathcal{P}$ by

(9) $$H(p) = < f(p), p > .$$

In terms of H, conditions (1) and (2) become, respectively,

(10) $$H(p) > (< p, f(q) >) \text{ if } p \neq q \text{ a.e. } \mu$$

and

(11) $$H(p) \geq (< p, f(q) >) \text{ for all } p, q \in \mathcal{P}.$$

Conditions (9), (10), and (11) can be expressed in terms of useful concepts found in the theory of convex analysis.

4. Review of Some Concepts of Convex Analysis.

Rockafellar (1970) gave definitions and theorems about convex functions and subgradients for the special case that the Hilbert space $\mathcal{X}$ is Euclidean. These concepts are sufficient for the finite discrete case of McCarthy's theorem. We have utilized some of Rockafellar's definitions and theorems for more general spaces, when they apply. Rather than prove that convex functions are continuous on the interior of their domains we must assume it. Theorem 5 states that the graph of any continuous convex function H over a nonempty open domain corresponds to the boundary of a convex set whose supporting hyperplanes correspond to subgradients of H. When $\mathcal{X}$ is Euclidean, Theorem 5 is implied by Theorem 23.4 of Rockafellar. Theorem 4 is given because, unlike Rockafellar, we have not assumed convexity as part of the definition of subgradient. The theory of convex sets in more general spaces is taken from Valentine (1964), Halmos (1957), and Köthe (1969).

Let $\mathcal{X}$ denote any Hilbert space. Then the space $\mathcal{X} \times R$ is a Hilbert space with inner product given by

$$\text{(12)} \qquad \langle (p, \alpha), (q, \beta) \rangle = \langle p, q \rangle + \alpha\beta .$$

The <u>epigraph</u> of a real-valued function H on a convex subset C of $\mathcal{X}$ will be denoted by epi (H) and defined by

$$\text{epi}(H) = \{(p, \alpha): p \in C, \alpha \in R, \alpha \geq H(p)\}.$$

H is a convex function iff epi (H) is a convex set. For this reason, the theory of convex sets, in Valentine (1964) for example, can be applied to convex functions. It is clear that the graph of H is contained in bdry(epi (H)) where the topology is given by the norm

(13) $\|(p, \alpha)\|^2 = \|p\|^2 + \alpha^2.$

Generalized concepts of tangency are those of supporting hyperplanes to a set in a topological linear space $\mathcal{L}$ and subgradients of functions on a Hilbert space $\mathcal{H}$. A hyperplane is a set $\{p \in \mathcal{L}: h(p) = \alpha\}$ where h is a linear functional on $\mathcal{L}$, $h \not\equiv 0$. The hyperplane $\{p \in \mathcal{L}: h(p) = \alpha\}$ supports a set C at $x_0 \in C$ if $h(x) \geq \alpha$ for all $x \in C$ and $h(x_0) = \alpha$. Theorem 1 below is implied by Theorems 2.15 and 4.1 of Valentine (1964). Valentine's Theorems 2.8 and 2.15 are incorrect since he does not state the interior of the convex set must be nonempty. Valentine's proofs however are valid since they depend on the correctly stated Theorem 2.7. An example of a convex set in which every point is a boundary point, yet no point except 0 has a hyperplane of support through it, is the space $\mathcal{L}_2^+(\mu)$ of Example 7, Section 5.

Theorem 1.

If C is a convex subset of a topological linear space $\mathcal{L}$ and if the interior of C is nonempty, then through each of its boundary points there passes a closed hyperplane of support. Conversely, if C is closed, if the interior of C is nonempty, and if through each of its boundary points there passes a plane of support, then C is convex.

If $\mathcal{L} = \mathcal{H}$ is a Hilbert space, then the hyperplanes of support of Theorem 1 can be characterized by using the Riesz representation theorem. We show this in Theorem 2.

Theorem 2.

If $\{p \in \mathcal{H}: h(p) = \alpha\}$ is a closed hyperplane contained in a Hilbert space $\mathcal{H}$, then there exists $q^* \in \mathcal{H}$ such that $h(p) = \langle p, q^* \rangle$ for all $p \in \mathcal{H}$.

Proof:

The set $\{p \in \mathcal{H}: h(p) = \alpha\}$ is closed iff the linear functional h is continuous on $\mathcal{H}$ with $h \not\equiv 0$ (Theorem 2.12, Valentine (1964)). Continuity of h on the compact set $\{p: \|p\| \leq 1, p \in H\}$ implies there exists $M > 0$ such that $h(\frac{p}{\|p\|}) \leq M$. Hence by the Riesz representation theorem (page 31, Halmos (1957)), there exists $q^* \in \mathcal{H}$ such that $h(p) = \langle p, q^* \rangle$ for all $p \in \mathcal{H}$. □

The supporting hyperplanes in $\mathcal{H} \times R$, with inner product given by (12), are of the form $\{(p, \alpha): h_1(p, 0) + h_1(0, \alpha) = \beta\}$ where h_1 is a linear functional on $\mathcal{H} \times R$. These hyperplanes are of two forms:

(i) $\{(p, \alpha): \alpha = \beta - h(p)\}$

(ii) $\{(p, \alpha): h(p) = \beta, \alpha \in R\}$.

The first set is a nonvertical hyperplane and the second is vertical. If H is convex on C then the supporting hyperplane at $(q, H(q))$ of epi (H) is seen to satisfy either

(14) $$H(p) \geq h(p - q) + H(q) \quad p \in C$$

or

(15) $$h(p) \geq \beta \quad p \in C, \ h(q) = \beta$$

where h is a linear functional on $\mathcal{H}$. (15) implies the only vertical supporting hyperplanes are on the boundary of C. The closed supporting hyperplanes in (i) which are given by linear functionals of the form $h(p) = < p, q^* >$ are of interest in this paper. The point q^* is a generalization of the gradient of H in R^n. We now generalize Rockafellar's definition of subgradient to the infinite-dimensional case.

<u>Definition 1</u>.

If H is defined on a convex set C contained in a vector space $\mathcal{L}$ and if there exists $q \in C$ and $q^* \in \mathcal{L}$ such that the inner product $< p, q^* >$ is defined for all $p \in C$ and

(16) $$H(p) \geq (< p - q, q^* >) + H(q) \quad \text{for all } p \in C$$

then q^* is a <u>subgradient</u> of H at q.

The following theorem shows that the subgradient is a generalization of the gradient when H is convex.

<u>Theorem 3</u>. (Theorem 25.1, Rockafellar (1970)).

Let H be a convex function on a convex set $C \subset R^n$. If H is differentiable at q, then $\nabla H(q)$ is the unique subgradient of H at q, so in particular

$$H(p) \geq (< p - q, \nabla H(q) >) + H(q) \quad \text{for all } p \in C.$$

Conversely, if H has a unique subgradient at q, then H is differentiable at q.

The following shows that the class of convex functions contains the functions which are "subdifferentiable" at each point in their domain.

Theorem 4.

If H has a subgradient q^* at each point q in a convex set C, then H is convex on C.

Proof:

Let $p \in C$, $q \in C$ and define $p_1 = (1-\lambda)p + \lambda q$. Let p_1^* be the subgradient of H at p_1. Then

$$H(p) \geq (< p - p_1, p_1^* >) + H(p_1)$$

and

$$H(q) \geq (< q - p_1, p_1^* >) + H(p_1).$$

Hence,

$$(1-\lambda)H(p) + \lambda H(q) \geq (< p_1 - p_1, p_1^* >) + H(p_1) = H(p_1). \square$$

The converse of Theorem 4 is not true in general. However, the converse is true whenever epi(H) contains an open set in $\mathscr{H} \times R$. This is given in Theorem 5. Lemma 1 restates the condition on epi(H) in terms of continuity. The proof follows directly from the definition of continuity.

Lemma 1.

Let C be a convex set in $\mathscr{H}$ whose interior is nonempty. Let H be a convex function on C which is continuous at a point $p \in \text{int}(C)$. Then epi (H) has a nonempty interior.

Theorem 5.

If C and H satisfy Lemma 1 then H has a subgradient $q^* \in \mathscr{H}$ at each point $q \in \text{int}(C)$.

Proof:

Theorems 1 and 2 imply that at every point $q \in C$ there exists a $q^* \in \mathscr{H}$ such that either (14) or (15) hold, where

$h(p) = < p, q^* >$. If $q \in \text{int}(C)$, then (15) cannot hold, so (14) holds and q^* is a subgradient of C. □

In Section 5 we give an example of a convex function H which has no subgradient at any point (Example 5), and an example of a continuous convex function whose subgradients are not members of $\mathcal{V}$ and $\text{int}(C) = \emptyset$ (Example 7). We also give an example where H does have a subgradient in $\mathcal{V}$ at each point in C, but H is not continuous (Example 6). In all the examples given, H is homogeneous.

If H is convex and homogeneous on a convex set C, then by putting $H(\lambda p) = \lambda H(p)$ the domain of H can always be extended to the <u>convex cone</u>

$$D = \{\lambda p: p \in C, \lambda > 0\}.$$

Thus without loss of generality we can assume H is defined on a convex cone.

<u>Lemma 2</u>.

If H is homogeneous on a convex cone D and if for each $q \in D$ there exists a subgradient $q^* \in \mathcal{V}$, then $H(q) = < q, q^* >$ for all $q \in D$.

<u>Proof</u>:

By letting $p_1 = \lambda p$, condition (16) and homogeneity imply

(17) $$H(p_1) \geq (< p_1 - \lambda q, q^* >) + \lambda H(q) \quad \text{for all } q, p_1 \in D.$$

Also condition (16) implies

(18) $$H(\lambda p) \geq (< \lambda p - q, q^* >) + H(q) \quad \text{for all } p, q \in D.$$

Taking the limit as $\lambda \to 0$ in (17) and (18) and letting $p_1 = q$ we obtain $H(q) = \langle q, q^* \rangle$ for all $q \in \mathcal{X}$. □

Definition 2.

H is said to be strictly convex on a convex set C if

(19) $$H((1-\lambda)p + \lambda q) < (1-\lambda)H(p) + \lambda H(q)$$

whenever $p \neq q$, $p \in C$, $q \in C$ and $0 < \lambda < 1$.

The following theorem is similar to page 94 of Valentine (1964).

Theorem 6.

The following statements are equivalent if H has a subgradient at each point in a convex set C:

(i) H is strictly convex;

(ii) H is nonlinear between any two distinct points in C;

(iii) Each nonvertical supporting hyperplane intersects epi (H) at exactly one point;

(iv) If $p \in C$ and $q \in C$ and $p \neq q$ then no subgradient of H at p is equal to a subgradient of H at q.

Proof:

(i) $\to$ (ii) $\longleftrightarrow$ (iii) $\longleftrightarrow$ (iv) are obvious. (ii) $\to$ (i) can be easily proved from the definition of convexity. □

The theory of this section is enough to prove the more general version of McCarthy's theorem, which we present in the next section. Some other concepts in convex analysis which can be applied to payoff functions will be given in Section 6.

5. McCarthy's Theorem.

McCarthy stated his theorem for the special case that $\mathcal{X} = R^n$ and $\mathcal{P}$ is the set of all n-dimensional probability vectors. The following is a restatement of McCarthy's theorem using our notation.

Theorem.

A payoff rule f satisfies (1) for all $p, p_1 \in \mathcal{P}$ iff there exists a homogeneous, convex function H such that $f(p)$ is the gradient of H at p. The function H satisfies (9).

The theorem is incorrect in that either condition (1) should be replaced by (2) or the convexity of H should be replaced by strict convexity. Also, f need not be continuous and H does not necessarily have a gradient at each point $p \in \mathcal{P}$. However, McCarthy does state that "the derivative has to be taken in a suitably generalized sense." Examples 3 and 4 are given below to illustrate McCarthy's conditions.

Example 3.

Let $f(p) = (c_1, c_2, \ldots, c_n)$. Then $H(p) = \sum_{k=1}^{n} c_k p_k$ satisfies McCarthy's theorem. Condition (2) holds but condition (1) doesn't because the payoff is independent of the forecaster's estimate.

Example 4.

Let $\mathcal{X} = \{E, E^c\}$ and $p = (p_1, p_2)$ where $p_2 = 1 - p_1$.

Define

$$f(p) = (f_1(p), f_2(p)) = \begin{cases} (1, 0) & \text{if } p_1 > \frac{1}{2} \\ (0, 1) & \text{if } p_1 \leq \frac{1}{2} . \end{cases}$$

Then $H(p) = \langle p, f(p) \rangle = \max \{p_1, p_2\}$ is convex and homogeneous. However, H is not differentiable at $p_1 = p_2$. There is some leeway in defining $f_k(p)$ at $p_1 = p_2 = \frac{1}{2}$; $f_k(p)$ need only satisfy $0 \leq f_1(p) \leq 1$, $f_1(p) + f_2(p) = 1$ at $p_1 = \frac{1}{2}$.

Examples of strictly convex functions which are not differentiable everywhere can be found by considering functions of the form $H(p) = \max \{H_1(p), H_2(p)\}$ where the $H_k(p)$ are strictly convex on $\mathcal{P}$.

A corrected and generalized version of McCarthy's theorem is given below. Again, we assume $\mathcal{L}_1$ and $\mathcal{P}$ are subsets of $\mathcal{L}$ in which the inner product (expectation) $< p, q^* >$ is defined if $p \in \mathcal{P}$ and $q^* \in \mathcal{L}_1$.

Theorem 7.

A payoff rule f mapping $\mathcal{P}$ into $\mathcal{L}_1$ satisfies condition (2) [or condition (1)] iff there exists a homogeneous and convex [or strictly convex on $\mathcal{P}$] function H defined on the convex cone $D = \{\lambda p: p \in \mathcal{P}, \lambda > 0\}$ such that $f(p)$ is the subgradient of H at p for all $p \in \mathcal{P}$. The function H satisfies condition (9).

Proof:

Assume f satisfies (2), and define $H(\lambda p) = < \lambda p, f(p) >$ for $p \in \mathcal{P}, \lambda > 0$. Then $f(q)$ is a subgradient of H at q (condition (11)). If f satisfies (1) then no subgradient of H at p is a subgradient of H at q (condition (10)). Apply Theorem 4 for the convexity of H and Theorem 6 for strict convexity.

Conversely, if H has $f(q)$ as a subgradient at q, for each $q \in \mathcal{P}$, and if H is homogeneous and convex, then applying Lemma 2 and Definition 1, we obtain condition (2). Condition (iii) of Theorem 6 implies (1). □

When condition (1) does hold, H is strictly convex not on D but on $\mathcal{P}$. In general, the notions of strict convexity and homogeneity are contradictory.

It might be asked what class of functions H satisfy the conditions of Theorem 7. Although every function which is convex on $\mathcal{P}$ can be extended to a homogeneous and convex function on the cone D of Theorem 7, every such function does not satisfy the additional requirement of having subgradients at each point $q \in \mathcal{P}$. When $\mathcal{L} = R^n$, this additional requirement is met on the interior of $\mathcal{P}$. We will prove a more general result for $\mathcal{P} \subset \mathcal{L}_2(\mu)$ (Theorem 8). The following example shows that H must be restricted.

Example 5.

Let $\mathcal{P}$ be the class of continuous, bounded densities ($\sup_x p(x) < \infty$) on $(R, \mathcal{B}, \mu)$ where μ is Lebesgue measure and $\mathcal{B}$ consists of the Borel sets. Define $H(p) = \sup_x p(x)$. Then H is clearly convex on $\mathcal{P}$. However, H is neither continuous at any $p \in \mathcal{P}$ (with respect to $\|p\|$) nor does H have a subgradient for any $p \in \mathcal{P}$.

For the remainder of this section, the Hilbert space $\mathcal{H}$ can be taken to be either $\mathcal{L}_2(\mu)$ or the smallest closed subspace of $\mathcal{L}_2(\mu)$ containing $\mathcal{P}$, where $\mathcal{P} \subset \mathcal{L}_2(\mu)$. According to Theorem 8 below, the functions satisfying Theorem 8 are contained in the class of functions which are maximum expectations of payoff functions which encourage honesty.

Theorem 8.

If the set of densities $D \subset \mathcal{X}$ is a convex set and if H is convex and homogeneous on D and continuous at a point p in the interior of D, then there exists f such that conditions (9) and (11) hold on the interior of D. The range of f may be taken in $\mathcal{X}$.

Proof:

Whenever $p \in \text{int}(P)$, apply Theorem 5 and let $f(p)$ be a subgradient of H at p. The proof follows from Lemma 2. □

The following theorem gives equivalent conditions on f for the conditions on H in Theorem 8.

Theorem 9.

If D is a convex cone in $\mathcal{X}$ whose interior is nonempty and if H and f satisfy (9) and (11) on D, then H is continuous at $p \in \text{int}(D)$ iff there exists a neighborhood of p on which $\|f(\cdot)\|$ is bounded.

Proof:

Let $p_n \in D$, $\|p_n - p\| \to 0$ as $n \to \infty$. Let $q_n = \dfrac{f(p_n)}{\|f(p_n)\|^2}$.

Then

$$H(p_n + q_n) \geq (< p_n + q_n, f(p_n) >)$$

$$= H(p_n) + < q_n, f(p_n) >$$

$$= H(p_n) + 1.$$

Thus, if H is continuous at p, we cannot have $\|q_n\| \to 0$ or $\|f(p_n)\| \to \infty$. Hence $\|f(\cdot)\|$ is bounded on a neighborhood of p.

If $\|f(\cdot)\|$ is bounded on a neighborhood of p, then by the Cauchy-Schwarz inequality,

(20) $$< p_n - p, f(p_n) > \to 0 \quad \text{if} \quad \|p_n - p\| \to 0.$$

(9) and (11) imply

(21) $$H(p_n) \geq (< p_n, f(p) >)$$

and

(22) $$H(p) \geq (< p, f(p_n) >).$$

(9) and (21) imply $\underline{\lim}\, H(p_n) \geq \lim < p_n, f(p) > = H(p)$ (Halmos, 1957, §17). (9), (20), and (22) imply

$$H(p) \geq \overline{\lim}\, H(p_n).$$

Hence $H(p) = \lim_{n\to\infty} H(p_n)$ if $\|p_n - p\| \to 0$. □

The results of Theorem 9 on local conditions on H and f easily give the following results on global conditions on H and f.

Corollary 1.

If f and H satisfy (9) and (11) on an open convex cone $D \subset \mathscr{H}$ then H is continuous on D iff $\|f(\cdot)\|$ is bounded on every closed set contained in D.

Proof:

Since $f(\lambda p) = f(p)$ if $\lambda > 0$, $\|f(\cdot)\|$ is bounded on every closed set contained in D is equivalent to $\|f(\cdot)\|$ bounded on every compact set in D which in turn is equivalent to the requirement of Theorem 9 that $\|f(\cdot)\|$ be "locally bounded" at each point $p \in \text{int}(D)$. □

Theorem 10 shows that the restriction of the maximum expected payoff functions to be continuous with respect to $\|p\|$ on all of $\mathcal{L}_2(\mu)$ is merely the restriction of payoff functions to be bounded when the condition of encouraging honesty is met.

Theorem 10.

If H and f satisfy conditions (9) and (11) for all points in a Hilbert space $\mathcal{H}$, then the following are equivalent:

(i) H is continuous;

(ii) H is bounded on the sphere $\{p \in \mathcal{H}: \|p\| = 1\}$;

(iii) f is bounded.

Proof:

(ii) follows from (i) since the sphere is compact. Suppose condition (ii) holds. As a consequence of conditions (9) and (11) and the fact that $f(q) \in \mathcal{H}$ if $q \in \mathcal{H}$ we have

$$H\left(\frac{f(q)}{\|f(q)\|} \right) \geq \left(< \frac{f(q)}{\|f(q)\|}, f(q) > \right) = \|f(q)\|$$

and hence (iii) follows. Condition (iii) implies (i) by the previous corollary. □

The following example shows that the class of maximum expected payoffs, H, satisfying the conditions of Theorem 6, contains discontinuous convex functions.

Example 6.

Let $\mathcal{X} = [0, 1]$, μ be Lebesgue measure, and $\mathcal{B}$ be the Borel subsets of $\mathcal{X}$. Let $H(p)$ be the norm

(23) $$H(p) = \left[\int |p(x)|^{\alpha} d\mu(x)\right]^{1/\alpha}$$

and define the payoff f to be

$$f(p) = \left[\frac{p(x)}{H(p)} \right]^{\alpha-1} \tag{24}$$

where $\alpha > 2$. Then H satisfies (9), and Holder's inequality is equivalent to (10), where $\mathcal{P}$ is taken to be the set of densities where H is finite. The domain of H is the familiar vector space $\mathcal{L}_\alpha(\mu) = \{p: H(p) < \infty\}$, and H is the usual norm. The space $\mathcal{L}_\alpha(\mu) \subset \mathcal{L}_2(\mu)$ is not closed with respect to the norm $\|p\|$, nor is the norm $H(p)$ continuous with respect to $\|p\|$.

Example 7, below, illustrates that Theorem 8 possibly can be generalized. The function H below cannot be extended to a convex domain whose interior is nonempty without losing its property of convexity; yet H satisfies (9) and (11) and is continuous.

<u>Example 7</u>.

Let $(\mathcal{X}, \mathcal{B}, \mu)$ be the space given in Example 6. Let $\mathcal{L}_2^+ = \mathcal{L}_2^+(\mu) = \{p: p(x) \geq 0 \text{ for all } x \in \mathcal{X}, \int p^2 d\mu < \infty\}$. Define f by

$$f(p)(x) = \begin{cases} \ln\left(\dfrac{p(x)}{\int p d\mu} \right) & \text{if } p(x) \neq 0,\ p \neq 0 \text{ a.e. } \mu \\ 0 & \text{if } p(x) = 0 \text{ or } p = 0 \text{ a.e. } \mu . \end{cases}$$

Define the function H by (9). Since

$$-e^{-1} \leq x \ln x \leq x^2$$

we have $pf(p)$ dominated above by $p^2(x)$ and below by $-e^{-1}$. Therefore, since μ is finite on $\mathcal{X} = [0, 1]$, $H(p)$ is finite for all $p \in \mathcal{L}_2^+$. H and f do satisfy (9) and (11) on $\mathcal{L}_2^+$.

The function H is continuous on $\mathcal{L}_2^+$ with respect to $\|\cdot\|$. To prove this, let $p_n \in \mathcal{L}_2^+$, $p \in \mathcal{L}_2^+$ and $\|p_n - p\| \to 0$ as $n \to \infty$. Give $\varepsilon > 0$ and let $\delta > 0$ be such that $|x \ln x| < \varepsilon$ if $0 < x < \delta$ and such that $M \geq 1$ where $M = \frac{1}{\delta} \ln \frac{1}{\delta}$. Then it can easily be shown that

$$|x \ln x - y \ln y| \leq M|x^2 - y^2|$$

if either $x \geq \delta$ or $y \geq \delta$. Thus, by integrating $|p \ln p - p_n \ln p_n|$ over the set $\{x: 0 \leq x \leq 1, p(x) < \delta, p_n(x) < \delta\}$ and the set $\{x: 0 \leq x \leq 1, p(x) \geq \delta$ or $p_n(x) \geq \delta\}$ we obtain

$$|H(p) - H(p_n)| \leq 2\varepsilon + M \int_0^1 |p^2(x) - p_n^2(x)| dx.$$

Since $\lim_{n \to \infty} \int_0^1 |p^2(x) - p_n^2(x)| dx = 0$ and since $\varepsilon > 0$ is arbitrary, $\lim_{n \to \infty} H(p_n) = H(p)$.

If $p \in \mathcal{L}_2^+$, then every neighborhood of p contains functions which are negative over sets of positive measure. Hence $\mathcal{L}_2^+$ has no interior points and Theorem 7 does not apply.

Although every member of $\mathcal{L}_2^+$ is a boundary point, the space $\mathcal{L}_2^+$ has no closed hyperplanes of support through any of its nonzero members. For, if $(< \lambda q, q^* >) \geq (< q, q^* >)$ for all $\lambda > 0$ then $< q, q^* > = 0$, and if $(< p, q^* >) \geq 0$ for all $p \in \mathcal{L}_2^+$ then $q^* = 0$ a.e. μ. Thus the convex set $\mathcal{L}_2^+$ is a counterexample to the set C in Theorem 2.15 of Valentine (1964), unless we add the requirement that the interior of C be nonempty.

6. Support Functions.

The maximum expectation function H of a payoff function f which encourages honesty, which was shown in section 5 to be convex and homogeneous, can be interpreted as a support function of a certain convex set. The theory of support functions given in this section is found in Part V of Valentine (1964). We again assume $\mathcal{X}$ is a Hilbert space, usually $\mathcal{L}_2(\mu)$.

<u>Definition 3</u>.

If C is a convex set then the function H defined on the set

$$D = \{p \in \mathcal{X}: \sup_{q \in C}(< p, q >) < \infty\} \tag{25}$$

by the equation

$$H(p) = \sup_{q \in C} < p, q > \tag{26}$$

is the <u>support function</u> of C.

The following theorem is a direct result of (25) and is Theorem 5.1 of Valentine (1964).

<u>Theorem 11</u>.

The domain D of definition of a support function H is a convex cone.

Note that (26) implies H is homogeneous and convex on D. We defined H in terms of C in (26). There is a one-to-one correspondence between homogeneous and convex functions H and closed convex sets C, where the domain of H is given by (25). We will prove this in Theorem 13.

For a given homogeneous and convex function H defined on a set $D \subset \mathcal{X}$, define C by the equation

$$C = \{p^*:(< p, p^* >) \leq H(p) \quad \text{for all} \quad p \in D\}. \tag{27}$$

Then C is closed because the half-space $\{p^*:(< p, p^* >) \leq H(p)\}$ is closed for each $p \in D$. We will prove Theorem 13 from the following theorem given in Köthe (1969), §20.7, Theorem 1.

<u>Theorem 12</u>.

Let C be a closed convex subset of a locally convex space L and let K be a compact convex set which is disjoint from A. Then there exists a closed hyperplane which separates A and K strictly.

Theorem 13 specializes to Theorem 5.3 of Valentine (1964) whenever $\mathcal{H}$ is Euclidean.

<u>Theorem 13</u>.

Let C be a nonempty closed convex set contained in a Hilbert space $\mathcal{H}$ and let H be its support function with nonempty domain of definition D. Then C satisfies (27).

<u>Proof</u>:

Let C_1 be given by (27) where H is the support function of C. Then $C \subset C_1$. Suppose there exists $q \in C_1 - C$. Since $\{q\}$ is compact, Theorem 12 and Theorem 2 imply there exists $p \in \mathcal{H}$ and $\alpha \in R$ such that $(< p, q >) > \alpha$ and $(< p, p^* >) < \alpha$ for all $p^* \in C$. Equation (25) implies $p \in D$ and (26) implies $H(p) \leq \alpha < (< p, q >)$. This contradicts $q \in C_1$. Thus $C_1 = C$. □

A theorem similar to that given by Theorem 7 can be given in terms of the closed convex set C in (27). To state the theorem, we need the following two definitions.

Definition 4. (pages 15 and 100, Rockafellar (1970)).

Let C be a closed convex set. Then p is normal to C at p^* if

$$< p, p^* > = \sup_{q \in C} < p, q >.$$

Definition 5. (Definition 7.5, Valentine (1964)).

A convex set $C \subset \mathscr{N}$ is smooth if at each of its boundary points there is a unique hyperplane of support to C. A convex function H is smooth (differentiable if $\mathscr{N} = R^n$) if H has a unique subgradient at every point of its domain.

Theorem 14.

A payoff function f on a space $\mathscr{P} \subset \mathscr{N}$ encourages honesty [strictly] if and only if there exists a closed convex [smooth] set $C \subset \mathscr{N}$ such that for each $p \in \mathscr{P}$, p is normal to C at f(p). The set C may be taken to be

(28) $$C = \{q: (< p, q >) \leq (< p, f(p) >) \text{ for all } p \in \mathscr{P}\}.$$

Proof:

If f encourages honesty, define C by (28). By definition, p is normal to C at f(p). Conversely, if C is closed and convex and p is normal to C at f(p), then $f(p) \in C$ (else C is not closed) so

(29) $$(< p, f(p) >) \geq (< p, f(q) >) \text{ for all } p, q \in \mathscr{P}.$$

Equality holds in (29) for some p and $q \in \mathscr{P}$, where $p \nmid q$, iff both p and q are normal to C at f(q). In other words, C has two hyperplanes of support through $f(q)$. Thus if (28) and (29) hold then C is smooth iff strict inequality holds in (29) for all $p, q \in \mathcal{P}$, $p \neq q$. □

As in the proofs of Theorem 7 and Theorem 14, a closed convex set C is smooth iff its support function H is strictly convex. This relationship is dual: C contains a line segment on its boundary (is not strictly convex) iff there exists p which is normal to C at two distinct points $f(p)$ and p^*, and this is true iff H is not smooth (there exists p such that H has two subgradients $f(p)$ and p^* at p).

The question, about which Theorem 8 is concerned, of when a convex homogeneous function H has a gradient at each point $p \in \mathcal{P}$, can be restated as the question of when a closed convex C has a closed hyperplane of support of the form $\{p^* : < p, p^* > = \alpha\}$ for each $p \in \mathcal{P}$. The following theorem, which is similar to Theorems 7 and 8, is given in Valentine (1964), Theorem 5.2.

<u>Theorem 14.</u>

If C is a bounded closed nonempty convex set in $\mathcal{X}$, then for each $p \in \mathcal{X}$ there exists a point $p^* \in C$ such that $H(p) = < p, p^* >$ where H is the support function of C.

The following example shows that the boundedness of C in Theorem 14 is necessary, even when its support function is bounded.

<u>Example 8.</u>

Let $\mathcal{X} = R^2$. Let $C = \{(p_1, p_2): p_1 < 0,\ p_2 \leq \frac{1}{p_1}\}$. Then the boundary of C is the graph of the function $h(x) = \frac{1}{x}$ for $x < 0$.

C is closed and convex. The set D given by (25) is the set $\{p: p_1 \geq 0, p_2 \geq 0\}$. If $p \in D$ and $p_1 \neq 0, p_2 \neq 0$ then p is normal to C at the point $f(p) = (-\sqrt{\frac{p_2}{p_1}}, -\sqrt{\frac{p_1}{p_2}})$. The support function H is defined for all $p \in D$ by $H(p) = -2\sqrt{p_1 p_2}$. H is bounded, continuous and convex on the set $\mathcal{P} = \{p \in D: p_1 + p_2 = 1\}$, yet H has no subgradients at the points (0, 1) or (1, 0).

The set $A(p_1)$ given in equation (4) of Section 2 is the intersection of $\mathcal{P}$ with a normal cone, defined below.

Definition 6. (page 135, Valentine (1964)).

Let C be a closed, convex set. Let $Q(p^*)$ be the set of all points which are normal to C at p^*. Then $Q(p^*)$ is the normal cone of C at p^*.

If f satisfies (2) and C satisfies (27), then $A(p_1) = Q(f(p_1)) \cap \mathcal{P}$ for each $p \in \mathcal{P}$. According to Theorem 11.1, Valentine (1964), the normal cone $Q(p^*)$ is convex for each p^* on the boundary of C. Thus $A(p_1)$ is convex for each $p_1 \in \mathcal{P}$. Alternatively, if (9) and (11) hold, then $A(p_1)$ is the projection on $\mathcal{V}$ of the intersection of the convex set epi $(H) \subset \mathcal{V} \times R$ with its supporting hyperplane $\{(p, \alpha): \alpha = \langle p, f(p_1) \rangle\}$. Thus $A(p_1)$ is convex.

In the following example, f, H, and C satisfy conditions (9), (11), and (27). The convex cones $Q(p^*)$ are illustrated in Figure 1 when they consist of more than one point--when the boundary of C is not smooth at p^*. Figures 1 and 2 illustrate the relationships between C and f and between H and f.

Example 9.

Let $\mathscr{X} = R^2$. Let

$$C = \{p\colon 0 \le p_1 \le p_2,\ \|p\| \le 1\} \cup \{p\colon 0 \le p_2 \le p_1,\ p_1 + p_2 \le \sqrt{2}\}.$$

Then C is a closed convex set. If p is normal to C at f(p), and if p is not of the form (0, -a), (-a, 0), or (a, a) where $a \geq 0$, then f(p) is uniquely determined:

$$f(p) = \begin{cases} \dfrac{p}{\|p\|} & \text{if } 0 \le p_1 < p_2 \\ (\sqrt{2},\ 0) & \text{if } 0 < p_1,\ p_2 < p_1 \\ (0,\ 0) & \text{if } p_1 < 0,\ p_2 < 0 \\ (0,\ 1) & \text{if } p_1 \le 0,\ p_2 > 0. \end{cases}$$

The support function H of C is defined by f in equation (9):

$$H(p) = \begin{cases} \|p\| & \text{if } 0 \le p_1 \le p_2 \\ \sqrt{2}\, p_1 & \text{if } 0 \le p_1,\ p_2 \le p_1 \\ p_2 & \text{if } p_1 \le 0,\ p_2 \ge 0 \\ 0 & \text{if } p_1 \le 0,\ p_2 \le 0. \end{cases}$$

Note the dual relationship between smoothness and strict convexity of H and C. Also note that f(p) lies on the boundary and is continuous where the boundary of C is strictly convex, and constant on the cone where the boundary is not smooth (rough). The boundary of C differs from the range of f only at the points where C is not strictly convex. See Figures 1 and 2.

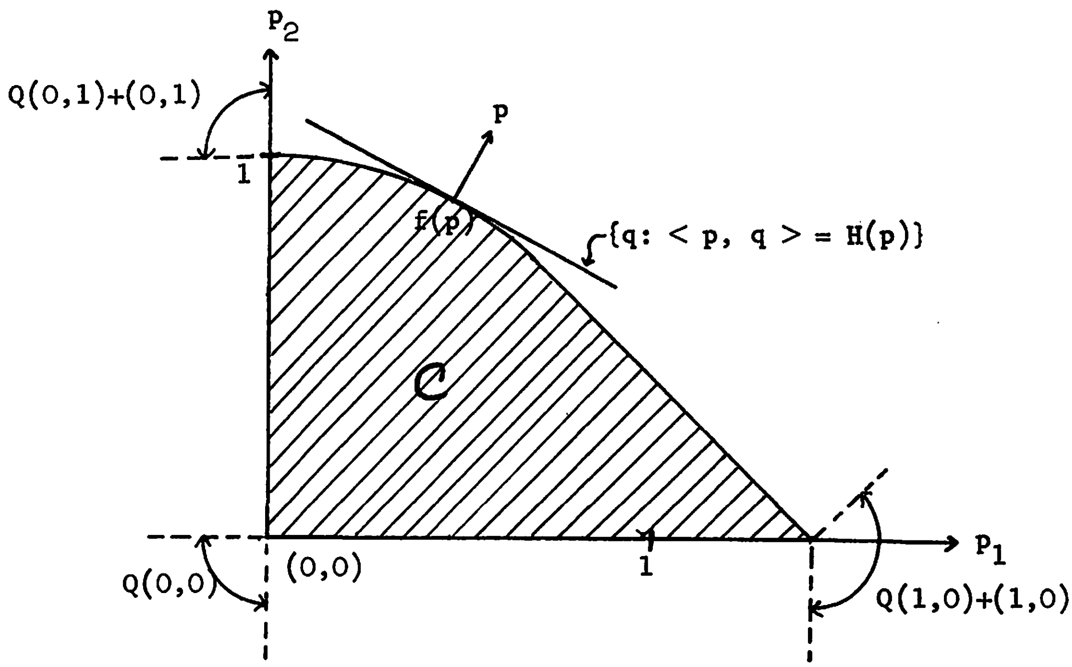

Figure 1. The convex set C of Example 9 and its cones of support.

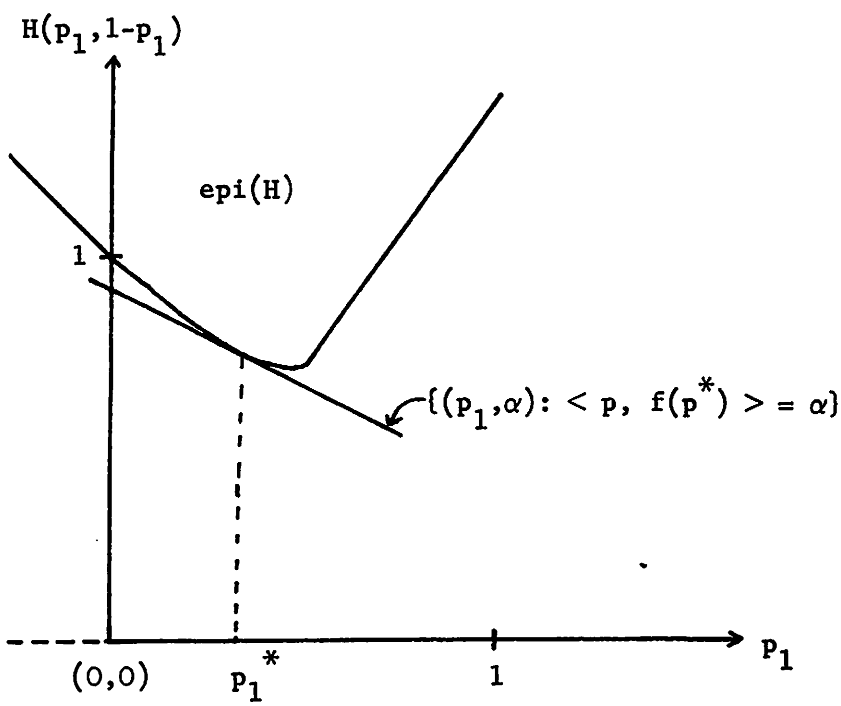

Figure 2. The support function H as a function of p_1, on the set $\{p: p_1 + p_2 = 1\}$.

7. Undominated Classes of Probability Measures.

The subject of encouraging honesty could be further generalized by taking $\mathcal{P}$ to be any set of probability measures on a space $(\mathfrak{X}, \mathfrak{a})$, not necessarily dominated, and taking $\mathcal{L}_1$ again to be the set, defined by (8), of random variables whose expectations are defined for every $P \in \mathcal{P}$.

Let f be a mapping of $\mathcal{P}$ into $\mathcal{L}_1$. We can use the same conditions (1) or (2) as before for f to encourage honesty. Let H be a mapping of $\mathcal{P}$ into R, and substitute $E(f(Q)|P)$ for the inner product in conditions (9), (10), and (11). We could make the following definition:

Definition 7.

$q^* \in \mathcal{L}_1$ is a subgradient of H at $Q \in \mathcal{P}$ if for all $p \in \mathcal{P}$,

$$H(P) \geq E(q^*|P) - E(q^*|Q) + H(Q).$$

In the present context, Theorem 7 is still true, and the proof is the same. The question of the existence of a subgradient q^* for each $Q \in \mathcal{P}$ again is a restriction on H. The class $\mathcal{P}$ no longer is a subset of $\mathcal{L}$, and $\mathcal{L}$ doesn't appear to have a topology.

REFERENCES

Aczel, J. and Pfanzagl, J. (1965). Remarks on the measurement of subjective probability and information. Metrika 11 91-105.

Good, I. J. (1952). Rational decisions. J. Royal Stat. Soc. Ser. B. 14 357-364.

Halmos, P. R. (1957). Introduction to Hilbert Space. Chelsea Publishing Co., New York.

Hendrickson, A. D. and Buehler, R. J. (1970). Some procedure for determining subjective probabilities by sequential choices. Technical Report No. 133, University of Minnesota, Minneapolis, Minnesota.

Köthe, G. (trans. D. J. H. Garling) (1969). Topological Vector Spaces I. Springer-Verlag, Berlin.

McCarthy, J. (1956). Measures of the value of information. Proc. Nat. Acad. Sci. 42 654-655.

Marschak, J. (1960). Remarks on the economics of information. Contributions to Scientific Research in Management. University of California Press, Berkeley.

Rockafellar, R. T. (1970). Convex Analysis. Princeton University Press, Princeton, New Jersey.

Valentine, F. A. (1964). Convex Sets. McGraw-Hill, New York.