THE ROLE OF SUBJECTIVE PROBABILITY AND UTILITY IN DECISION-MAKING

PATRICK SUPPES
STANFORD UNIVERSITY

1. Introduction

Although many philosophers and statisticians believe that only an objectivistic theory of probability can have serious application in the sciences, there is a growing number of physicists and statisticians, if not philosophers, who advocate a subjective theory of probability. The increasing advocacy of subjective probability is surely due to the increasing awareness that the foundations of statistics are most properly constructed on the basis of a general theory of decision-making. In a given decision situation subjective elements seem to enter in three ways: (i) in the determination of a utility function (or its negative, a loss function) on the set of possible consequences, the actual consequence being determined by the true state of nature and the decision taken; (ii) in the determination of an *a priori* probability distribution on the states of nature; (iii) in the determination of other probability distributions in the decision situation.

These subjective factors may be illustrated by a simple example. A field general knows he is faced with opposing forces which consist of either (s_1) three infantry divisions and one armored division, or (s_2) two infantry divisions and two armored divisions. Thus the possible states of nature are s_1 and s_2. The possible consequences are a tactical victory (v), a stalemate (t), and a defeat (d). He subjectively estimates utilities as follows: $u(v) = 3, u(t) = 2, u(d) = -1$. On the basis of his intelligence he subjectively estimates the probability of s_1 as $\frac{1}{3}$, and of s_2 as $\frac{2}{3}$. Also in his view there are two major possible dispositions of his forces (f_1 and f_2). Using military experience and knowledge he now estimates the probability of victory, stalemate or defeat if he decides for disposition f_1 and s_1 is the true state of nature. Corresponding estimates are made for the pairs (f_1, s_2), (f_2, s_1) and (f_2, s_2). He then presumably decides on f_1 or f_2 depending on which yields the greater expected utility with respect to his estimated *a priori* distribution on s_1 and s_2.

In connection with this example, it may properly be asked why probabilities and utilities play such a prominent role in the analysis of the general's problem. The most appropriate initial answer, it seems to me, is that we expect the general's decision to be rational in some definite sense. The probabilities are measures of degree of belief, and the utilities measures of value. To be rational he should try to maximize expected value or utility with respect to his beliefs concerning the facts of the situation. The crucial problem is: what basis is there for introducing numerical probabilities and utilities? Clearly methods of measurement and a theory which will properly sustain the methods

This research was supported by the Office of Ordnance Research, U.S. Army. I am indebted to Professor Herman Rubin for a number of helpful suggestions.

are needed. Our intuitive experience is that at least in certain limited situations, like games of chance, such measurement is possible. The task for the decision theorist is to find unobjectionable postulates which will yield similar results in broader situations. It would be most unusual if any set of postulates which guaranteed formally satisfactory measures of probability and utility also was unequivocally intuitively rational. As we shall see in section 3, compromises of some sort must be reached.

Because of the many controversies concerning the nature of probability and its measurement, those most concerned with the general foundations of decision theory have abstained from using any unanalyzed numerical probabilities, and have insisted that quantitative probabilities be inferred from a pattern of qualitative decisions. A most elaborate and careful analysis of these problems is to be found in L. J. Savage's recent book, *Foundations of Statistics* [17]. The present paper gives an axiomatization of decision theory which is similar to Savage's. The summary result concerning the role of subjective probability and utility is the same: one decision is preferred to a second if and only if the expected value of the first is greater than that of the second.

The theory presented here differs from Savage's in two important respects: (i) the number of states of nature is arbitrary rather than infinite; (ii) a fifty-fifty randomization of two pure decisions is permitted; this does not presuppose a quantitative theory of probability. More detailed differences are discussed in section 3. Since the present scheme is offered as an alternative to Savage's it is perhaps worth emphasizing that the intuitive ideas at its basis were developed in collaboration with Professor Donald Davidson in the process of designing experiments to measure subjective probability and utility [5], [6]. I suspect that experimental application of Savage's approach may be more difficult. It should also be mentioned that the approach developed here goes back to the early important, unduly neglected work of Ramsey [11].

The proof of adequacy of the axioms in section 4 depends on previous work by Mrs. Muriel Winet and me [18], and unpublished results by Professor Herman Rubin [14]; it is unfortunate that Rubin's important results are still unpublished. His work differs from the present in that he assumes a quantitative theory of probability.

Finally it should be remarked that the theory developed in the present paper is presumed susceptible of either prescriptive or descriptive use.

2. Primitive and defined notions

The four primitive notions on which our axiomatic analysis of decision-making is based are very similar to the four used by Savage in [17]. Our first primitive is a set S of states of nature; the second, a set C of consequences; and the third, a set D of decision functions mapping S into C. Savage's first three primitive notions are identical. His fourth primitive is a binary relation of preference on D. In contradistinction, our fourth primitive $\geq$ is a binary relation of preference on the *Cartesian product* $D \times D$. ($D \times D$ is the set of all ordered couples (f, g) such that f and g are in D.) This apparently slight technical difference reflects the introduction of a restricted notion of randomization which does not require a quantitative concept of probability. Thus if f, g, f' and g' are in D, the intended interpretation of $(f, g) \geq (f', g')$ is that the decision-maker (weakly) prefers a half chance on f and a half chance on g to the mixed decision consisting of a half chance on f' and a half chance on g'. For application of the apparatus developed here it must be possible to find a chance event which is independent of the state of nature and which has a subjective probability of $\frac{1}{2}$ for the decision-maker.[1] In most applications of decision theory it should be relatively easy to find such a chance event, since we are usually dealing with what Savage calls small-world situations, and not the fate of the whole universe.

To illustrate the intended interpretation of our primitive notions we may consider the following example. A certain independent distributor of bread must place his order for a given day by ten o'clock of the preceding evening. His sales to independent grocers are affected by whether or not it is raining at the time of delivery, for if it is raining, the grocers tend to buy less on the reasonably well-documented evidence that they have fewer customers. On a rainy day the maximum the distributor can sell is 700 loaves; on such a day he makes less money if he has ordered more than 700 loaves. On the other hand, when the weather is fair, he can sell 900 loaves. If the simplifying assumption is made that the consequences to him of a given decision with a given state of nature (rainy or not) may be summarized simply in terms of his net profits, the situation facing him is represented in table I.

TABLE I

	d_1—buy 700 loaves	d_2—buy 800 loaves	d_3—buy 900 loaves
s_1—rain	\$21.00	\$19.00	\$17.00
s_2—no rain	21.00	24.00	26.50

The distributor's problem is to make a decision. Decision d_2 is a kind of hedge. We also permit him the hedge of randomizing fifty-fifty between two pure decisions. He may own a coin which he believes is fair, and he does not believe that flipping this coin has an effect on the weather. Thus he may choose the mixture (d_1, d_3) over d_2. On a particular morning he might prefer the possible course of action open to him as follows:

$$(1) \qquad d_1 > (d_1, d_2) > (d_1, d_3) > d_2 > (d_2, d_3) > d_3 .$$

The use of the relation $>$ in this example is made precise by two definitions. Since the mixture (f, f) in the intended interpretation just means decision (or action) f, it is natural to extend the field of $\geq$ to D.

DEFINITION 1. *$(f, g) \geq h$ if and only if $(f, g) \geq (h, h)$; $h \geq (f, g)$ if and only if $(h, h) \geq (f, g)$; and $h \geq g$ if and only if $(h, h) \geq (g, g)$.*

For α and β either mixtures or pure decisions we now define the relation $>$ of strong preference.

DEFINITION 2. *$\alpha > \beta$ if and only if $\alpha \geq \beta$ and not $\beta \geq \alpha$.*

For later work we also need the definition of equivalence in preference (that is, indifference).

DEFINITION 3. *$\alpha \sim \beta$ if and only if $\alpha \geq \beta$ and $\beta \geq \alpha$.*

For the statement of our axioms on decision-making two further definitions are needed. The first is the definition of a notion we need for the statement of the Archimedean axiom (A.7).

[1] The term "mixed decision" is used here in the very restricted sense of referring to gambles involving just this special chance event independent of the state of nature; formally such gambles are the elements of $D \times D$.

DEFINITION 4. *$(f, g)\ L\ (f,' g')$ if and only if $f \sim f'$ and $(f, g) \sim g'$.*

The Archimedean axiom makes use of *powers* of L. We have that $(f, g)\ L^2\ (f', g')$ if and only if there exist decisions f'' and g'' such that $(f, g)\ L\ (f'', g'')$ and $(f'', g'')\ L\ (f', g')$, which situation is represented in figure 1. (Note that f, f' and f'' all occupy the same position.)

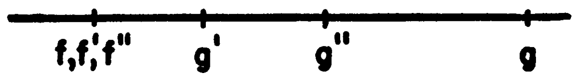

FIGURE 1

The nth power of L is defined recursively:

(1) $(f, g)\ L^1\ (f', g')$ if and only if $(f, g)\ L\ (f', g')$;

(2) $(f, g)\ L^n\ (f', g')$ if and only if there are elements f'' and g'' in D such that $(f, g)\ L^{n-1}\ (f'', g'')$ and $(f'', g'')\ L\ (f', g')$. The numerical interpretation of the relationship $(f, g)\ L^n\ (f', g')$ is that $f = f'$ and

$$\frac{2^n - 1}{2^n} f + \frac{1}{2^n} g = g' .$$

Finally, we need the notion of a *constant* decision function, that is, a function which yields the same consequence independent of the state of nature.

DEFINITION 5. *If $x \in C$ then x^* is the function mapping S into C such that for every $s \in S$, $x^*(s) = x$.*

As we shall see, the constant decisions play an all too important role in the theory developed in this paper.

3. Axioms

Using the primitive and defined notions just considered we now state our axioms for what we shall call *rational subjective choice structures.*

A system $< S, C, D \geq >$ is a RATIONAL SUBJECTIVE CHOICE STRUCTURE *if and only if the following axioms 1–11 are satisfied for every f, g, h, f', g', h', f'' and g'' in D:*

A.1 $(f, g) \geq (f', g')$ *or* $(f', g') \geq (f, g)$;

A.2 *If* $(f, g) \geq (f', g')$ *and* $(f', g') \geq (f'', g'')$ *then* $(f, g) \geq (f'', g'')$;

A.3 $(f, g) \sim (g, f)$;

A.4 $f \geq g$ *if and only if* $(f, h) \geq (g, h)$;

A.5 *If* $(f, g) \geq (f', g')$ *and* $(h, g') \geq (h', g)$ *then* $(f, h) \geq (f', h')$;

A.6 *If* $(f, g) > (f', g')$ *and* $g > g'$ *then there is an h in D such that* $g > h$ *and* $h > g'$ *and* $(f, g) \geq (f', h)$;

A.7 *If* $f > g$ *and* $f' > g'$, *then there is an h in D and a natural number n such that* $(f, g)\ L^n\ (f, h)$ *and* $(f', h) \geq (f, g')$;

A.8 *For every x in C, $x^* \in D$;*

A.9 *If for every s in S,* $(f(s)^*, g(s)^*) \geq (f'(s)^*, g'(s)^*)$, *then* $(f, g) \geq (f', g')$;

A.10 *There is an h in D such that for every s in S,* $h(s)^* \geq f(s)^*$ *and* $h(s)^* \geq g(s)^*$;

A.11 *There is an h in D such that for every s in S,* $(f(s)^*, g(s)^*) \sim h(s)^*$.

The interpretation of the first two axioms is clear: they require a simple ordering of decisions. The third axiom guarantees that our special chance event independent of the state of nature has subjective probability $\frac{1}{2}$. To see this, let $f > g$, and let E^* be our special chance event. The interpretation of (f, g) is that decision f is taken if E^* occurs and g if $\tilde{E}^*$ occurs (that is, if E^* does not occur). If the subjective probability of E^* (in symbols: $s(E^*)$) is greater than that of $\tilde{E}^*$, (f, g) will be preferred to (g, f). On the other hand, if $s(E^*) < s(\tilde{E}^*)$, then (g, f) will be preferred to (f, g). Hence, A.3 corresponds to saying that $s(E^*) = s(\tilde{E}^*) = \frac{1}{2}$. (For further discussion of this, see Davidson and Suppes [6].)

Axiom A.4 states an obvious substitution property. It is a special case ($\alpha = \frac{1}{2}$) of an axiom introduced by Friedman and Savage (see axiom P3, p. 468 [7]). It also is essentially a special case of Samuelson's strong independence axiom [16]. A kind of domination property is expressed by A.5. If the mixture (f, g) is at least as desirable as the mixture (f', g'), and h is sufficiently preferred over h' to reverse this preference in the sense that (h, g') is weakly preferred to (h', g), then it is reasonable to expect that (f, h) is weakly preferred to (f', h'). The content of this axiom is made clearer by considering particular cases among the possible orderings of the decisions. An example which brings out the implications of the axiom is given by the supposition that we have the following ordering: $f' \geq f \geq g \geq g'$. Now we must then have $h \geq h'$ since $(h, g') \geq (h', g)$; furthermore, the latter implies that the difference between h and h' is greater than between f' and f, since when h is coupled with the least desirable decision g', the mixture (h, g') is preferred to (h', g), but in the case of f', the mixing with g' leads to (f, g) being preferred. Hence, we expect to find that (f, h) is weakly preferred to (f', h'), which is what the axiom requires.

Axiom A.6 I regard as a blemish which should be eliminated or changed in form. It says nothing essentially new about the structure of any model of our axioms; just that if (f, g) is preferred to (f', g'), then we may find a decision h slightly better than g' such that we still have (f, g) preferred to the new mixture (f, h). Axiom A.7 is an Archimedean axiom of the sort necessary to get measurability. Its existence requirements are not unreasonable in view of the plenitude of decisions guaranteed by A.10 and A.11. The meaning of A.7 is very simple. No matter how great the interval between f and g, the interval may be subdivided sufficiently to find an h closer to f than g' is to f'. The axiom could be weakened by adding to the hypothesis the condition that $(f, g') \geq (f', g)$.

Axiom A.8 requires that all constant decisions, that is, decisions whose consequences are independent of the state of nature, be in D. The inclusion of such constant decisions, or of something essentially as strong, is necessary to obtain the summary result we want: f is preferred to g if and only if the expected value of f with respect to a utility function on consequences and an *a priori* distribution on states of nature is greater than the corresponding expected value of g. The inclusion of these constant decisions is not peculiar to the theory of decision-making developed here, but is also essential to Rubin's [14] and Savage's [17] theories.[2] The difficulties surrounding the inclusion of these decisions may be illustrated by considering one of Savage's colorful examples (see [17], p. 14). We have before us an egg. One of two states of nature obtains: the egg is good (s_1) or the egg is rotten (s_2). We are making an omelet and five good eggs have already been broken into the bowl. We may take one of three actions: break the egg in the bowl (f), break the egg in a saucer and inspect it (g), simply throw the egg away (h). The various consequences are easy to describe: $f(s_1)$ = six-egg omelet, $g(s_1)$ = six-egg omelet and saucer to wash, etc. But now suppose we add the constant decisions. How are we to think about

[2] An analogue of our A.8 is not included among Savage's seven axioms unless his set F of acts (corresponding to our set D of decisions) is meant to be the set of *all* functions mapping S into C, which is of course a stronger assumption than A.8. In any case it is essential to his formal developments to have such decisions at hand (see [17], from p. 25 on).

the decision which guarantees us a six-egg omelet? If the true state of nature is s_2, it is not clear that we are considering an action which makes any kind of sense. Certainly we are in no position to push the ultrabehavioristic interpretation of decision-making favored by Savage when we consider the constant decisions. I can, for instance, imagine no behavioristic evidence which would persuade me that an individual in the situation just described had chosen the constant decision guaranteeing a six-egg omelet. As far as I can see, about the most reasonable way to analyze a preference involving a constant decision such as the above one is to regard it as a nonbehavioristic subjective evaluation of consequences. Axioms A.8–A.11 have the effect intuitively of requiring such direct evaluations of consequences.

Axiom A.9 corresponds closely to Savage's seventh postulate and to Rubin's sixth axiom [14]. If for every state of nature the consequences of the mixture of decisions f and g are preferred to the consequences of the mixture of f' and g', then the mixture of f and g should be preferred to that of f' and g'. As Savage remarks, the kind of sure-thing principle expressed by this axiom is one of the most acceptable postulates of rational behavior. Axiom A.10 asserts that given any two decisions there is a third at least as good as either of the two with respect to every state of nature. This axiom is weaker than the assumption that the set of consequences of any decision f has an upper bound, that is, there is an x in C such that for every s in S, $x^* \geq f(s)^*$. It is possible that the main theorem of section 4 can be proved without this axiom, but I have not succeeded in finding such a proof.

Axiom A.11 should probably be regarded as the strongest axiom of the group. Given any two decisions f and g, A.11 asserts there is another decision h with the property that for each state of nature the consequence of h is halfway between the consequence of f and the consequence of g. This axiom may be regarded as a very strong form of Marschak's continuity axiom [10]. His axiom is that if $f > g$ and $g > h$ then there is a numerical probability α such that the mixture of f and h with probability α and $1 - \alpha$ respectively is equivalent to g. The significance of A.11 is discussed in more detail below.

Now that the analysis of individual axioms is complete, some general remarks are pertinent. Compared to Savage's axiomatization in [17], we may say of the present theory that there are more axioms but perhaps less complicated definitions. A more important kind of comparison between Savage's and the present analysis is the rather radical difference in what I like to call the *structure* axioms (as opposed to the *rationality* axioms). By and large, a structure axiom is an existential assertion.[3] Axiom A.11 is the main structure axiom in the present axiomatization. If we consider the situation facing the independent distributor of bread, which was discussed in the last section, it is clear that A.11 is not satisfied. In fact, it is easy to show that if there are two decisions, one of which is strictly preferred to the other, then A.11 and certain of the other axioms imply that there is an infinity of decisions. However, I for one am reluctant to call the distributor irrational because an insufficient number of decisions is available to him. I prefer to say that the situation the distributor is in does not permit the structure axioms to be satisfied, and hence the present theory is inapplicable; we cannot use it to decide if the distributor is regularly choosing an action or decision solely in terms of its expected value. In a given axiomatic analysis of decision-making it is not always easy or even possible clearly to separate the axioms into the two categories of rationality axioms

[3] This is certainly not always the case. The strong structure axiom in [6], which asserts that consequences are equally spared in utility, is not existential in character.

and structure axioms. Of the eleven axioms used in this paper, I would say that A.1–A.5 and A.9 are "pure" rationality axioms which should be satisfied by any rational, reflective man in a decision-making situation. On the other hand, A.8, A.10 and A.11 are "pure" structure axioms which have little directly to do with the intuitive notion of rationality. They are to be considered as axioms which impose limitations on the kind of situations to which our analysis may be applied. Axiom A.6 is a technical structure axiom which tells us little intuitively about restrictions on applicability of the theory. Without A.11, the Archimedean axiom, A.7, would need to be considered a structure axiom, but in the presence of A.11, I regard it as a rationality axiom.

Of Savage's seven postulates, two are structure axioms (P5 and P6), and the rest are rationality axioms. His P5 excludes the trivial case where all consequences are equivalent in utility and thus every decision is equivalent to every other. Postulate P6 is his powerful structure axiom corresponding to my A.11. Essentially his P6 says that if event B is less probable than event C (B and C are subsets of S, the set of states of nature), then there is a partition of S such that the union of each element of the partition with B is less probable than C. As Savage remarks, this postulate is slightly stronger than the axiom of de Finetti and Koopmans which requires the existence of a partition of S into arbitrarily many events which are equivalent in probability. Thus the consequence of his P6 is that there must be an infinity of states of nature, and as a consequence an infinity of decisions; whereas the consequence of A.11 is that there must be an infinity of decisions, with the number of states of nature wholly arbitrary. Such infinite sets, either of decisions or states of nature, can be eliminated by various kinds of special structure axioms. Davidson and I [6] eliminated them by requiring that all consequences be equally spaced in utility—an assumption which has proved manageable in some controlled experiments on decision-making at Stanford [5], but is not realistic in general.

Savage defends his P6 by holding it is workable if there is a coin which the decision-maker believes is fair for any finite sequence of flips (see p. 33 [17]). However, if the decision-maker does not believe the flipping of the coin affects what is ordinarily thought of as the state of nature, such as raining or not raining in the case of the bread distributor, then it seems to me that it is misleading to construct the states of nature around the fair coin. Once *repeated* flips of a fair coin are admitted, we can extend the single act of randomization admitted in the interpretation of the axiomatization given here, and directly introduce all numerical probabilities of the form $k/2^n$. With this apparatus available we can give an axiomatization very similar to Rubin's [14] and drop any strong structure axioms on the number of states of nature or the number of decisions.

To illustrate further the nature of the structure axiom A.11, and at the same time to argue by way of example that it does not make our theory impossible of application, I would like to modify one of Savage's finite examples (see pp. 107–108 [17]) which does not, even as modified, satisfy his P6. A man is considering buying some grapes in a grocery store. The grapes are in one of three conditions (the three states of nature): green, ripe, or rotten. The man may decide to buy any rational number of pounds between 0 and 3. If, for example, the state of nature is that the grapes are rotten and he makes the decision to buy two pounds, then the immediate consequence is possession of two pounds of rotten grapes and the loss of a certain small amount of capital. If the man is at all intuitively rational in his preferences concerning the amount of grapes to buy, it will not be hard for him to satisfy A.1–A.11—provided, of course, that he has at hand some simple random mechanism, such as a coin he believes to be fair for single tosses (he need not believe that any finite *sequence* of outcomes is as likely as any other). This example is discussed further in section 5.

By way of summary my own feeling is that Savage's postulates are perhaps esthetically more appealing than mine, but this fact is balanced by two other considerations: my axioms do not require an infinite number of states of nature, and their intuitive basis derives from ideas which have proved experimentally workable.

4. Adequacy of axioms

We now turn to the proof that our axioms for decision-making are adequate in the sense that decision f is weakly preferred to decision g if and only if the expected value of f is at least as great as the expected value of g. The actual result is not quite this strong. As might be expected, the theorem holds only for bounded decisions (precisely what is meant by a bounded decision is made clear in the statement of the theorem). On the basis of A.1–A.11 uniqueness of the *a priori* distribution on the states of nature cannot be proved, since the constant decisions alone constitute a realization of the axioms. If S is assumed finite, various conditions which guarantee uniqueness are easy to give. In stating the theorem, we use the notation: $U \circ f$ for the *composition* of the functions U and f.

THEOREM. *If* $\langle S, C, D, \geq \rangle$ *is a rational subjective choice structure, then there exists a real-value function* ϕ *on* D *such that*

(i) *for every* f, g, f' *and* g' *in* D $(f, g) \geq (f', g')$ *if and only if* $\phi(f) + \phi(g) \geq \phi(f') + \phi(g')$,

(ii) ϕ *is unique up to a linear transformation, and*

(iii) *if* U *is the function defined on* C *such that for every* x *in* C

$$U(x) = \phi(x^*), \tag{1}$$

then there exists a finitely additive probability measure P *on* S *such that for every* f *in* D *if* $U \circ f$ *is bounded, then*

$$\phi(f) = \int_S (U \circ f)(s)\, dP(s). \tag{2}$$

PROOF. The proof of (i) and (ii) follows rather easily from some previous results obtained by Mrs. Muriel Winet and me. Using a notion R of utility differences and a notion Q of preference, we established in [18] that, on the basis of axioms similar to A.1–A.7 and A.11 of this paper, there exists a real-valued function ψ unique up to a linear transformation such that

(3) $f\, Q\, g$ if and only if $\psi(f) \geq \psi(g)$,

$f, g\, R\, f', g'$ if and only if $|\psi(f) - \psi(g)| \geq |\psi(f') - \psi(g')|$.

If we introduce the two defining equivalences[4]

(4) $f\, Q\, g$ if and only if $(f, f) \geq (g, g)$;

(5) $f, g\, R\, f', g'$ if and only if either (i) $f \geq g, f' \geq g'$ and $(f, g') \geq (f', g)$, or (ii) $g \geq f$, $f' \geq g'$ and $(g, g') \geq (f, f')$, or (iii) $f \geq g$, $g' \geq f'$ and $(f, f') \geq (g, g')$, or (iv) $g \geq f$, $g' \geq f'$ and $(g, f') \geq (f, g')$,

[4] In [18] the inequalities of (3) are actually reversed, but trivial changes in the axioms given there yield (3) as a consequence.

then on the basis of A.1–A.7, A.9 and A.11 we may prove the axioms of [18] on Q and R as theorems, as well as the equivalence

(6) $(f, g') \geq (f', g)$ if and only if either (i) $f \geq g, f' \geq g'$ and $f, g\ R f', g'$, or (ii) $f \geq g$ and $g' \geq f'$, or (iii) $g \geq f, g' \geq f'$ and $f', g'\ R f, g$.

Parts (i) and (ii) of our theorem then follow immediately from the main theorem in [18].

The proof of (iii), concerning the existence of an *a priori* distribution on S, essentially uses Rubin's results in [14]. However, certain extensions of D are required in order to apply his main theorem.

By means of the utility function U on the set of consequences C, as defined in the hypothesis of (iii), we define the set F of all numerical income functions

(7) $$F = \{\rho\colon \text{ there exists } f \text{ in } D \text{ such that } \rho = U \circ f\},$$

and we define the functional η on F

(8) $$\eta(U \circ f) = \phi(f).$$

We observe first that if $\rho, \sigma \in F$, then

(9) $$\tfrac{1}{2}\rho + \tfrac{1}{2}\sigma \in F,$$

for let $\rho = U \circ f$ and $\sigma = U \circ g$, then by (i) and (ii) of our theorem, A.9, and A.11, there exists an h in D such that for every s in S

(10) $$\tfrac{1}{2}(U \circ f)(s) + \tfrac{1}{2}(U \circ g)(s) = (U \circ h)(s).$$

Hence,

(11) $$\tfrac{1}{2}\rho + \tfrac{1}{2}\sigma = U \circ h,$$

and $U \circ h$ is in F. Also, since

(12) $$\eta(U \circ h) = \phi(h) = \tfrac{1}{2}\phi(f) + \tfrac{1}{2}\phi(g) = \tfrac{1}{2}\eta(U \circ f) + \tfrac{1}{2}\eta(U \circ g),$$

we have

(13) $$\eta(\tfrac{1}{2}\rho + \tfrac{1}{2}\sigma) = \tfrac{1}{2}\eta(\rho) + \tfrac{1}{2}\eta(\sigma).$$

From (9) and (13) it easily follows that if $\rho, \sigma \in F$ and k and n are positive integers such that $k \leq 2^n$, then

(14) $$\frac{k}{2^n}\rho + \left(1 - \frac{k}{2^n}\right)\sigma \in F,$$

and

(15) $$\eta\left[\frac{k}{2^n}\rho + \left(1 - \frac{k}{2^n}\right)\sigma\right] = \frac{k}{2^n}\eta(\rho) + \left(1 - \frac{k}{2^n}\right)\eta(\sigma).$$

We now extend F by the following definition: $\rho \in \bar{F}$ if and only if there is a finite sequence $<a_1, \cdots, a_n>$ of real numbers and a finite sequence $<\rho_1, \cdots, \rho_n>$ of elements of F such that

(16) $$\rho = \sum_n a_i \rho_i.$$

(It is clear from (16) that $\bar{F}$ is a linear space.)

In order to extend η in a well-defined manner to $\bar{F}$, we need to prove that if

$$\sum_n a_i\rho_i = \sum_m b_j\sigma_j \tag{17}$$

then

$$\sum_n a_i\eta(\rho_i) = \sum_m b_j\eta(\sigma_j)\,. \tag{18}$$

Clearly without loss of generality we may assume

$$a_i,\ b_j > 0 \qquad \text{for } 1 \le i \le n,\ 1 \le j \le m\,. \tag{19}$$

We shall first establish (18) under the restriction that

$$\sum_n a_i = \sum_m b_j\,. \tag{20}$$

If a_i and b_j are rational numbers of the form $k/2^n$ with $k \le 2^n$, then (18) follows from (15) by a straightforward inductive argument (which we omit), provided

$$\sum_n a_i = \sum_m b_j = 1\,. \tag{21}$$

But the requirement of (21) is easily weakened to

$$\sum_n a_i = \sum_m b_j < 1\,, \tag{22}$$

for we may add $c\rho$ with $c = 1 - \sum_n a_i$ to both sides of (17), and then (21) will be satisfied. Furthermore, (22) is readily extended to arbitrary positive rationals, since two finite sequences of positive rationals can be reduced to (22) by multiplying through and dividing by a sufficiently high power of 2.

We are now ready to consider the case where the a_i's and b_j's are arbitrary positive real numbers. There are rational numbers r_i and s_j such that

$$r_i < a_i \qquad \text{and} \qquad s_j > b_j\,. \tag{23}$$

It is an immediate consequence of A.10 that there is a τ in F such that

$$\tau \ge \rho_i \qquad \text{and} \qquad \tau \ge \sigma_j\,. \tag{24}$$

From (23) and (24) we have, by a regrouping of coefficients

$$\sum_n r_i\rho_i + \Big[\sum_n (a_i - r_i) + \sum_m (s_j - b_j)\Big]\tau \ge \sum_m s_j\sigma_j\,. \tag{25}$$

Since the coefficient of τ is rational, we obtain by our previous results

$$\sum_n r_i\eta(\rho_i) + \lambda\eta(\tau) \ge \sum_m s_j\eta(\sigma_j)\,, \tag{26}$$

where

$$\lambda = \sum_n (a_i - r_i) + \sum_m (s_j - \beta_j)\,. \tag{27}$$

By suitable choice of the r_i's and s_j's, we may make λ arbitrarily small, and we thus infer from (23) and (26)

$$\sum_n a_i \eta(\rho_i) \geq \sum_m b_j \eta(\sigma_j) . \tag{28}$$

By an exactly similar argument, we get

$$\sum_m b_j \eta(\sigma_j) \geq \sum_n a_i \eta(\rho_i) . \tag{29}$$

To establish (18) in full generality it remains only to consider the case where

$$\sum_n a_i \neq \sum_m b_j . \tag{30}$$

Suppose, for definiteness, that

$$\sum_n a_i > \sum_m b_j . \tag{31}$$

There are elements x and y in C such that $U(x) > U(y)$ (if there are no two such elements, the proof of the whole theorem is trivial). Furthermore, in view of A.11, we may choose x and y such that $U(x) > 0$ and $U(y) > 0$, or $U(x) < 0$ and $U(y) < 0$. Let $\mu = U \circ x^*$ and $\nu = U \circ y^*$. Then μ and ν are in F, and there are nonnegative numbers a_0 and b_0 such that

$$a_0 + \sum_n a_i = b_0 + \sum_m b_j \tag{32}$$

and

$$a_0 \mu = b_0 \nu . \tag{33}$$

Then by our previous result under the restriction (20), we have

$$a_0 \eta(\mu) + \sum_n a_i \eta(\rho_i) = b_0 \eta(\nu) + \sum_m b_j \eta(\sigma_j) , \tag{34}$$

but from (33), (8) and the definition of U

$$a_0 \eta(\mu) = b_0 \eta(\nu) , \tag{35}$$

and thus

$$\sum_n a_i \eta(\rho_i) = \sum_m b_j \eta(\sigma_j) , \tag{36}$$

which establishes (18) in full generality.

On the basis of (18) we extend η to $\bar{F}$. The argument from (30) on has closely followed Rubin's proof in [14]. His proof may now be used to complete the proof of (iii). We sketch the main steps. Clearly η is a linear functional on $\bar{F}$, and it is easily shown that η is nonnegative, and hence that $\eta(\rho)$ is between $\inf_{s \in S} \rho(s)$ and $\sup_{s \in S} \rho(s)$. Let G be the space of all functions on S bounded by elements of $\bar{F}$. Then by the Hahn-Banach theorem (see pp. 27–28 [1]) η can be extended to G. Finally, it can be shown [12], [13] that such a linear functional on G is, for bounded functions in F, their expected value with respect to an *a priori* distribution on S which is in general finitely additive. (A result closely related to the existence of such a distribution is established in theorem 2.3 [19].)

5. Critical remarks

The theory of decision developed in the previous sections is no doubt defective in a number of ways, some of which I am well aware of. In this final section I briefly examine what I consider to be its gravest weakness, at least for normative applications. It is laudable to wish to base a theory of decision on behaviorally observable choices, but the decision-maker is interested in something more. He wants advice on how to choose among alternative courses of action. He wants to have at hand a theory which tells him how to use initial information. The result of the analysis in this paper and in Savage's book is that if certain structure axioms are satisfied, any rational man acts as if he had an *a priori* distribution on the states of nature. But what the rational man wants is a method for selecting that *a priori* distribution which *best* uses his *a priori* information. The present theory or Savage's offers little help on this point. The importance of this problem is testified to by the over-all situation in statistical decision theory: we have clear ideas of optimality only when given an *a priori* distribution on the states of nature. Bayesian principles of choice seem naturally to dominate the scene. (For some penetrating reasons, see chapter 4 [2].)

In recent years a serious attempt has been made by philosophical logicians to develop a theory of confirmation which is closely related to the problem under discussion. The theory of confirmation is concerned with precisely characterizing the degree to which a given hypothesis is supported by given evidence. The confirmation function which is usually introduced is very similar in its formal properties to the standard notion of conditional probability. Perhaps because the theory of confirmation has usually been stated in logical or linguistic terms, its connections with decision theory have not been made as clear as they could. Thus viewed, the purpose of confirmation theory is to develop methods for codifying prior information to yield an *a priori* distribution on the states of nature. The available evidence is our prior information, and a hypothesis corresponds to asserting that a given state of nature is the true one.

For concreteness we may consider the grape example of section 3. In Savage's discussion of this example (see p. 108 [17]) he assigns subjective probabilities to the three states of nature, and then goes on to consider what action the decision-maker should take after observing a sample of one grape. But the point at issue here is: given certain prior information is one *a priori* distribution as reasonable as any other? As far as I can see there is nothing in my or Savage's axioms which prevents an affirmative answer to this question. Yet if a man had bought grapes at this store on fifteen previous occasions and had always got green or ripe but never rotten grapes, and if he had no other information prior to sampling the grapes, I for one would regard as unreasonable an *a priori* distribution which assigned a probability of $\frac{2}{3}$ to the rotten state. Unfortunately, though I am prepared to reject this one distribution as unreasonable, I am not prepared to say what I think is optimal.

The most thoroughgoing analysis of confirmation theory has been made by Carnap [3], but his chosen confirmation function c^* is beset with many technical difficulties which give rise to counterintuitive examples (see, for example, [8], [9], [15]). Here I am not concerned to scrutinize the current problems of confirmation theory but merely to argue for the relevance of the theory to decision theory.[5] An adequate confirmation

[5] A central problem in confirmation theory is what *a priori* distribution to choose when there is no information whatsoever. Chernoff [4] has shown that if certain reasonable postulates are accepted and if the number of states of nature is finite, then the distribution to choose is that one which makes each state equally probable.

theory would not discredit the kind of axiomatization of decision-making given in this paper; it would not disturb the central role of subjective probability and utility.[6] It would stand to the theory of this paper more as statistical mechanics stands to macroscopic thermodynamics: a decision theory which included a confirmation function would have the axioms of the present paper (or of a similar theory such as Savage's) forthcoming as theorems. Such an enlarged decision theory would remain subjective but an important element of counterintuitive arbitrariness would have been eliminated.

In conclusion, I should like to acknowledge my indebtedness to Professor Herman Rubin for a number of helpful suggestions, as well as to Professor Donald Davidson, Professor Robert McNaughton and Dr. Jean Rubin for their useful comments.

REFERENCES

[1] S. Banach, *Théorie des Opérations Linéaires*, Warsaw, 1932.
[2] D. Blackwell and M. A. Girshick, *Theory of Games and Statistical Decisions*, New York, John Wiley and Sons, 1954.
[3] R. Carnap, *Logical Foundations of Probability*, Chicago, University of Chicago Press, 1950.
[4] H. Chernoff, "Rational selection of decision functions," *Econometrica*, Vol. 22 (1954), pp. 422–443.
[5] D. Davidson, S. Siegel and P. Suppes, "Some experiments and related theory on the measurement of utility and subjective probability," Stanford Value Theory Report No. 4, August, 1955.
[6] D. Davidson and P. Suppes, "A finitistic axiomatization of subjective probability and utility," to be published in *Econometrica.*
[7] M. Friedman and L. J. Savage, "The expected utility hypothesis and the measurability of utility," *Jour. of Political Economy*, Vol. 60 (1952), pp. 463–474.
[8] J. G. Kemeny, Review of [3], *Jour. of Symbolic Logic*, Vol. 16 (1951), pp. 205–207.
[9] ———, "A logical measure function," *Jour. of Symbolic Logic*, Vol. 18 (1953), pp. 289–308.
[10] J. Marschak, "Rational behavior, uncertain prospects, and measurable utility," *Econometrica*, Vol. 18 (1950), pp. 111–141.
[11] F. P. Ramsey, *The Foundations of Mathematics and Other Logical Essays*, London, Kegan Paul, 1931.
[12] H. Rubin, "An axiomatic approach to integration" (abstract), *Bull. Amer. Math. Soc.*, Vol. 55 (1949), p. 1064.
[13] ———, "Measures and axiomatically defined integrals" (abstract), *Bull. Amer. Math. Soc.*, Vol. 55 (1949), p. 1064.
[14] ———, "Postulates for rational behavior under uncertainty," unpublished.
[15] H. Rubin and P. Suppes, "A note on two-place predicates and fitting sequences of measure functions," *Jour. of Symbolic Logic*, Vol. 20 (1955), pp. 121–122.
[16] P. A. Samuelson, "Probability, utility, and the independence axiom," *Econometrica*, Vol. 20 (1952), pp. 670–678.
[17] L. J. Savage, *Foundations of Statistics*, New York, John Wiley and Sons, 1954.
[18] P. Suppes and M. Winet, "An axiomatization of utility based on the notion of utility differences," *Management Science*, Vol. 1 (1955), pp. 259–270.
[19] K. Yosida and E. Hewitt, "Finitely additive measures," *Trans. Amer. Math. Soc.*, Vol. 72 (1952), pp. 46–66.

[6] This remark is controversial. In the opinion of many competent investigators an adequate confirmation theory would dispense with any need for subjective probability. I cannot here state my reasons for disagreeing with this view.