Recent Advancement in Game Theory, Princeton University Conference, 1962, pp. 61-72.

RECENT DEVELOPMENTS IN UTILITY THEORY[1]

Patrick Suppes
Stanford University
Stanford, California

Excellent recent surveys of the subject of this paper are to be found in Adams [1959], Arrow [1958] and Edwards [1961]. In preparing this lecture, I have also drawn heavily upon a list of references Duncan Luce uses in connection with his seminar on choice and measurement.

I have organized the material I want to discuss under seven headings: empirical studies of utility measurement, empirical studies of subjective probability measurement, mathematical results in utility theory, stochastic theories of choice, derivation of utility functions from more general behavioral theories, models to replace utility, and scaling problems. What I shall have to say under these headings will probably not be as judiciously balanced as it should be but will reflect my own interests and predilections. For other viewpoints and a wider survey I urge you to refer to the articles of Adams, Arrow and Edwards.

1. Empirical studies of utility measurement.

Since the classical study of Mosteller and Nogee in 1951, there have been a number of experimental investigations aimed at utility measurement. For example, Ward Edwards ([1953], [1954a], [1954b], [1955]) has made a series of studies that show the difficulty of separating the measurement of utility and subjective probability. Clyde Coombs has also written a series of papers (e.g. [1958a] [1958b]) bringing out some of the difficulties of utility measurement and examining alternative theories, some of which I mention below. Donald Davidson, Sidney Siegel and I attempted to apply the theoretical approach of F. P. Ramsey to subjective probability and utility to actual experimental situations, and the results of these studies were published in our 1957 book. The three of us have since that time each done further work in the measurement of utility working with other colleagues (Davidson and Marschak [1959], Siegel [1956], Becker and Siegel [1958]).

Without entering into details, it seems to me that the rough conclusion to be drawn from these various studies is that for certain narrowly defined situations it is possible to show that subjects' behavior may be interpreted in terms of their using a utility function to make behavioral

[1]This paper has been written in connection with research supported by the Group Psychology Branch of the Office of Naval Research.

decisions. On the other hand, it can scarcely be denied that this experimental literature is as yet entirely too fragile in relation to the massive claims sometimes made for utility theory. The kind of difficulties of measurement Ward Edwards has emphasized seems particularly important to understand better. It is perhaps not unexpected that relatively naive subjects behave in such a way that it is not possible to claim that they have a separable utility function and subjective probability function. The point is that the utilities for a given outcome interact with the subjective probabilities of the possible states of natures. It would be important to know to what extent this same kind of behavior is deep-seated in sophisticated subjects. More generally, I feel that the most important problem in the measurement of utility is not the refinement of additional models, but the prosecution of more robust attempts to measure utility. I mean by this that what we need above all are attempts to measure utility in situations in which people are making weighty and significant decisions. There are admittedly serious experimental difficulties to be faced by such a program, but I think they are surmountable. In fact, it is my impression that some of the people concentrating upon operations research in government and industry are on the verge of undertaking studies of the kind I am suggesting.

2. Empirical studies of subjective probability measurement.

Emir Shuford ([1959a], [1959b], [1961]) has been conducting some rather elaborate studies of subjective probability in which subjects are asked to estimate proportion, as for example the proportion of horizontal and vertical lines in a square array. S. S. Stevens and E. H. Galanter [1957] published a similar study. Ward Edwards has also performed some experiments of this character. J. Cohen and his associates in England have conducted an active program of investigation ([1956a], [1956b], [1957]); they have attempted to investigate among other things the empirical validity of the addition and multiplication laws of probability. Unfortunately, the experiments have been conducted in such a way that it is not precisely clear how their results should be interpreted.

A standard problem in the measurement of subjective probability that goes back to the Mosteller-Nogee study is the problem of separating the measurement of subjective probability from that of utility. For example, by assuming that utility is linear in money, the results of the Mosteller-Nogee experiment may be interpreted as measurements of subjective probability. Davidson, Siegel and I attempted to avoid this in our 1957 studies. The difficulty of those studies was the fragility of empirical results, because the measurement of subjective probability was only made on the basis of earlier measurements of utility, and the range of utility measurements was severely

limited. In 1958 I noted that it was possible to disentangle the two on the basis of the following simple arithmetical equivalent. Let us represent the decision problem by the following matrix. The roles of the three possible states of nature E_1, E_2 and the complement of the union of E_1 and E_2, the two possible decisions d_1 and d_2, and the outcomes a and b are indicated.

	d_1	d_2
E_1	a	b
E_2	b	a
$\overline{E_1 \cup E_2}$	0	0

On the standard expected utility hypothesis the expectation of the decision d_1, for example, is:

$$\mathcal{E}(d_1) = s(E_1)\, u(a) + s(E_2)\, u(b) + s(\overline{E_1 \cup E_2})\, u(0) ,$$

where s is the subjective probability function and u the utility function. Suppose $u(a) > u(b)$, which is only an ordinal assumption. Then the decision-maker should prefer d_1 to d_2 if and only if $s(E_1) > s(E_2)$. In the fall of 1958 I used this equivalence to conduct a study of the measurement of subjective probability. The situation was the following. Students interested in football at Stanford were asked each week at the beginning of the season to make a large series of bets on the number of games Stanford would win during the entire season. A typical bet would be the choice between waging a nickel on Stanford winning five, six or seven games and waging a nickel on the event of Stanford winning one, four or ten games. As each week passed, one number was eliminated from the sets of numbers of possible games Stanford might win. Assuming the additivity of subjective probability in terms of the atomic events, each wager could be converted into an inequality in the atomic events. It was then a straightforward problem to ask if the inequalities in terms of the atomic events had a non-trivial solution. In case it did not, I used the linear programming approach already followed in Chapter 3 of the book with Davidson and Siegel. Unfortunately, the subjects were in certain respects entirely too consistent in their responses. I mean by this that by assigning a probability of 1 to one atomic event, in this case winning six games, a large part of the subjects' data was then accounted for. In designing the experiment, I had not foreseen this possibility, and among a large number of comparisons of events I had not included the comparison of each atomic event with its complement in the entire possible set. This defect of the unpublished study was not a weakness of the general method exemplified by the equivalence above, and new experiments are now under design to pursue the matter in greater detail.

Because of the intrinsic difficulty of utility measurement experiments with outcomes of serious consequence, the experimental investigation of subjective probability seems to me in many respects to be more promising. I see no theoretical or practical reasons for not being able to conduct experiments on the measurement of subjective probability for events of a heterogeneous but significant character.

3. New mathematical results.

I will not attempt anything like a detailed survey of this topic. I would like to mention the work of Marschak and his collaborators (e.g., Block and Marschak [1960], Marschak [1960]) in studying the relations between various utility indicators, the elegant axiomatizations considered by Gerard Debreu (e.g., [1958], [1959]) and the very extensive study of stochastic theories of choice by John Chipman [1960]. Also of interest is the study by Ernest W. Adams and Samuel Messick of classical Thurstone scaling methods [1958].

Dana Scott and I have tried to lay down a general theory of measurement [1958]. One of the motivations of this general theory has been the necessity of having an exact framework in order to prove negative results in the theory of rational behavior. I won't attempt a technical description here of our investigations but simply remark that our results show some of the difficulties of axiomatizing the theory of rational behavior when specific structural axioms like the assumption that the states of nature may be partitioned as finely as one pleases are omitted. Mention should also be made of a specific negative result by Kraft, Pratt and Seidenberg [1959] on the existence of a strictly agreeing probability measure for finite fields of events in which de Finetti's qualitative axioms are satisfied. It seems fairly clear that their results can be used to obtain some general negative metamathematical results on the axiomatizability of the theory of qualitative probability by using the kind of arguments pursued by Scott and me.

4. Stochastic theories of choice.

Perhaps the most important single development in utility theory in the past few years is the move away from an algebraic to a stochastic theory. The motivation for this move has come in large part from an increasing concentration on the psychological background of the theory of choice. The recognition that when individuals are presented with the same choice in essentially the same circumstances they will not always choose in the same fashion is classical in psychophysics and psychometrics. Let p_{ab} be the probability that a will be chosen over b. In terms of this function of two arguments, it is possible to state at the level of probabilities

rather than at the level of actual choices conditions similar to the familiar transitivity postulates. For example, weak stochastic transitivity is defined as follows:

$$\underline{\text{If}} \quad p_{ab} \geq \frac{1}{2} \quad \underline{\text{and}} \quad p_{bc} \geq \frac{1}{2} \quad \underline{\text{then}} \quad p_{ac} \geq \frac{1}{2} \ .$$

The stricter condition of strong stochastic transitivity is defined thus:

$$\underline{\text{If}} \quad p_{ab} \geq \frac{1}{2} \quad \underline{\text{and}} \quad p_{bc} \geq \frac{1}{2} \quad \underline{\text{then}} \quad p_{ac} \geq \underline{\max} \ (p_{ab} \ , \ p_{bc}) \ .$$

It is apparent that strong stochastic transitivity implies weak stochastic transitivity. Tests of these transitivity postulates are to be found in Papandreou, et al. [1957], and Davidson and Marschak [1959]. Obviously the satisfaction of either one of these postulates does not guarantee in itself the existence of a utility function connected with the probability of choosing a or b . There are at least three distinct approaches to the problem of the existence of such a function in the context of stochastic theories of choice. One arises naturally from the algebraic axiomatization of utility given by Muriel Wood Winet and myself [1955]. The condition on the utility function in this case is that

$$u(a) - u(b) \leq u(c) - u(d) \quad \text{if and only if} \quad p_{ab} \leq p_{cd} \ .$$

A second interesting possibility is an assumption due to Bradley and Terry originally but much exploited in recent years by Luce [1959]. This approach asserts the existence of a response strength function v such that

$$p_{ab} = \frac{v_a}{v_a + v_b} \ .$$

Still a third approach is to express p_{ab} in terms of Thurstone's theory of paired comparison. In this case, we obtain the following double integral for p_{ab} .

$$p_{ab} = \int_{-\infty}^{\infty} \int_{y}^{\infty} n_a(x) \ n_b(x) \ dx \ dy \ ,$$

where n_a and n_b are normal densities.

One puzzle that does not seem to have been satisfactorily solved for stochastic theories of choice is how to introduce a subjective probability function as well as a stochastic utility function. In the experimental situations as yet studied there has been no real need for a probability function, but for the more general kinds of applications intended it is necessary to integrate subjective probability with the existence of a stochastic utility function. A typical problem would be to re-do the kind of postulates given in Savage's Foundations of Statistics. It is interesting to note that probabilities enter in a way different from that envisaged by Savage,

namely, in terms of a description of the behavior of the decision maker. Presumably this addition supplements rather than replaces the subjective probability functions considered by Savage.

5. Derivation of utility functions from more general behavioral theories.

This topic seems to be a more important one for the future than it has yet turned out to be. The difficulty, of course, is that of finding adequate theories from which to derive the more special theory of utility. Here I will mention two different beginnings of such developments. The first is to be found in Duncan Luce's recent book, Individual Choice Behavior, in which he derives utility functions from more general choice postulates. It is indeed surprising that from such simple assumptions about choice behavior, very similar in fact to Rényi's [1955] axiomatization of conditional probability, one can obtain such far-reaching results not only about utility theory but also about learning theory as well.

I have attempted a derivation of stochastic utility functions in a somewhat different vein from Luce by deriving them from the postulates of stimulus sampling theory (Suppes [1961]). The empirical validity of these stochastic functions is examined in Chapter 11 of Suppes and Atkinson [1960]. The remark made earlier about the role of subjective probability once stochastic theories are used applies as well to the derivations from general behavioral axioms by Luce and myself. Neither derivation shows much promise of providing a lead for the introduction of subjective probability functions. A conceptual clarification of this situation seems to me to be a pressing problem.

6. Concepts to replace utility.

Economists are familiar with Allais' [1955] animadversions against the concept of expected utility and also with Shackle's [1955] concept of potential surprise. Among psychologists Coombs and his associates (for example, Coombs and Pruitt [1960]) have studied variance preferences and related behavioral measures as alternative methods for predicting decisions. To a certain extent they have been able to show that in some experimental situations a more refined prediction of behavior can be obtained by these methods.

In conjunction with Halsey Royden and Karol Walsh [1959] I attempted to substitute a notion of utility of gambling for the standard utility notion. The utility-of-gambling "surfaces" we obtained for our subjects were complicated. For some subjects the utility-of-gambling model predicted behavior better than the expected utility model, but this was not true for most subjects.

As these remarks are meant to indicate, the experimental evidence behind any substantial

class of models to replace utility models is as yet extremely fragile. Certainly this is not surprising in view of the rather limited experimental evidence in the whole field. It is my own belief that the relatively simple concepts of utility including those of a stochastic kind, thus far developed, will not be adequate to explain much of the fine structure of decision behavior. On the other hand, the intrinsic reasonableness of the general idea of the expected utility hypothesis is too overwhelming for new theories not to have to derive it as a first approximation.

7. Some scaling problems.

The concept of cardinal utility as a function which is defined on consequences or outcomes and which is unique up to a linear transformation, is thoroughly imbedded in the literature and has been overwhelmingly accepted since the first appearance of the von Neumann and Morgenstern axiomatization of utility. It is perhaps worth emphasizing that there is nothing in the theory essentially requiring such a result. We may illustrate the ideas by considering Luce's utility function. Recall that the probability p_{ab} of choosing a over b is expressed in terms of the response strength or utility v as follows:

$$p_{ab} = \frac{v_a}{v_a + v_b} . \tag{1}$$

It is at once apparent that v is unique up to a ratio scale on the weakest sort of assumptions. On the other hand, we can express essentially the same sorts of facts by the following equation:

$$p_{ab} = \frac{1}{1 + e^{v_b - v_a}} . \tag{2}$$

The same assumptions that make v a ratio scale in (1) make v a difference scale, that is, unique up to a translation, in (2). Furthermore, by introducing an arbitrary positive power, we may obtain a log interval scale:

$$p_{ab} = \frac{v_a^k}{v_a^k + v_b^k} .$$

As far as I can see, there are no particular mathematical arguments for choosing one scale rather than another. It is simply a matter of using that which is most convenient in developing other aspects of the theory or in making computations. Not every one who has thought about this matter will accept this arbitrariness concerning the scale type but I have yet to see any convincing arguments to the contrary, and there is one powerful argument in favor of the viewpoint I advocate. It may be illustrated by considering alternative representations of extensive quantities like mass or distance by a multiplicative scale. This alternative representation rests on the

fact simply of the isomorphism between additive semi-groups of positive real numbers and multiplicative semi-groups of positive real numbers (I omit exact definitions). These structural isomorphisms are fundamental, and I see no way for a theory of measurement to escape them.

Most of the remarks in this paper have necessarily been superficial and cursory. I have meant to indicate some main lines of development. A serious study of these matters certainly could not proceed from examination of this paper, but would have to begin with the references I have listed.

[See References beginning on next page]

REFERENCES

Adams, E. W. Survey of Bernoullian utility theory. In Mathematical Thinking in the Measurement of Behavior (Solomon, H., Ed., Free Press, Glencoe, Ill., 314 pp., 1959)

Adams, E. W., and Messick, S. An axiomatic formulation and generalization of successive intervals scaling. Psychometrika, 23 (1958)

Allais, M. Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'ecole Americaine. Imprimerie Nationale, Paris, 55 pp. (1955)

Arrow, K. J. Utilities, attitudes, choices: a review note. Econometrica, 26 1-23 (1958)

Becker, S. W., and Siegel, S. Utility of grades: level of aspiration in a decision theory context. J. Exptl. Psychol., 55, 81-85 (1958)

Block, H. D., and Marschak, J. Random orderings and stochastic theories of responses. In Contributions to Probability and Statistics (Olkin, Ghurge, Hoeffding, Madow, Mann, Eds., Stanford University Press, Stanford, Calif., 1960)

Chipman, J. S. The foundations of utility. Econometrica, 28, 193-224 (1960)

Cohen, J., Dearnaley, E. J., and Hansel, C.E.M. The addition of subjective probabilities: the summation of estimates of success and failure. Acta Psychol., 12, 371-80 (1956a)

Cohen, J., and Hansel, C.E.M. Risk and Gambling; the Study of Subjective Probability (Philosophical Library, New York, N. Y., 153 pp., 1956b)

Cohen, J., Dearnaley, E. J., and Hansel, C.E.M. Measures of subjective probability: estimates of success in performance in relation to size of task. Brit. J. Psychol., 48, 272-75 (1957)

Coombs, C. H. On the use of inconsistency of preferences in psychological measurement. J. Exptl. Psychol., 55, 1-7 (1958a)

Coombs, C. H., and Komorita, S. S. Measuring utility of money through decisions. Am. J. Psychol., 71, 383-89 (1958b)

Coombs, C. H., and Pruitt, D. G. A Study of Decision Making under Risk. Rept. No. 2900-33-T, Willow Run Laboratories (Univ. of Michigan, Ann Arbor, Mich., 25 pp., April 1960)

Davidson, D., and Marschak, J. Experimental tests of a stochastic decision theory. In Measurement: Definitions and Theories (Churchman, C. W., and Ratoosh, P., Eds., John Wiley and Sons, New York, N. Y., 274 pp., 1959)

Davidson, D., Suppes, P., and Siegel, S. Decision-making: An Experimental Approach (Stanford Univ. Press, Stanford, Calif., 121 pp., 1957)

Debreu, G. Stochastic choice and cardinal utility. Econometrica, 26, 440-44 (1958)

Debreu, G. Cardinal utility for even-chance mixtures of pairs of sure prospects. Rev. Econ. Studies, 71, 174-77 (1959)

Edwards, W. Probability-preferences in gambling. Am. J. Psychol., 66, 349-64 (1953)

Edwards, W. Probability preferences among bets with differing expected values. Am. J. Psychol., 67, 57-67 (1954a)

Edwards, W. Variance preferences in gambling. Am. J. Psychol., 67, 441-42 (1954b)

Edwards, W. The prediction of decisions among bets. J. Exptl. Psychol., 51, 201-14 (1955)

Edwards, W. Behavioral decision theory. Ann. Rev. Psychol., 12, 473-98 (1961)

Kraft, C. H., Pratt, J. W., and Seidenberg, A. Intuitive probability on finite sets. Ann. Math. Stat., 30, 408-419 (1959)

Luce, R. D. Individual Choice Behavior (John Wiley and Sons, New York, N. Y., 153 pp., 1959)

Marschak, J. Binary choice constraints on random utility indicators. In Stanford Symposium on Mathematical Methods in Social Sciences (Stanford Univ. Press, Stanford, Calif., 1960)

Papandreou, A. G. A test of a stochastic theory of choice. University of California Publications in Economics, 16, No. 1, 1-18 (1957)

Rényi, A. On a new axiomatic theory of probability. Acta Math., 6, 285-335 (1955)

Royden, H. L., Suppes, P., and Walsh, K. A model for the experimental measurement of the utility of gambling. Behav. Sci., 4, 11-18 (1959)

Scott, D., and Suppes, P. Foundational aspects of theories of measurement. J. Sym. Logic, 23, 113-28 (1958)

Shackle, G.L.S. Uncertainty in Economics (University Press, Cambridge, England, 267 pp., 1955)

Shuford, E. H. A comparison of subjective probabilities for elementary and compound events. U. N.C. Psychom. Lab. Rep., No. 23 (1959a)

Shuford, E. H., and Wiesen, R. A. Bayes estimation of proportion: The effect of stimulus distribution and exposure time. U. N.C. Psychom. Lab. Rep., No. 23 (1959b)

Shuford, E. H. Percentage estimation of proportion as a function of element type, exposure time, and task. J. Exptl. Psychol., 61, 430-36 (1961)

Siegel, S. A method for obtaining an ordered metric scale. Psychometrika, 21, 207-16 (1956)

Stevens, S. S., and Galanter, E. H. Ratio scales and category scales for a dozen perceptual continua. J. Exptl. Psychol., 54, 377-409 (1957)

Suppes, P. Behavioristic foundations of utility. Econometrica, 29, 186-202 (1961)

Suppes, P., and Atkinson, R. C. Markov Learning Models for Multiperson Interactions. (Stanford Univ. Press, Stanford, Calif., 1960)

Suppes, P., and Winet, M. An axiomatization of utility based on the notion of utility differences. J. Manage. Sci., 1, 259-70 (1955)