SURROGATES FOR UNCERTAIN DECISION PROBLEMS: MINIMAL INFORMATION FOR DECISION MAKING[1]

BY STANLEY REITER

1. INTRODUCTION

PROBLEMS OF choosing optimal decisions in uncertain situations are of increasing importance in economics. In the past, theoretical approaches to rational behavior in the face of uncertainty have largely ignored the problem of acquiring and using the information required in order to solve decision problems.[2] Our most highly developed theoretical treatment of the problem, statistical decision theory, seems to call for more information than is usually available in practical applications [1] [3]. In some situations actually arising in practice it has been possible to simplify the uncertain decision problem so as to show clearly just what information was really needed to solve the problem. The purpose of this paper is to expose the logic of this simplification, and by means of the concept of "surrogate" developed below, to find out just what constitutes minimal information required for decision-making in a certain wide class of uncertain decision problems.

Recent attempts to apply decision theory in practical situations have pointed the direction for us. Professor H. Theil has recently considered a problem in which "economic welfare" depends on two quantities, one to be chosen by a (wise) policy-maker, the other subject to a probability distribution. Professor Theil has stated that if "welfare" is a quadratic function of these two variables, then it is possible to choose an optimal decision knowing only the expected value of the random variable [7].

A second example of interest, leading to a generalization of Professor Theil's results, arose in the work of Professors Holt, Modigliani, and Simon on the scheduling of the production of paint for a succession of time periods [5]. In this problem, as in many others like it, the future sales were unknown and the probability distributions of future sales were also unknown. In this situation Professor Simon obtained a result which enabled the practical application of the scheduling technique to go forward. Professor Simon was able to show that if the planning is for a finite number of time periods into the future, and if the total cost function for the planning period is quadratic, then an optimal decision for the first period can be chosen if one knows only the expected sales in each future time period [6].[3]

Another example which should perhaps be mentioned is afforded by Professor

[1] I am indebted to my colleagues at Purdue and to H. A. Simon of the Carnegie Institute of Technology and L. J. Savage of the University of Chicago for helpful comments and criticisms.

[2] On the other hand, in [2] we find lack of complete information about the structure of costs given a central role in the problem of choosing optimal production plans for a firm. We also find there a prior use of the term "surrogate" in a sense analogous to its use here.

[3] No comparison of the generality of problems or results is intended here.

Hicks who, in attempting to assimilate the theory of the firm under uncertainty to the certainty theory, in effect restricted his theory to those cases in which optimal output decisions would be reached by the firm if it knew only "the most probable price $\pm$ an [unspecified] allowance for risk" [**4**]. This is a simplification of an uncertain decision problem of the type which interests us, but, unfortunately the class of cases in which this simplification will result in optimal decisions is not specified.

We are led by these examples to ask:

(1) How can we formulate the problem of finding what constitutes minimal information for solving uncertain decision problems?

(2) Can we characterize a broad class of cases in which the information necessary to reach an optimal decision is substantially simpler than what is now understood from decision theory to be required?

We may summarize our answer to the first question by saying that we introduce the concept of a "surrogate" for an uncertain decision problem, namely, a related problem in which the troublesome probability distribution of the original problem is replaced by something simpler, and whose solution is a decision which is optimal in the original problem.

We can then approach our second question by studying the properties of the surrogate problem. We shall show (Theorem I) that if the loss (or utility or pay-off) functions can be factored in a certain way, then the information required by the surrogate problem takes a relatively simple form. We can therefore determine, by examining the loss function, the minimum information needed to solve an uncertain decision problem, even in the (typical) case where we do not know the relevant probability distributions.

We shall also show that even in cases not covered by Theorem I, even when the loss function cannot be factored in the sense of that theorem, it is still possible, and may be convenient, to approximate optimal decisions by means of a surrogate problem. This result may be of some practical importance, for it is rare in economics that we can claim to know cost or profit functions exactly. We can be assured that if we have a good approximation to the true loss function, we will reach a decision which carries an expected loss close to the minimum expected loss.

2. THE UNCERTAIN DECISION PROBLEM AND ITS SURROGATE

In this section we shall formulate an uncertain decision problem and define the concept of a surrogate. We shall also show that every uncertain decision problem has a surrogate.

Following the formulations of statistical decision theory, we may specify an uncertain decision problem as follows.

We have a set X whose elements are to be thought of as possible actions or decisions.

Thus, $x (x \in X)$ might be a scalar denoting the rate of output of a firm; alternatively x might be a vector whose rth component denotes the firm's rate of output of commodity r.

We have a set Y whose elements we think of as possible "actions" of the environment, or "nature." We shall later apply probability measures to subsets of Y. We shall assume that the class of measurable subsets is full enough to avoid technical difficulties.

Thus, $y(y \in Y)$ might be a scalar random variable denoting the rate of sales of a certain commodity; or a random vector denoting, say, the rates of sales of several commodities.

Finally we have a real valued function f defined on the Cartesian product set

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}.$$

We interpret $f(x, y)$ as denoting the loss to the decision-maker of taking action x when "nature's action" is y.

We let P be a set of probability distributions so that $p \in P$ is a probability distribution of the random variable y.

We can now formulate an uncertain decision problem.

Problem I. Given $p \in P$ choose $x \in X$ so as to minimize

$$F_p(x) = \int_Y f(x, y)\, dp(y).$$

We denote the set of solutions or optimal decisions in Problem I by $X^*(p)$.

We are looking for possibilities of simplification in the direction of replacing the distributions p by some simpler quantities, about which information may be more readily available.

We consider replacements of the probability distributions by points in a certain set. Analytically this may be viewed in terms of a transformation T of P into a set H.

$$\eta = T(p), \qquad p \in P.$$

Thus, by means of T we associate to each distribution of y a point $\eta \in H$. Let G be a function defined on $X \times H$.

Then we consider

Problem II. Given $\eta \in H$, choose $x \in X$ so as to minimize $G(x, \eta)$.

For each $\eta \in H$, denote the set of solutions of Problem II by $X^{**}(\eta)$.

We shall call Problem II a *surrogate* for Problem I if

$$\eta = T(p) \tag{1}$$

and

$$X^{**}(\eta) \subset X^*(p), \tag{2}$$

i.e., every solution of Problem II given η is a solution of Problem I with a p to which η corresponds.

A surrogate for an uncertain decision problem is then a minimization problem in which the probability distribution in the original problem is replaced by a point in another space, it being specified that both problems require minimization over the same set of possible actions, and that an action which is a solution of the surrogate is also a solution of the original problem.

Having defined a surrogate problem, one wonders whether every uncertain decision problem has one. To show that every uncertain decision problem has a surrogate, we take H to be P itself, and take for T the identity transformation. Thus, T takes p considered as a probability distribution into $T(p) = \eta$ which is the same thing considered as a point. Then

$$G(x, \eta) \equiv F_p(x).$$

3. A CLASS OF SURROGATE PROBLEMS

We now delineate a class of uncertain decision problems which admit a surrogate problem that constitutes a real simplification. The characterizing condition involves only the function f, namely that it be factorable in a certain way. Hence, our conclusions can be arrived at in complete ignorance of the probability distribution.

THEOREM I. *If, in an uncertain decision problem having the form of Problem I, the loss function f can be written in the form*

$$f(x, y) = \sum_{i=1}^{n} A_i(x)B_i(y)$$

where all functions are positive, then there is a surrogate problem with

$$(i) \qquad \eta = (\eta_1, \eta_2, \cdots, \eta_n), \qquad \eta_i = \int_y B_i(y)\, dp(y), \qquad i = 1, 2, \cdots, n,$$

and

$$(ii) \qquad G(x, \eta) = \sum_{i=1}^{n} A_i(x)\eta_i .$$

PROOF:

$$F_p(x) = \int_y \sum_{i=1}^{n} A(x)B_i(y)\, dp(y) = \sum_{i=1}^{n} A_i(x) \int_y B_i(y)\, dp(y).$$

Hence if we take

$$G(x, \eta) = \sum_{i=1}^{n} A_i(x)\eta_i$$

where

$$\eta_i = \int_y B_i(y)\, dp(y),$$

we have

$$G(x, \eta) \equiv F_p(x).$$

Hence F and G must have the same extreme.

Thus, Theorem I tells us that if the loss function in uncertain decision problems can be factored, the information about the uncertain events needed in order to solve the problem is specified by the vector η, and the dependence of η on the form of the factors of the loss functions, as well as on the unknown distributions, is given explicitly in conclusion (i) of Theorem I.

Theorem I assures us that for a certain class of loss functions there is a surrogate problem with an η of relatively simple form. Suppose we specify a minimization problem involving an η of appropriate form and ask, "For which uncertain decision problems is the given problem a surrogate?" It is interesting to know whether we get anything outside the class of functions considered in the hypotheses of Theorem I. In other words, we wish to know whether the conditions of Theorem I are necessary as well as sufficient conditions. To answer this question we need only look at the case when η is the expected value of a one-dimensional y, for this case gives us the strongest nontrivial hypotheses for necessity. Hence, if we cannot show necessity in this case, we cannot show it under weaker hypotheses.

As the following counter-examples show, the factorability condition is not necessary; the counter-examples are instructive and suggest the difficulties involved in finding necessary conditions.

First, if the function f is independent of y, then a surrogate involving only the expected value of y can be constructed, no matter what the form of f, e.g., add the term $\mathcal{E}y$ to f. Similarly, in the case of continuously differentiable f if the partial derivative f_x is independent of y, the same conclusion holds.

Second, a more subtle possibility occurs when the set of points satisfying the necessary condition for a minimum with respect to x is independent of y, even though f and f_x are not independent of y.

For example, let

$$f_x(x, y) = A(x - c)B(x, y), \qquad B \neq 0,$$

where c is independent of y. Then

$$A(x - c)B(x, y) = 0$$

implies

$$A(x - c) = 0$$

and hence that the minimizing x is independent of y. Thus for any function of the form

$$f(x, y) = \int^x A(w - c)B(w, y)\, dw$$

a surrogate problem can be constructed which does not involve y; the constant, $\mathcal{E}y$, can be added without affecting the problem.

4. APPROXIMATE SOLUTIONS BY MEANS OF SURROGATE PROBLEMS

It may happen in uncertain decision situations arising in practice that the loss function is not known exactly, but only an approximation is available. Or it may happen that the loss function is not factorable and so does not satisfy the hypotheses of Theorem I. In such cases, it may be convenient to approximate optimal decisions by means of surrogate problems.

If the loss function f is continuous, it can be approximated uniformly on a bounded set by a polynomial of degree m, where m depends on the desired error of approximation, ϵ. Thus, given $\epsilon > 0$ there is an m so that

$$\left| f(x, y) - \sum_{i=0}^{m} A_i(x) y^{m-i} \right| < \epsilon$$

where A_i are polynomials in x.

Then, the surrogate problem with

$$G(x, \eta) = \sum_{i=0}^{m} A_i(x) \eta_i$$

and

$$\eta_i = \int_Y y^{m-i}\, dp(y), \qquad i = 0, 1, 2, \cdots, m$$

will yield a decision which is approximately optimal in the original problem. We have no guarantee that a decision which is optimal for the surrogate problem will be "close" to a decision optimal in the original problem. We are, however, assured by the uniform approximation that a decision obtained from the surrogate problem cannot result in additional loss of more than the preassigned amount ϵ.

Here an additional element is brought into the decision problem, for the decision-maker must now weigh the value of a better guarantee $\epsilon' < \epsilon$ against the increased cost of estimating a higher dimensional vector η' necessitated by the consequent higher degree polynomial approximation. Alternatively, if the amount of information is restricted, so that, say, only η_1 is available, the error of the best polynomial approximation for which η_1 appears in the surrogate problem will give an estimate of the expected loss due to insufficient information about the uncertain events.

Purdue University

REFERENCES

[1] BLACKWELL, D. AND M. A. GIRSHICK, *Theory of Games and Statistical Decisions*, John Wiley and Sons, New York, 1954.

[2] CHARNES, A., W. W. COOPER, AND B. MELLON, "A Model for Optimizing Productions by Reference to Cost Surrogates," *Econometrica*, July 1955, pp. 307–323.

[3] Dvoretsky, A., J. Kiefer, and J. Wolfowitz, "The Inventory Problem I and II," *Econometrica*, April 1952, p. 187; July 1952, p. 450.
[4] Hicks, J. R., *Value and Capital*, 2nd ed., Clarendon Press, Oxford, 1946, pp. 125–126.
[5] Holt, C. C., F. Modigliani, and H. A. Simon, "Linear Decision Rule for Production and Employment Scheduling," *Management Science*, Vol. 2, October 1955, pp. 1–30.
[6] Simon, H. A., "Dynamic Programming Under Uncertainty with a Quadratic Criterion Function," *Econometrica*, Vol. 24, No. 1, January 1956, pp. 74–81.
[7] Theil, H., "Econometric Models and Welfare Maximization," *Weltwirtschaftliches Archiv*, 1954. (I am indebted to Professor H. A. Simon for calling this paper to my attention.)