A REHABILITATION OF THE PRINCIPLE OF INSUFFICIENT REASON*

HANS-WERNER SINN

It is shown that two of the axioms necessary for the expected utility rule imply the Principle of Insufficient Reason. Whenever a decision maker knows the possible states of the world, but completely lacks information about the plausibility of each single state, he has to behave as if all states occurred with the same objective probability, known with certainty. The result is applied to decision trees and used to solve a problem formulated by Savage in order to discredit the classical version of the Principle of Insufficient Reason.

I. INTRODUCTION

This paper is concerned with a very old subject dating back to J. Bernoulli [1713, pp. 88–89] and Laplace [1814, pp. iv, vii]: The Principle of Insufficient Reason. This principle says that if there is no reason to believe that out of a set of possible, mutually exclusive events no one event is more likely to occur than any other, then one should assume that all events are equally probable. For example, consider that you are throwing a die. Since you do not believe that one side is more likely to occur than any other, you regard all probabilities as equal. And indeed, repetitive throwing shows that this is correct.

Unfortunately, such an experimental verification is usually impossible in economic decision problems. Nevertheless, from a normative point of view, another rationalization for the Principle of Insufficient Reason can be given by the use of two axioms that in various versions are widely accepted in risk theory, the implications of which, however, have not yet been completely elaborated upon.

II. THE BASIC CHOICE PROBLEM

Our choice-theoretic approach is the usual one: The decision maker chooses one out of an opportunity set of mutually exclusive acts, $A_1, \ldots, A_m$, by referring to the (ex ante) results of these acts, $R_1, \ldots, R_m$, each of which represents a row of the decision matrix. In general we write a particular act's result R, which can be regarded as a lottery ticket, as

* I gratefully acknowledge comments by Ann Schwarz-Miller, Peter Howitt, an unknown referee, and particularly, Hans Heinrich Nachtkamp. The responsibility for all shortcomings, however, is entirely mine.

© 1980 by the President and Fellows of Harvard College. Published by John Wiley & Sons, Inc.
The Quarterly Journal of Economics, May 1980 0033-5533/80/0094-0493$01.40

$$(1) \qquad R = \begin{bmatrix} S_1 & S_2 & \cdots & S_n \\ r_1 & r_2 & \cdots & r_n \end{bmatrix}.$$

Here r_j is the subresult or lottery prize resulting if, given the particular act, the state of nature turns out to be S_j.

We do not place any substantial restrictions on what a subresult may be. It is defined as containing all information of interest to the decision maker concerning the situation that will prevail after the state of nature has been revealed. Thus, special state preference in the sense of Hirshleifer [1965, esp. p. 522] is excluded. The state of nature, under which a particular subresult r_j occurs, is of interest *only* for its information about the likeliness or plausibility of this subresult. Whenever the subresults of two states of nature, bearing the same plausibility information, are interchanged, this is meaningless to the decision maker.

Sometimes the plausibility information might be so good that the decision maker even knows objective probabilities, $p_1, p_2, \ldots, p_n$, for the states of nature. In this case we can also write the lottery ticket as

$$(2) \qquad R = \begin{bmatrix} p_1 & p_2 & \cdots & p_n \\ r_2 & r_2 & \cdots & r_n \end{bmatrix}.$$

Unfortunately, such a decision problem under "risk" is by no means typical in economics. Instead the decision maker has normally at best a very bare knowledge of the plausibility of the states of nature. If he has no idea at all which of the states of nature is more likely to obtain than any other, then we have a decision problem under "complete uncertainty." This is the case with which we are dealing.

Again in line with most choice-theoretic approaches, we assume the

AXIOM OF ORDERING.[1] There exists a complete weak transitive ordering of the results R_i, $i = 1, \ldots, m$.

Here "weak" means that the decision maker is able to decide whether one result is at least as good as another one and "complete" indicates that all results can be compared in this way. This axiom is the first of the two axioms mentioned above. As is well-known, it implies the existence of a preference function $\psi(R)$, defined up to a strictly increasing monotonic transformation.

1. We do not really require completeness here, but refer to this version of the axiom of ordering only, since it is a very common one. In fact, for our result to hold, all one needs is transitivity and indifference.

Several of the approaches to establishing the function $\psi(\cdot)$ were constructed especially for the case of unknown probabilities. We should mention the classical rules such as Wald's [1945] minimax criterion, the optimism-pessimism index of Hurwicz [in Milnor, 1954], or the minimax regret principle of Niehans [1948] and Savage [1951], none of which require probabilities.[2] The most frequently accepted is the alternative approach of Savage [1953, 1954],[3] according to which the decision maker has to assess subjective probabilities for the states of nature and then to maximize von Neumann-Morgenstern expected utility. Unfortunately, however, this approach is still quite vague, since it does not tell the decision maker *which* subjective probabilities he should assess. The aim of this study is to fill this gap. We intend to show that under complete ignorance it is wise to replace $S_1, S_2, \ldots, S_n$ in vector (1) each by an objective probability $1/n$; i.e., we shall show the equivalence,

$$(3)\quad \begin{bmatrix} S_1 & S_2 & \cdots & S_n \\ r_1 & r_2 & \cdots & r_n \end{bmatrix} \sim \begin{bmatrix} 1/n & 1/n & \cdots & 1/n \\ r_1 & r_2 & \cdots & r_n \end{bmatrix}.$$

We do not, however, intend to specify the function $\psi(\cdot)$ completely.

Attempts to rationalize a similar result were already provided in different ways by Chernoff [1954] and Milnor [1954], but their axioms are quite technical and intuitively not very appealing, at least according to their critics, Luce and Raiffa [1957, esp. pp. 291 and 296]. Moreover, Chernoff and Milnor assume that the subresults are already transformed into von Neumann-Morgenstern utilities. They thereby implicitly require the Archimedes or continuity axiom, which is not accepted in lexicographic preference theory [Georgescu-Roegen, 1954, esp. p. 525] and thus should not be used if not necessary.

III. THE ELLSBERG PARADOX

Before deriving our result, we digress to the so-called Ellsberg [1961] paradox, since it provides a good basis for understanding the problem. Consider the following (slightly modified) game: There is an urn containing white and black balls. After the player has chosen a color, one ball is drawn. If it is of the chosen color, \$100 is paid to the player, otherwise nothing. The player is then asked to declare his

2. It should be noted that according to Niehans [1948] and Savage [1951] the preference index $\psi(\cdot)$ is not only a function of the ex ante result that is to be evaluated, but also of all other ex ante results within the opportunity set, a special feature that does not seem very convincing.

3. For earlier works see Ramsey [1931] and de Finetti [1937]. Extensions of the concept can be found in Luce and Raiffa [1957], and Schlaifer [1959].

maximum willingness to pay to participate in the game given one of the two alternatives:

(a) It is known that there are an equal number of white and black balls in the urn (risk).

(b) The ratio of white and black balls is unknown (complete uncertainty).

Obviously in case (a) there are objective probabilities of 1/2 for each color and in case (b) completely unknown probabilities. Thus, the extreme possibilities "risk" and "complete uncertainty" are compared.

The usual answer that can be observed indicates a clear preference for game (a), i.e.,

$$\pi_a > \pi_b, \tag{4}$$

where π_a and π_b are the maximal prices in the risk and uncertainty case, respectively. This answer seems to contradict the hypothesis that people assess subjective probabilities at all, for if they did so, they would pay at least as much for the uncertainty game as for the risk game. The reason is that, while the risk game provides a winning chance of 1/2, the uncertainty game provides the same chance if the subject probabilities are equal and a greater chance if they differ.

Several attempts have been made to explain the observed behavior: all are certainly fruitful for positive theory of human behavior.[4] However, they cannot be accepted as a guide to wise action, for, as Raiffa [1961] has pointed out, the people have simply made a mistake.[5] Consider Raiffa's extension of the game: Suppose that the player has already given his maximal prices for the risk (a) and uncertainty (b) game. Then he is asked whether he has any preference for betting black or white in the uncertainty case. If the answer is negative, the game is modified such that

(c) a coin decides which color to bet in game (b).

When asked for his maximum willingness to pay (π_c) for this modified game, the player's typical reply is that he would pay the same as for the game (b), i.e.,

$$\pi_c = \pi_b. \tag{5}$$

4. Roberts [1963] argues that the decision maker misinterprets his task. Brewer [1963] and Schneeweiß [1968] suppose that the decision maker thinks himself to be taking part in a game against the experimenter. Krelle [1968] accepts instead the decision maker's choice as an outcome of a specific uncertainty aversion:

5. It is indeed reasonable to state a mistake, for after an explanation by the experimenter that is how people usually explain their choice themselves.

This answer, however, shows an inconsistency on the part of the decision maker, for the modified game (c) is in fact a genuine risk game with equal probabilities like game (a) so that one should expect the answer

$$\pi_c = \pi_a. \tag{6}$$

In order to see this identity, let us assume that "heads" means to bet "black," and "tails" to bet "white" and call the (unknown) relative shares or probabilities of black and white balls in the urn p_b and p_w. Then the decision maker has two chances to win. The first is that "heads" comes up and a black ball is drawn. Its probability is $0.5p_b$. The second is that "tails" comes up and a white ball is drawn, with the probability $0.5p_w$. Thus, the total probability of winning is $0.5(p_b + p_w) = 0.5$, and of course that of losing is 0.5 as well. This result holds regardless of what the "true" probabilities or relative shares of white and black balls are.

The question now becomes which is the mistake that people make. Are they wrong in saying that Raiffa's modified game and the uncertainty game are alike, or are they wrong in saying that uncertainty is something different from risk? In the following, I show the latter to be true by providing a theoretical basis for Raiffa's trick. To be specific, it will be demonstrated that the uncertainty game must be evaluated as if there were equal objective probabilities for each color, which are known with certainty.

IV. THE ROLE OF THE AXIOM OF INDEPENDENCE

The crucial role in the argument is played by the famous Axiom of Independence, which is the second of the two axioms initially mentioned:

AXIOM OF INDEPENDENCE.[6] Let p denote an objective probability, $0 < p \leq 1$, and r_α, r_β, and r_γ three arbitrary results. Then

$$r_a \begin{Bmatrix} < \\ \sim \end{Bmatrix} r_\beta \Leftrightarrow \begin{pmatrix} p & 1-p \\ r_\alpha & r_\gamma \end{pmatrix} \begin{Bmatrix} < \\ \sim \end{Bmatrix} \begin{pmatrix} p & 1-p \\ r_\beta & r_\gamma \end{pmatrix}.$$

6. The axiom is presented in the popular strong version of Samuelson [1953], although in fact we require only the weak version

$$r_\alpha \sim r_\beta \Leftrightarrow \begin{bmatrix} p & 1-p \\ r_\alpha & r_\gamma \end{bmatrix} \sim \begin{bmatrix} p & 1-p \\ r_\beta & r_\gamma \end{bmatrix},$$

which was first established with Marschak's [1950] "Postulate IV." Together with the Archimedes or continuity axiom and a nonsaturation axiom, the axioms of ordering and strong independence imply the expected utility rule.

The axiom says that if there is a choice between two lotteries, both of which provide the same subresult r_γ with the probability $1 - p$, but different subresults with probability p, then the ordering of the two lotteries should be the same as that of the two different subresults. In particular, if these subresults, although different in kind, are evaluated equally from an isolated point of view, both lotteries should also be equally desirable. With the exception of some earlier criticism by Allais [1953a,b,c], the axiom is almost universally accepted. Friedman and Savage [1952, p. 468] call it "unique among maxims for wise action" and even speculate that "the Greeks must surely have had a name for it."

We begin with some formal, though quite simple reasoning. Consider the following result vectors:

$$
\begin{aligned}
(7) \qquad R \equiv R^1 &\equiv \begin{bmatrix} S_1 & S_2 & \cdots S_n \\ r_1 & r_2 & \cdots r_n \end{bmatrix} \\
R^2 &\equiv \begin{bmatrix} S_1 & S_2 & S_3 & \cdots & S_n \\ r_n & r_1 & r_2 & \cdots & r_{n-1} \end{bmatrix} \\
\vdots \\
R^{n-1} &\equiv \begin{bmatrix} S_1 & \cdots & S_{n-2} & S_{n-1} & S_n \\ r_3 & \cdots & r_n & r_1 & r_2 \end{bmatrix} \\
R^{n-1} &\equiv \begin{bmatrix} S_1 & \cdots & S_{n-2} & S_{n-1} & S_n \\ r_3 & \cdots & r_n & r_1 & r_2 \end{bmatrix} \\
R^n &\equiv \begin{bmatrix} S_1 & S_2 & \cdots & S_{n-1} & S_n \\ r_2 & r_3 & \cdots & r_n & r_1 \end{bmatrix},
\end{aligned}
$$

where the decision maker has no idea how likely the single states of nature are, but knows that one of these must occur:

$$(8) \qquad p(S_1 \cup S_2 \cup \cdots \cup S_n) = 1.$$

The first vector is that of the real life decision problem. Our aim is to replace its S_i's by $1/n$, as expressed by (3) above. The other result vectors are artificial. We construct them by bringing the last subresult in front and moving the other subresults one step to the right. For instance, in R^2 we say that subresult r_2, which occurs in the real decision situation if state of nature 2 obtains, now occurs if state 3 obtains. For convenience, we may think of the R^j's as lottery tickets that we offer to the decision maker.

Now we ask the decision maker to evaluate the lottery tickets. For example, following the Axiom of Ordering, we ask: Do you agree that R^j is not worse than R^k? And vice versa: Do you agree that R^k is not worse than R^j? The answer to both questions must be "yes," thus implying equivalence of R^j and R^k. The reason is simply that a negative answer by the decision maker indicates that he thinks one state more plausible than another and does not suffer from complete ignorance, as we had previously assumed. Thus, we cannot escape the conclusion that, generally, all lottery tickets are equivalent:[7]

$$R^j \sim R^k \;\forall\; j,k. \tag{9}$$

The next step of the argument is to introduce a higher level lottery ticket giving the decision maker the chance of winning other lottery tickets with (completely known) objective probabilities that are produced by a random machine. Like the real world decision problem, the random machine has n states. They are denoted as E_1, $E_2, \ldots, E_n$ and the corresponding probabilities are $p_1, p_2, \ldots, p_n$. This higher level lottery ticket will now be transformed in a stepwise procedure, leaving its position in the preference scale constant.

In the beginning the higher level lottery ticket is identical to the result vector of the real life decision problem:

7. Despite complete ignorance, this equivalence would not hold if there were Hirschleifer's [1965] "state preference," which, as we pointed out in the beginning of the article, arises from a misspecification of the results. Consider the following decision matrix of an automobile insurance buyer who has no idea at all whether an accident leading to a monetary loss of $1,000 will take place or not. (This need not be a realistic assumption.)

	Accident (loss of $1,000)	No accident (no loss)
A_1	$0	$0
A_2	$800–$$\pi$	–$$\pi$
A_3	–$$\pi$	$800–$$\pi$

A_1 is the action of not buying insurance, whereas A_2 indicates that an 80 percent coverage insurance is bought for a premium of $$\pi$. Action A_3 means to buy an artificial lottery ticket in the above sense that costs $$\pi$ and provides a prize of $800 if there is no accident. Obviously we cannot assume that A_2 and A_3 are equivalent, for a risk-averse agent would clearly prefer to receive $800 if an accident takes place; i.e., he displays state preference. However, the state preference vanishes if the monetary loss arising from the accident itself is included in a description of the subresults:

	Accident	No accident
A_1	–$1000	$0
A_2	–($200 + $$\pi$)	–$$\pi$
A_3	–$$\pi$	–($200 + $$\pi$)

If $1,000 measures the whole money equivalent loss, actions A_2 and A_3 are indeed equally desirable, as we require. The crucial point is that a subresult describes *all* interesting aspects of the situation ex post, whereas the state of nature is only an indicator of its plausibility or probability ex ante. In the case at hand the states of nature are very bad indicators of probabilities. Yet they are *equally* bad indicators. So there is no meaningful difference between actions A_2 and A_3, although they are, of course, formally distinguishable. (It should be noted that the examples could easily be extended to nonmonetary results.)

$$R^1 = \begin{bmatrix} p_1 & p_2 & \cdots & p_n \\ R^1 & R^1 & \cdots & R^1 \end{bmatrix}; \tag{10}$$

i.e., the decision maker gets ticket R^1 regardless of the random machine's state. In order to prepare the first transformation step, consider the following way to write R^1:

$$R^1 = \begin{bmatrix} p_2 & 1 - p_2 \\ R^1 & \begin{pmatrix} \frac{p_1}{1-p_2} & \frac{p_3}{1-p_2} & \cdots & \frac{p_n}{1-p_2} \\ R^1 & R^1 & \cdots & R^1 \end{pmatrix} \end{bmatrix}. \tag{11}$$

The purpose of this transformation is to allow the utilization of the Independence Axiom, for its formulation is similar. Because of this axiom and the relationship (9), we can substitute the first subresult by R^2, defined in (7). Retransformation yields

$$R^1 \sim \begin{bmatrix} p_1 & p_2 & p_3 & \cdots & p_n \\ R^1 & R^2 & R^1 & \cdots & R^1 \end{bmatrix}. \tag{12}$$

After this, further transformation steps are performed in an analogous way, so that we finally obtain

$$R^1 \sim \begin{bmatrix} p_1 & p_2 & \cdots & p_n \\ R^1 & R^2 & \cdots & R^n \end{bmatrix}. \tag{13}$$

With this expression the main work is done. However, a simplification might be possible if we remember that the artificial result vectors have in common that their subresults are the real world subresults $r_1, r_2, \ldots, r_n$. So we may try to replace the last equivalence by

$$R^1 \sim \begin{bmatrix} p(r_1) & p(r_2) & \cdots & p(r_n) \\ r_1 & r_2 & \cdots & r_n \end{bmatrix}, \tag{14}$$

where $p(r_1), p(r_2), \ldots, p(r_n)$ denote the total inherent probabilities of getting the corresponding real world subresults. Indeed, at least generally, it is not difficult to calculate the probabilities. We obtain

$$
\begin{aligned}
p(r_1) &= p[(E_1 \cap S_1) \cup (E_2 \cap S_2) \cup \cdots \\
&\qquad \cup (E_{n-1} \cap S_{n-1}) \cup (E_n \cap S_n)] \\
p(r_2) &= p[(E_1 \cap S_2) \cup (E_2 \cap S_3) \cup \cdots \\
&\qquad \cup (E_{n-1} \cap S_n) \cup (E_n \cap S_1)] \\
&\ \ \vdots \\
p(r_2) &= p[(E_1 \cap S_n) \cup (E_2 \cap S_1) \cup \cdots \\
&\qquad \cup (E_{n-1} \cap S_{n-2}) \cup (E_n \cap S_{n-1})].
\end{aligned} \tag{15}
$$

Note that up to now special probabilities were not required for the states of the random machine. The equivalence shown above always holds. Therefore, we may arbitrarily set all probabilities equal:

$$p_1 = p_2 = \cdots = p_n = 1/n. \tag{16}$$

The advantage of this assumption is that (15) simplifies to

$$
\begin{aligned}
p(r_1) = p(r_2) = \cdots = p(r_n) &= p[E_i \cap (S_1 \cup S_2 \cup \cdots \cup S_n)] \\
&= p_i\, p(S_1 \cup S_2 \cup \cdots \cup S_n) = 1/n, \quad i = 1, 2, \cdots, n,
\end{aligned} \tag{17}
$$

since even though the decision maker has no idea how plausible the single states of nature are, he knows with certainty that one of these states must occur (see (8)). Inserting (17) into (14) and noting that $R^1 \equiv R$, we arrive at our final result:

$$R \sim \begin{bmatrix} 1/n & 1/n & \cdots & 1/n \\ r_1 & r_2 & \cdots & r_n \end{bmatrix}. \tag{18}$$

Q.E.D.

Verbally this conclusion is the following: *Under complete ignorance of any probabilities for the states of the world a row of the decision matrix must be evaluated as if each state would obtain with the same objective probability, known with certainty.*

V. Tree Diagrams

In many practical situations the decision problem has a structure resembling that illustrated in Figure I; i.e., the states of the world (S) are obtained if cases, subcases, subcases of subcases, etc., are distinguished.

An interesting question is which probabilities should be assigned to the states of the world if the decision maker has no idea at all how plausible the branches of a fork are. According to our previous result it seems adequate to distribute the probability sum of unity equally

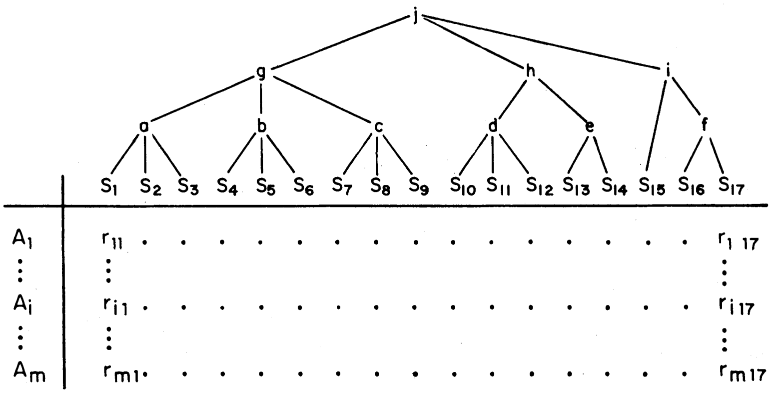

FIGURE I

among all branches of a fork. Then, according to the multiplication theorem of probabilities, the probability of a certain state of the world could easily be calculated by multiplying the probabilities of all branches over which one has to go in order to get from the trunk to the last branchlet defining the state in question. For the example of Figure I, this method would yield the following probabilities:

state number	1	2		9	10	11	12	13	14	15	16	17
probability	$\frac{1}{27}$	$\frac{1}{27}$	...	$\frac{1}{27}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$.

If correct, the important feature of this result is that the Principle of Insufficient Reason yields not only equal probabilities but also differently shaped probability distributions. The question is, however, whether the result does indeed follow from our axioms. We are going to show that this is the case.

For brevity we refer only to the special case illustrated in Figure I and assume that a certain action A, leading to a particular result vector,

$$R = \begin{bmatrix} S_1 & S_2 & \cdots & S_n \\ r_1 & r_2 & \cdots & r_n \end{bmatrix},$$

is chosen. We shall consider several subdivisions of this vector that are figuratively represented by the complete bushes below the forkings $a, b, \ldots, g$. The subdivisions are indicated by the letter labeling the corresponding forking.

Our demonstration starts with forking a and the corresponding vector a, consisting of the elements r_1, r_2, and r_3. According to the result of the previous section, we can assess equal probabilities to all branches below forking a without changing the evaluation of vector a. We proceed analogously with forkings b, c, d, e, and f and call the result vectors that are assigned objective probabilities in this way a', $b', \ldots, f'$. Without already integrating these vectors at this stage into the tree diagram, we now look at forking g and regard it as consisting of the elements a, b, and c. Thereby we again have a problem of the kind considered in the previous section, for there we did not place any restrictions on what the subresults r_{ij} are. Thus, we can assess equal probabilities for all branches below g (each 1/3). Analogous results can be obtained for the branches below h and i. Now we replace elements a, b, and c by a', b', and c' within vector g in a stepwise procedure, referring to the Axiom of Independence:

$$\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ a & b & c \end{bmatrix} \sim \begin{bmatrix} 1/3 & 2/3 \\ a & \begin{pmatrix} 1/2 & 1/2 \\ b & c \end{pmatrix} \end{bmatrix} \sim \begin{bmatrix} 1/3 & 2/3 \\ a' & \begin{pmatrix} 1/2 & 1/2 \\ b & c \end{pmatrix} \end{bmatrix}$$

$$\sim \begin{bmatrix} 1/3 & 2/3 \\ b & \begin{pmatrix} 1/2 & 1/2 \\ a' & c \end{pmatrix} \end{bmatrix} \sim \begin{bmatrix} 1/3 & 2/3 \\ b' & \begin{pmatrix} 1/2 & 1/2 \\ a' & c \end{pmatrix} \end{bmatrix} \sim \begin{bmatrix} 1/3 & 2/3 \\ c & \begin{pmatrix} 1/2 & 1/2 \\ a' & b' \end{pmatrix} \end{bmatrix}$$

$$\sim \begin{bmatrix} 1/3 & 2/3 \\ c' & \begin{pmatrix} 1/2 & 1/2 \\ a' & b' \end{pmatrix} \end{bmatrix} \sim \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ a' & b' & c' \end{bmatrix} \equiv g'. \tag{19}$$

The transformed vector is called g'. Analogously we transform h to h' and i to i'. We then proceed as follows: At first we assess equal probabilities for the branches below j and then replace g, h, and i by g', h', and i'. Hence for the example of Figure I, the Principle of Insufficient Reason has been meaningfully utilized for the assessment of probabilities in tree diagrams. We forgo the pure mechanical work of a generalization for arbitrary tree diagrams and state the following:

In case studies for the states of the world, whenever no subcase is more plausible than another, then each subcase must be assigned an equivalent objective probability equal to the reciprocal value of

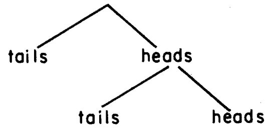

FIGURE II

the number of subcases. The probability of a certain state of the world is then the product of the (conditioned) probabilities of all cases and subcases that have to be distinguished in order to define this state.

VI. CRITICISM

In this section a well-known criticism of the classical Principle of Insufficient Reason is scrutinized for its applicability to our results. A coin is thrown twice. What is the probability that tails come up both times? If we distinguish the states of the world "tails, tails" and "not: tails, tails," then the probability sought is 1/2. If, however, we distinguish the states "tails, tails," "tails, heads," "heads, tails," and "heads, heads" then the probability is 1/4, a contradiction. Here, the correct solution is obvious, but ascertaining the probability of getting tails at least once can be more confusing. Accordingly, D'Alembert (according to Todhunter [1865]), the nightmare of classical mathematics, argues that if "heads" comes up with the first throwing, a second throwing is superfluous. For this reason the states "heads," "tails, heads," and "tails, tails" should be distinguished, and the probability sought is 2/3 instead of 3/4, the correct probability.

These examples lead us to the problem of which states of the world have to be distinguished, as was already clearly discussed by von Kries [1886]. Obviously, a calculation of objective probabilities according to the Principle of Insufficient Reason demands correctly distinguished states of the world. In the light of classical probability theory, this is a very important problem that was unfortunately never satisfactorily solved. However, our results are only slightly affected, for we sought subjective probabilities rather than objective ones, although, of course, the former have the form of equivalent objective probabilities. In order to make the point very clear: If D'Alembert does not see any reason why one of his three cases is more plausible than the others, he should indeed assess probabilities of 1/3 for each.

This, however, does not mean that there is no reason. Had D'Alembert considered the next tree diagram (Figure II), he would

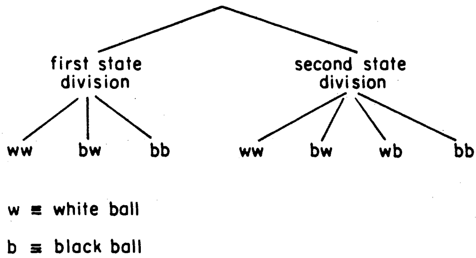

FIGURE III

have found that no one branch is more plausible than the others and thus would have calculated the correct probabilities 1/2, 1/4, 1/4.

A problem closely related to D'Alembert's mistake was presented by Savage [1954, p. 65]: The decision maker knows several possible state divisions but does not know which is the right one. In this case the Principle of Insufficient Reason seems to fail, for different probabilities can be calculated for a special event. Consider Savage's example. Two balls are drawn from an urn that is known to contain either two white balls, two black balls, or a white ball and a black ball. If we regard these three possibilities as the states of the world, the probability of, for instance, drawing a white and a black ball is 1/3. For Savage, however, it also seems possible to distinguish the states "white, white," "black, black," "black, white," and "white, black," so that the probability in question is 1/2. Fortunately, we can help Savage. If he does not know any reason why one state division is more likely than the other, he may refer to the third tree diagram (Figure III) and assess the probabilities according to the rule developed in the previous section. As a result, he will obtain an equivalent objective probability of 5/12 for drawing a white and a black ball.

UNIVERSITIES OF MANNHEIM AND WESTERN ONTARIO

REFERENCES

Allais, M., "Fondements d'une théorie positive des choix comportant un risque," in *Colloques internationaux du centre national de la recherche scientifique,* XL, *Econométrie,* Paris 12–15 May 1952, published in Paris 1953a, pp. 127–40.

——, "Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'école américaine," in *Colloques internationaux du centre national de la recherche scientifique,* XL, *Econométrie,* Paris 12–15 May 1952, published in Paris, 1953b, pp. 257–332.

——, "Le Comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école américaine," *Econometrica,* XXI (1953c), 503–46.

Bernoulli, J., *Wahrrscheinlichkeitsrechnung,* third and fourth parts, Ostwald,

Klassiker der exakten Wissenschaften, Vol. 108 (Leipzig, 1896), translation of *"ars conjectandi,"* published in 1713.
Brewer, K. R. W., "Decisions under Uncertainty," this *Journal,* LXXVII (1963), 159–61.
Chernoff, H., "Rational Selection of Decision Functions," *Econometrica,* XXII (1954), 422–43.
de Finetti, B., "La prévision: Ses lois logiques, ses sources subjectives," *Annales de l'institut Henri Poincaré,* Receuil de conférences et mémoires de calcul des probabilités et physique théoretique, VII (Paris, 1937).
de Laplace, P. S., *Théorie analytique des probabilités* (Paris, 1814).
Ellsberg, D., "Risk, Ambiguity, and the Savage Axioms," this *Journal,* LXXV (1961), 643–69.
Friedman, M., and L. J. Savage, "The Expected-Utility Hypothesis and the Measurability of Utility," *Journal of Political Economy,* LX (1952), 463–74.
Georgescu-Roegen, N., "Choice Expectations and Measurability," this *Journal,* LXVIII (1954), 503–32.
Hirshleifer, J., "Investment Decision Under Uncertainty: Choice-Theoretic Approaches," this *Journal,* LXXIX (1965), 509–36.
Krelle, W., *Präferenz- und Entscheidungstheorie* (Tübingen, 1968).
Luce, R. D., and H. Raiffa, *Games and Decisions, Introduction and Critical Survey* (New York, London, and Sydney, 1957).
Marschak, J., "Rational Behavior, Uncertain Prospects, and Measurable Utility," *Econometrica,* XVIII (1950), 111–41.
Milnor, J., "Games Against Nature," in R. M. Thrall, C. H. Coombs, and R. L. Davis, eds., *Decision Processes* (New York and London, 1954), pp. 49–59.
Niehans, J., "Zur Preisbildung bei ungewissen Erwartungen," *Schweizerische Zeitschrift für Volkswirtschaft und Statistik,* LXXXIV (1948), 433–56.
Raiffa, H., "Risk, Ambiguity, and the Savage Axioms: Comment," this *Journal,* LXXV (1961), 690–94.
Ramsey, F. P., "Truth and Probability," in *The Foundations of Mathematics and Other Logical Essays* (London, 1931), pp. 151–98.
Roberts, H. V., "Risk, Ambiguity, and the Savage Axioms: Comment," this *Journal,* LXXVII (1963), 327–36.
Samuelson, P. A., "Utilité, préférence et probabilité," in *Colloques internationaux du centre national de la recherche scientifique,* XL, *Économétrie,* Paris 12–17 May 1952, published Paris 1953, pp. 140–50.
Savage, L. J., "The Theory of Statistical Decision," *Journal of the American Statistical Association,* XLVI (1951), 55–67.
——, "Une axiomatisation de comportement raisonnable face à l'incertitude," *Colloques internationaux du centre national de la recherche scientifique,* XL, *Économétrie,* Paris 12–17 May 1952, published Paris, 1953, pp. 29–33.
——, *The Foundations of Statistics,* (New York and London, 1954).
Schlaifer, R., *Probability and Statistics for Business Decisions. An Introduction to Managerial Economics under Uncertainty* (New York, Toronto, and London, 1959).
Schneeweiß, H., "Spieltheoretische Analyse des Ellsberg-Paradoxon," *Zeitschrift für die gesamte Staatswissenschaft,* CXXIV (1968), 249–56.
Todhunter, I., *A History of the Mathematical Theory of Probability: From the Time of Pascal to That of Laplace* (Cambridge, 1865), cited from D'Alembert, *Croix ou Pile,* Encylopédie ou Dictionaire Raisonné . . . 1754.
von Kries, J., *Die Principien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung* (Freiburg, 1886), esp. pp. 1–23.
Wald, A., "Statistical Decision Functions Which Minimize the Maximum Risk," *Annals of Mathematics,* XLVI (1945), 265–280.