# 3 Rational Choice

"Though the 'Look before you leap' principle is preposterous if carried to extremes, I would none the less argue that it is the proper subject of our further discussion, because to cross one's bridges when one comes to them means to attack relatively simple problems of decision by artificially confining attention to so small a world that the 'Look before you leap' principle can be applied there." Leonard J. Savage[1].

## 3.1 Decision Making and the Economics Paradigm

The basic ingredient in modern economics is that individuals have well-defined goals, and that they make choices consistent with those goals given the information and resource constraints they face. This approach can be traced back to Hobbes (1651)'s great work *Leviathan*, which begins with a description of the qualities of man. The genius of his argument is that one takes the qualities of man as given, and then derives the role of government and the sovereign in creating civil society. The empirical content of economic models arises from exploring how behavior varies with changes in the resources available. As Becker (1976) has argued, this approach has been a very useful way to organize and think about a large volume of socio-economic phenomena, and is, in the sense of Kuhn (1970), a leading paradigm for social sciences.

For the purposes of this book we are interested in decision theory to the extent that it can *represent* behavior. In other words, contract theory needs a model of behavior that allows one to predict how individuals respond to different environments. Given this representation, one can then ask how to design a contract with certain desirable properties. We shall show that the rational choice model provides a very simple and elegant representation of behavior that is sufficient for a wide variety of problems. Just as the low resolution maps discussed in Chapter 2do not perfectly represent the a area on the ground, neither will the rational choice model perfectly represent human behavior.

More importantly, there does not exist any other model of human behavior that is *comprehensive* in the sense that it can be applied in any context. The benefit of a comprehensive model is that it can form a benchmark against which more domain specific models of behavior can

1 Savage (1954), page 16.

be evaluated. This is the approach pioneered by Kahneman and Tversky (1979). Their work lays the foundations for behavioral economics which organizes observed behavior as *deviations* from the rational choice model. When individual choices are assumed to satisfy some well-defined properties, then one can explicitly test for these properties. When the failures of the rational choice model are systematic, then this allows one to modify the existing theory to build a better descriptive model.

However, this approach can also lead to confusion that can result in a mixture of both positive and normative analysis. For example, when Leonard Savage, one of the founders of modern decision theory, was informed that he violated his own axiom for rational choice, he promptly responded that he would change his answer to be consistent with the theory.[2] This illustrates that individuals do not necessarily take actions that are consistent with "rational choice" - if one of the founders of the theory cannot do this, what hope have the rest of us? By deriving the actions that a rational person would follow, the theory has a strong normative element - individuals should follow what the theory prescribes.[3] When the failures of the theory are systematic, then one can extend the model to produce a more descriptive theory.

Thus, regardless of one's perspective, all economic models of human behavior begin with the rational choice model. This chapter provides a review of this model and emphasizes its usefulness as a good first order *representation* of behavior. The next section makes precise the distinction between behavior and preference. This distinction is analogous to Arrow (1958)'s observation that psychologists distinguish between choice and decision. The former is what individuals actually do in some situations, while the later is the result of a conscious process of reflection. From this perspective the economist's notion of rational choice can be viewed as a parsimonious (and very useful) model of choice.

Section 3 then extends the basic model to address the issue of decision making in the face of risk - where risks are events with well-defined probabilities of occurring. When individuals are averse towards risk this generates a demand for insurance contracts. Incorporating risk into the

2 See the discussion in Machina (1982) on page 289.

3 There is a literature that explores the impact of economics education upon behavior. See for example Frank et al. (1993)and Ferraro et al. (2005).

model of decision making is essential for understanding the role of insurance contracts. Frank Knight (1921) introduced the distinction between "risk" and "uncertainty". He used these terms to distinguish between events for which past experience would allow us to have a reliable estimate of the probability of the event occurring in the future.

By "uncertainty" he meant events for which there is insufficient experience to accurately assess the likelihood of the event. The issue then is how a business person who is launching a new product whose success is uncertain should make a "rational" choice. Section 4 outlines Savage's theory of decision making which explicitly addresses this issue and shows that one can indeed have a model of rational choice, even in the face of uncertainty. This theory forms the foundation for the theory of strategic choice that we discuss in some detail in chapter 4.

## 3.2 Behavior and the Rational Choice Model

The rational choice model is based upon the idea that the pursuit of well-defined goals provides a way to describe behavior in a wide variety of situations. To see why. let us begin with an abstract, but general, model of individual behavior. As a matter of convention the word individual is used to denote a person, firm, or any other decision making entity that may be pursuing well-defined goals. Each period $t$ the individual is faced with a menu of actions, say $A_t$, from a set of possible actions $A$. The precise form that $A_t$ takes depends upon resource and information constraints. For example in the standard consumer choice problem, $A$ is the set of possible consumption choices and $A_t$ is the set of consumption bundles the individual can afford in period $t$. For example, when faced with a fruit bowl, should one choose an apple, banana or orange?

Formally, a decision taken in period $t$ is denoted by $d_t \in A_t$. If we observe $d_t$ then the individual has, in the sense of Samuelson (1938), revealed a preference for $d_t$ over the other actions in the set $A_t$. An immediate difficulty is that the individuals may be indifferent over several of the options in $A_t$, and in this case the individual is in fact choosing randomly from a set of acceptable decisions that can be denoted $D_t$. The possibility of indifference is an essential ingredient for some of the contract models described later in this book. The reason for this is that in some situations one may ask a party to reveal information. In order to ensure truthful revelation, the contract must ensure that the individual has no incentive to bias her report.

An individual's behavior is formally described by a function $D_t = B(A_t, t)$ that defines the set of acceptable decisions at date $t$ given the set of feasible choices $A_t$. In order for the model to have some empirical bite we need to add some additional restrictions. The first of these is that an individual makes a decision at each date, and hence $B(A_t, t) \neq \emptyset$. This may seem straightforward, but it is not an innocuous assumption. In particular, it implies that an individual acts even if she does not have all the information or knowledge necessary to make a decision.

Second, we need some hypothesis to connect data collected at one point in time with actions at another point in time. The two most common assumptions to achieve this are unit homogeneity and time invariance. In the context of choice, unit homogeneity can be captured by supposing that individuals with similar characteristics have similar preferences. At this stage we have no basis for such a hypothesis. For much of the analysis in this book we shall assume *time invariance* - namely a person's behavior as defined by $B$ does not change over time and is given by the function[4]:

$$B : 2^A \backslash \{\emptyset\} \times T \to 2^A \backslash \{\emptyset\}. \tag{3.1}$$

Thus at date $t$ we suppose that the observed decision is given by some $d_t \in D_t = B(A_t)$. The assumption that behavior is time invariant does not imply that behavior does not change due to the acquisition of new information. Rather it is assumed that information changes the payoffs associated with different actions, and hence the set $A_t$.

The theory of rational choice builds upon the idea that one can learn about an individual's behavior from the way an individual makes *binary* choices. Namely, for any pair $\{x, y\} \in A \times A$, the individual is said to prefer $x$ over $y$, written $x \succsim y$, if and only if $x \in B(\{x, y\})$. A person is said to be *rational* if they also satisfy the following definition:

*Axiom* A preference relation $\succsim$ defined on $A \times A$ is *rational* if:

4 The notation $2^A$ denotes the set of functions mapping from $A$ to the set $\{0, 1\}$. Each function corresponds to the subset $B$ of $A$ where the function takes on the value of 1 (and hence has a value of 0 on the complement, $B^c$). Such functions are called *correspondences* when viewed as a mapping from $2^A$ to $A$.

1. It is complete, for every $x, y \in A$, $x \succsim y$ or $y \succsim x$.
2. It is transitive, for every $x, y, z \in A$, $x \succsim y$ and $y \succsim z$ implies $x \succsim z$.

Completeness ensures that it is always possible to rank two alternatives, and hence the individual is always capable of making a decision. This is equivalent to supposing $B(\{x, y\}) \neq \emptyset$. The condition of transitivity is at the core of the theory of rational choice, and is the condition that allows one to deduce how individuals are likely to behave when faced with any number of alternatives. Thus if we are given three alternatives $x, y$ and $z$, and $x \succsim y$ and $y \succsim z$, then we conclude that $x \succsim z$. This means that given the set $\{x, y, z\}$ the individual would choose $x$ over both $y$ and $z$.

When $\{x, y\} = B(\{x, y\})$, then $x \succsim y$ and $y \succsim x$ and the individual is *indifferent* between $x$ and $y$, which is written $x \sim y$. An alternative, $x$ is strictly preferred to $y$, is written $x \succ y$, if $x \succsim y$, but not $y \succsim x$. Preferences can be used to define behavior for any set $A_t$ by supposing individuals choose the most preferred outcome available to them. More formally, an individual's behavior given preferences $\succsim$ is defined by:

$$B(A', \succsim) = \{x | x \succsim y \text{ for every } y \in A'\}. \tag{3.2}$$

The agent chooses actions that are preferred to all other actions in the feasible set.

The axiom of transitivity ensures that this set is always well-defined. For example, suppose the choices are $s, v$ and $c$, corresponding to strawberry, vanilla and chocolate ice cream respectively. If $v \succ c \succ s \succ v$, then an agent faced with the set $\{s, v, c\}$ would never be able to make up her mind because no matter which choice she makes, there is always a superior alternative, and hence $B(\{v, c, s\}, \succsim) = \emptyset$. Thus the requirement of transitivity is intimately linked to the requirement that individuals are able to make a decision when the number of alternatives is greater than two. When preferences are transitive, not only does this ensure that $B(A', \succsim)$ is never empty, it also implies that observed behavior is consistent with the individual maximizing some *utility function.*

**Proposition 3.1 .** *Suppose $A$ is a finite set, if preferences, $\succsim$, are rational then:*

1. $B(A', \succsim) \neq \emptyset$ for all $A' \subseteq A$.
2. There exists a function $U : A \to \Re$, called a *utility function* such that $x \succsim y$ if and only if $U(x) \geq U(y)$ and $B(A', \succsim) = \arg\max_{x \in A'} U(x)$.

*Proof.* See exercise 3. □

This result shows that when behavior is described by rational preferences, then the individual behaves as if she maximizes utility. In this theory, the values that a utility function provide are used only to make comparisons among alternatives, and do not have any independent normative value in and of themselves. In this case with a finite number of alternatives, a completely equivalent way to represent preferences is with a list $\{P_1, P_2, ..., P_n\}$ where $P_i \subseteq A$, and all $x \in P_i$, $y \in P_j$ satisfy $x \succsim y$ if and only if $i \geq j$. In social choice theory it is quite common to suppose that preferences do not allow for any indifference, in which case they are simply represented by a list $\{x_1, x_2, ..., x_n\}$, rather than with a utility function. In particular, utility does not correspond to any notion of happiness, it is simply a way to rank alternatives. The prisoner may be very miserable, but she will likely choose drinking water over no water, even if the impact on her overall well being may be small.

### 3.2.1 Is the rational choice model restrictive?

The reason we began with a general model of behavior, rather than the more traditional approach of defining preferences immediately, is to highlight the fact that rational choice theory provides a compact *representation* of behavior. In this section we make this idea more explicit by asking how much information is needed to represent any possible behavior relative to behavior described by rational preferences.

Suppose that there are $n$ alternatives in $A$. For simplicity restrict attention to behavior with the property that for any set $A_t \subset A$, the individual makes a unique choice, and hence $B(A_t)$ is a function from $2^A \backslash \emptyset$ to $A$. In particular, this implies that the individual is never indifferent between two choices, and thus only strict orders on $A$ need to be considered (for any $\succsim$ and $x, y \in A$, $x \neq y$ then $x \succ y$ or $y \succ x$). The number of such rankings can be computed by observing there can be $n$ top ranked alternatives, $n-1$ second ranked alternatives and so on. Hence the total number of possible rankings is $n!$. Thus with $n$ alternatives, there are at most $n!$ different behaviors that are consistent with the rational choice hypothesis.

Now consider the different types of behavior that are possible that do not necessarily conform to the rational choice hypothesis. Here we are explicitly allowing *framing effects.* There is a great deal of evidence from behavioral economics that choices can depend upon the other options in the choice set. For example, given a set $\{a, b\}$ a person might choose $a$, but when offered the set $\{a, b, c\}$ chooses b. Such behavior is clearly inconsistent with rational choice (why?), but certainly in the range of possibility. For example, the choice might be a customer choosing between suits. Upon viewing suit $c$, it causes the person to re-evaluate her preferences over $a$ and $b$.

Consider now how the potential for framing increases the number of possible different behaviors. Observe that there are $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ different sets $A_t$ of size $k$. For each of these sets the function representing behavior, $B(A_t)$, can take on $k$ values, hence the total number of different possible functions is given by:

$$\prod_{k=2}^{n} k^{\left(\frac{n!}{k!(n-k)!}\right)}.$$

The following table summarizes the results of these computations for $n = 3, 4$ and $5$ :

| n | Number of events | Number of possible different behaviors | Number of behaviors consistent with rational choice |
|---|---|---|---|
| 3 | 7 | 24 | 6 |
| 4 | 15 | $20,736$ | 24 |
| 5 | 31 | $3.0959 \times 10^{11}$ | 120 |

As the number of possible alternatives increases, the rational choice model is extremely restrictive. When there are 5 alternatives, only 0.00000003% of possible behaviors are consistent with the rational choice hypothesis. Given that there are many examples of individuals violating the axioms of rational choice theory, this example illustrates that merely adding framing effects would lead to many more degrees of freedom that would allow one to more accurately represent individual behavior.

However, as discussed in chapter 2, one role of a model is to provide a concise representation of data. The rational choice model provides a representation for individual behavior that, as Becker (1976) argues, has provided a very useful way to think about a wide variety of economic institutions. At the moment there is simply no alternative model of behavior that provides an equally concise representation of behavior that is systematically better. For this reason behavioral economics since the work of Kahneman and Tversky (1979) has been organized around deviations from rational choice.

### 3.2.2 Continuity of Preferences

The fact that in practice individuals make mistakes, forget how they have acted in the past, and so on, implies that rational choice theory is rarely, if ever, a perfect representation of individual behavior. In practice, what one needs is that it is a *useful* model that provides a parsimonious approximation of actual behavior. In order to discuss "approximate" behavior we need to allow for some notion of "close". This section extends the model to allow for a continuum of choices, in order to allow for a meaningful discussion of "approximately" rational choice. The theory for choice over a finite set of elements extends naturally to a continuum by supposing that small variations in quantities have correspondingly small effects upon one's preference for a good. This is captured formally by requiring preferences to be continuous:

**Definition 3.2 .** Suppose that $A$ is a closed subset of $\Re^n$ then a rational preference order, $\succsim$, defined on $A$ is continuous if for any two convergent sequence of decisions, $\{x'_n\}$ and $\{x''_n\}$ in $A$ such that $x'_n \succsim x''_n$ for every $n$, then [5]

$$\lim_{n\to\infty} x'_n = x' \succsim x'' = \lim_{n\to\infty} x''_n. \tag{3.3}$$

Gerard Debreu (1954), Theorem I, proves that this condition combined with rationality is sufficient to ensure the existence of a continuous utility function. The reader is referred to Debreu (1954) for the proof. The statement of the theorem is as follows:

**Theorem 3.3 .** *Suppose that preferences $\succsim$ are continuous and rational on a closed and connected subset $A$ of $\Re^n$, then there exists a (continuous) utility function $U : A \to \Re$, such that for every $x', x'' \in A$, $x' \succsim x''$ if and only if $U(x') \geq U(x'')$.*

For the rest of the book, it is a maintained hypothesis that individual behavior can be well approximated by the rational choice model with

5 Here the limits are defined with the topology induced by the standard Euclidean norm.

continuous preferences. At various points we discuss how this hypothesis may be relaxed. In section 5 of this chapter, a model of boundedly rational choice is presented to explain why in the short run individuals may not choose optimal actions. - though for individuals with sufficient experience, the model predicts that behavior is well modeled by the utility maximization hypothesis. These formal results are consistent with a large body of work that finds that even though individuals may not be perfect optimizers, with learning the hypothesis of utility maximization is a very good first order approximation.

## 3.3 Risk

One motive for writing a contract is the enforcement of insurance contracts. For example, there is a small chance that one's house will burn down in the future, causing great harm to the affected family. However the total number of houses that burn from year to year is relatively stable, and the reasons for these accidents are (hopefully) random. Hence, it is possible for an insurance company to *diversify* the risk, and sell policies to individuals that will reimburse them in the event that their house catches fire. In this case an individual is choosing between facing the risk of a complete loss of the house and a consumption bundle that in effect shares the losses from each individual with the group of homeowners, so that one's net consumption does not depend upon whether or not one's house burns down. The theory of decision making in the face of risk begins with a rational choice model that can represent individual aversion towards risk. From this representation we can compute how much an individual would be willing to pay for insurance.

As in the case of the basic rational choice model, the analysis is greatly simplified if we restrict attention to a finite number of consumption bundles defined by the set $C = \{c_1, \ldots, c_n\}$. For most cases considered, these outcomes represent money payoffs and effort levels by individuals, though they may also represent an allocation of goods, the distribution of income between individuals, or even one's work environment. For example $c_i$ might be one's salary or a bundle of groceries. It may also represent working 60 hours in a week and receiving a paycheck of \$1000 at the end of the week.

Risk is introduced by supposing that the probability of receiving bundle $c_i$ is a potential choice variable. For example, an insurance company can ensure that one never loses income due to a car accident or accidental burning of one's house. Similarly, lottery companies regularly sell tickets that pay off in millions of dollars with a low probability. Formally, the set of *lotteries* over $C$ is denoted by $\Delta(C)$, and defined as the set of probability distributions over $C$:

$$\Delta(C) = \left\{ p = \{p_1, \ldots, p_n\} \in \Re^n | \sum_{i=1}^{n} p_i = 1, p_i \geq 0 \right\}. \tag{3.4}$$

Now suppose that in each period the feasible choice set is a *closed* subset $L_t \subset \Delta(C)$, from which the agent chooses an allocation $B(L_t)$. Suppose that this behavior is represented by a preference relation, $\succsim$, that is continuous and satisfies the axioms of rational choice. From proposition 3.3 there exists a continuous utility function, $U : \Delta(C) \to \Re$, representing these preferences. Without additional structure, general preferences over lotteries do not provide much in the way of testable implications. Our concern here is to find a way to describe aversion towards risk, and to differentiate between the preferences over goods and preferences over risks. Von Neuman and Morgenstern (1944) solve this problem with their celebrated *expected utility theory.*

Before formally deriving this theory, it is useful to consider an early attempt to model behavior when lottery outcomes represent money payoffs. Suppose that $c_i \in \Re$ represents a monetary return. Then an individual is an *expected monetary value maximizer* (EMVer), if preferences are represented by a utility function defined to be the expected monetary value of a trade:

$$l' \succsim l'' \text{ iff } \sum_{i=1}^{n} p_i' c_i \geq \sum_{i=1}^{n} p_i'' c_i. \tag{3.5}$$

For example, suppose that an individual can buy a lottery ticket which pays off \$1 million dollars with a probability of .00005. The EMV of this ticket is .00005×\$1million = \$50. In other words the individual buys such a ticket if and only if its price is less than \$50. Though this is in many respects an appealing criterion, it does not do a good job representing observed individual choice in many situations. Bernoulli (1738) was the first to point this out with the following example: a coin is tossed repeatedly until a head appears, and if there are $n$ tosses before the head appears, then the individual is paid $2^n$ ducats. The EMV of this lottery is:

$$\frac{1}{2} \cdot 2 + \left(\frac{1}{2}\right)^2 \cdot 2^2 + \left(\frac{1}{2}\right)^3 \cdot 2^3 + \cdots = \infty. \tag{3.6}$$

Bernoulli argued individuals are unlikely to offer more than 20 ducats for such a lottery, even though the expected value is infinite. He solved this problem by positing a declining marginal utility for money, and supposing that individuals maximize *expected utility:*

$$U(l) = \sum_{i=1}^{n} p_i u(c_i), \tag{3.7}$$

where $u(x)$ is called a *Bernoulli utility.* In this example if one supposes $u(x) = \log(x)$ ,then the utility of this lottery is:

$$\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \log(2^n) = \log(2)(\sum_{n=1}^{\infty} \frac{n}{2^n}) = log(4). \tag{3.8}$$

Thus, in this case the individual would be willing to pay at most 4 ducats for this lottery, much less than the expected monetary value of this lottery. Von Neuman and Morgenstern (1944) provide an *axiomatization* for Bernoulli's idea.

By axiomatization one means finding a set of principles of behavior that imply a particular utility function. Axioms provide a set of principles of behavior that are both easily interpreted and tested. For example, if $a \succ b$ and $b \succ c$ then we would consider a person irrational or erratic if he chooses $c$ when offered a choice set $\{a, b, c\}$.

In the context of risky decisions, the first normative principle is that probabilities are correctly interpreted. This idea is formalized by requiring individuals to correctly evaluate compound lotteries. A compound lottery is denoted by $\left\{\alpha_1, l^1; \alpha_2, l^2; ...; \alpha_n, l^n\right\}, \alpha_i \geq 0, \sum \alpha_i = 1$, and corresponds to running a two stage lottery in which a number $m \in \{1, ..., n\}$ is drawn with probability $\alpha_m$, after which the lottery $l^m$ is played. If these two lotteries are independent, then the probability that the individual receives consequence $c_i \in C$ is $\sum_m \alpha_m p_i^m$. A rational person should be indifferent between the compound lottery and the simple lottery paying $c_i$ with probability $\sum_m \alpha_m p_i^m$, for each $i = 1, ..., n$. Formally we have:

**Definition 3.4 .** An individual with preferences over simple lotteries $\Delta(C)$ satisfies *reduction of compound lotteries* if she is indifferent between the compound lottery $\left\{\alpha_1, l^1; \alpha_2, l^2; ...; \alpha_n, l^n\right\}$ and the simple lottery $\{\sum_m \alpha_m p_1^m, \sum_m \alpha_m p_2^m, ..., \sum_m \alpha_m p_n^m\} \in \Delta(C)$.

The implicit assumption is that the compound lottery is carried out in a sufficiently short time span that time preferences do not matter. The substantive content of this axiom is that individuals are able to use the rules of probability to judge outcomes, an issue that is discussed in more detail later in this chapter. The next axiom states a substantive restriction upon preferences that does not follow from any principle of rationality:

**Definition 3.5 .** The preference relationship, $\succsim$, on $\Delta(C)$ satisfies *independence* if for all lotteries $l, l', l'' \in \Delta(C)$ and for every $\alpha \in (0,1)$

$$l \succsim l' \text{ if and only if } \{\alpha, l; (1-\alpha), l''\} \succsim \{\alpha, l'; (1-\alpha), l''\}. \tag{3.9}$$

This axiom requires that when one prefers lottery $l$ over $l'$, then combining each of these with another lottery, $l''$, does not change the ranking. Suppose that $L_1 \sim L_2$, then notice that for all $\alpha \in (0,1)$

$$L_1 \sim \{\alpha, L_1; (1-\alpha), L_1\} \sim \{\alpha, L_1; (1-\alpha), L_2\} \sim \{\alpha, L_2; (1-\alpha), L_2\} \sim L_2.$$

Figure 3.1 illustrates this axiom when there are three consequences satisfying $c_1 \succ c_2 \succ c_3$, and provides a pictorial representation of the simplex $\Delta\{c_1, c_2, c_3\}$. The length of each side of the equilateral triangle is $2/\sqrt{3}$, and hence for any point in the triangle the sum of the distances to each side is equal to 1. Therefore each point in the triangle represents a lottery where the probability of choosing $c_i$ is the distance to the line between the other two consequences. For example, the point of the triangle indexed by $c_i$ corresponds to choosing $c_i$ with probability 1. Given that $L_1 \sim L_2$ then the reduction of compound lotteries and the independence axioms imply that an individual would be indifferent between any points on the straight line between $L_1$ and $L_2$.

If preferences are required to satisfy these axioms in addition to the standard axioms of rational choice, one has a generalization of Bernoulli's idea, namely that preferences can be represented by expected utility.

**Proposition 3.6 .** *Suppose that preferences $\succsim$ over $\Delta(C)$ are rational, continuous, and satisfy reduction of compound lotteries and independence, then there is a function $u : C \to \Re$ (called a Bernoulli utility*

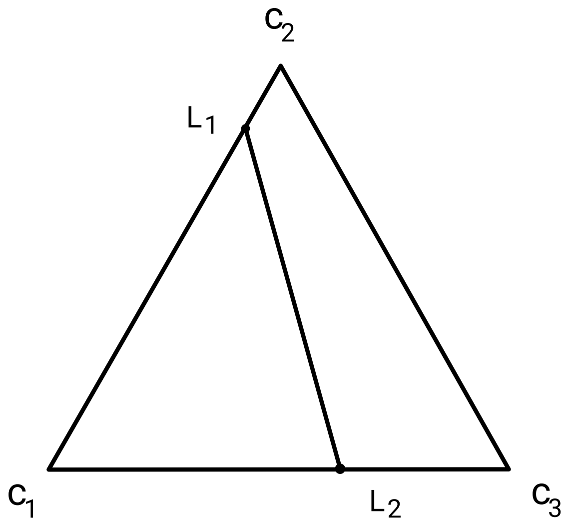


**Figure 3.1**
Lottery Simplex

*function) such that for all* $l = (p_1, ..., p_n) \in \Delta(C)$ *the function* $U$ *defined by:*

$$U(l) = \sum_{c_i \in C} u(c_i) p_i, \tag{3.10}$$

*is a utility function representing the preferences* $\succsim$.

*Proof.* (sketch) Given that preferences are continuous, then there exists a continuous utility function representing $\succsim$. The continuity of utility and the compactness of $\Delta(C)$ implies that there exists lotteries, $l_{\min}$ and $l_{\max}$, the least and most preferred lotteries respectively.

If $l_{\min} \sim l_{\max}$ then the individuals are indifferent over all lotteries, and we can set $u(c_i) = \bar{u}$,where $\bar{u}$ is some constant, and we are done.

Suppose now that $l_{\max} \succ l_{\min}$, and observe that the independence hypothesis implies that for all $\alpha \in (0,1)$ that $l_{\max} \succ \{\alpha, l_{\max}; (1-\alpha), l_{\min}\} \succ l_{\min}$. From this result and independence again it follows that:

$$\{\alpha, l_{\max}; (1-\alpha), l_{\min}\} \succ \{\beta, l_{\max}; (1-\beta), l_{\min}\} \tag{3.11}$$

for all $\alpha, \beta \in [0,1]$, $\alpha > \beta$. To see this let $l = \{\alpha, l_{\max}; (1-\alpha), l_{\min}\}$, and let $\gamma = \beta/\alpha \in (0,1)$, then

$$l \sim \{\gamma, l; (1-\gamma), l\} \succ \{\gamma, l; (1-\gamma), l_{\min}\} \sim \{\beta, l_{\max}; (1-\beta), l_{\min}\}.$$

By continuity, this implies that for every $l \in \Delta(C)$ there is a number, denoted $U(l) \in [0,1]$ such that $l \sim \{U(l), l_{\max}; (1-U(l)), l_{\min}\}$. It readily follows from the above that $U(l)$ is a utility function representing $\succsim$.

In the final step observe that 3.11 and reduction of compound lotteries imply that for any $\alpha \in [0,1]$ and $l', l'' \in \Delta(C)$ then $U(\{\alpha, l'; (1-\alpha), l''\}) = \alpha U(l') + (1-\alpha) U(l'')$. In particular this implies that we can let $u(c_i) = U(c_i)$ and rewrite utility in the expected utility form: $U(l) = \sum_{c_i \in C} u(c_i) p_i$. □

One of the real benefits of an axiomatization is that it makes expected utility theory readily testable. One can present subjects with a finite set of choices for which the axioms make clear predictions. A famous example is the *Allais Paradox:*[6] Allais asked a number of individuals to consider the following choice problem. Suppose that there are three outcomes:

A lottery is then given by triples of the form $\{p_1, p_2, p_3\}$, where $p_i$ is the probability of receiving $c_i$. One way of generating lotteries is to suppose that there are 100 lottery tickets, upon which $c_1$, $c_2$ or $c_3$ is written. An individual has an equal probability of receiving one of these tickets. Let us now consider the choice between different lotteries that allocate the payoffs among the tickets as follows:

From Table 3.1, situation 1 corresponds to the choice between lottery $L_1 = \{0,1,0\}$ and $L_2 = \{0.1, 0.89, 0.01\}$, where $c_i = \{\$0, \$0.5\text{million}, \$2.5\text{ million}\}$. Situation 2 differs from situation 1 only in that the payment of \$0.5 million for tickets $12-100$ is replaced by no payment. Since the only difference between the two situations is the payoffs for tickets $12-100$, then the independence axiom implies that any person choosing $L_1$ over $L_2$, must also choose $L_3$ over $L_4$. In experiments it is often observed that individuals rank $L_1$ above $L_2$, while ranking $L_4$ above $L_3$.

6 The discussion here is taken from Savage (1954).

| | | Ticket Number | | |
|---|---|---|---|---|
| | | 1 | 2-11 | 12-100 |
| Situation 1: | $L_1$ | \$0.5 million | \$0.5 million | \$0.5 million |
| | $L_2$ | \$0 | \$2.5 million | \$0.5 million |
| Situation 2: | $L_3$ | \$0.5 million | \$0.5 million | \$0 |
| | $L_4$ | \$0 | \$2.5 million | \$0 |

**Table 3.1**
Example of Independence Axiom

These observations imply that this theory does not in general correctly describe choices of many individuals in the face of risk, and hence in the spirit of Popper (1963) the theory has been refuted. Despite this, expected utility theory is the work horse model of modern microeconomics. The reason is that the model is a useful *representation* of behavior. As we observed in chapter 2, a useful model is one that provides a parsimonious representation of behavior, not necessarily a perfect representation. Harless and Camerer (1994) and Hey and Orme (1994) have reviewed a great deal of experimental evidence and have found that although expected utility theory is not strictly correct, it is broadly consistent with observed behavior.

There is active research in the theory of choice that explores ways to relax the axioms of expected utility, while maintaining the basic axioms of transitivity, completeness and continuity (see the review by Epstein (1992)). As with the case of behavioral economics, a great deal can be learned by exploring the implications of the theory given expected utility theory, while remaining open to the possibility that some of the failures of the theory may be due to the way preferences are represented.

### 3.3.1 Risk Aversion

In this section we discuss the problem of how to evaluate or measure a person's willingness to bear risk. This discussion is typically carried out within the context of monetary rewards, and hence we let the agent's choice set be given by $C = (-\infty, \infty) = \Re$. The expected utility model extends naturally to incorporate preferences over random variables, $X$, with payoffs in $\Re$. Under the von Neumann-Morgenstern axioms the individual's utility function has the form:

$$U(X) = E\{u(X)\} = \int_{-\infty}^{\infty} u(x)\, dF(x), \tag{3.12}$$

where $u : \Re \to \Re$, is a Bernoulli utility function, $F(\centerdot) \in \Delta(\Re)$ is the cumulative probability density function for $X$, and $E\{\cdot\}$ the expectations operator. Demand for insurance is modeled by assuming individuals prefer sure outcomes to risky outcomes, a characteristic called risk aversion:

**Definition 3.7 .** An individual with preference ordering $\succsim$ over random monetary payoffs is *risk averse* if for every $X \in \Delta(\Re)$:

$$E\{X\} \succsim X. \tag{3.13}$$

An individual is strictly risk averse if this preference is strict whenever $var\{X\} > 0$. The individual is *risk neutral* (an expected monetary value maximizer) if:

$$E\{X\} \sim X. \tag{3.14}$$

When preferences over monetary outcomes satisfy the von Neumann-Morgenstern axioms, then it is possible to obtain a complete characterization of risk aversion in terms of the individual's Bernoulli utility function.

**Proposition 3.8 .** *An individual whose preferences are represented by expected utility is risk averse if and only if her Bernoulli utility function is concave. Moreover, she is strictly risk averse if and only if her Bernoulli utility function is strictly concave.*

*Proof.* Suppose the Bernoulli utility function is concave, then by Jensen's inequality one has:

$$U(X) = E\{u(X)\} \leq u(E\{X\}), \tag{3.15}$$

from which we conclude that the agent is risk averse. Conversely, suppose the preference order satisfies risk aversion, but $u$ is not concave. Then there exists an $x$ and $y$ such that:

$$\alpha u(x) + (1-\alpha)\, u(y) > u(\alpha x + (1-\alpha)\, y). \tag{3.16}$$

Let $X$ denote the lottery that pays $x$ with probability $\alpha$ and $y$ with probability $(1-\alpha)$, then it follows that $X \succ E\{X\}$, a contradiction. $\square$

**Example 3.9 .** Consider an individual who wishes to buy home earthquake insurance, and whose preferences are described by a strictly concave and increasing Bernoulli utility function $u(\cdot)$. Let $I > 0$ denote the individual's income, $\rho$ the probability of an earthquake and $d$ the damage sustained to the house in the event of an earthquake. Let the price of earthquake insurance be $q$ per dollar of coverage. The amount of insurance that the individual will buy is a solution to the following problem:

$$\max_{x \geq 0} (1-\rho) u(c_1) + \rho u(c_2),$$

subject to:

$$c_1 = I - qx, \tag{3.17}$$

$$c_2 = I - d + x - qx. \tag{3.18}$$

The amount that the agent consumes when there is an earthquake is $c_2$, while $c_1$ is consumption when there is no earthquake. Notice that the first order condition for this problem is:

$$\frac{u'(c_1)}{u'(c_2)} = \frac{\rho}{1-\rho}\frac{1-q}{q}. \tag{3.19}$$

Thus the individual buys full insurance if and only if $\rho = q$, that is the price per unit of insurance is equal to the probability of an earthquake. Notice that at this price the expected profit of the insurance company is:

$$profit = income - outlay \tag{3.20}$$

$$= \rho x - \rho x = 0. \tag{3.21}$$

This price is called *actuarially fair* since it ensures that a risk neutral insurance company earns zero profits. When the price $q$ is higher than $\rho$, then the individual under-insures ($c_1 > c_2$), while conversely we get over-insurance when the price is lower than the actuarially fair price.

When a person is risk averse he or she is willing to accept an amount for a lottery $X$ that is less than its EMV. This amount, called the *certainty equivalence,* is denoted by $CE(X)$, and is defined by:

$$u(CE\{X\}) = E\{u(X)\}. \tag{3.22}$$

Under the hypothesis that the Bernoulli utility function is increasing, then the certainty equivalence also defines a utility function for preferences: $X \succsim Y$ if and only if $CE(X) \geq CE(Y)$. It has the attractive property that if $X = \bar{x}$ is a constant amount then $CE(X) = \bar{x}$, and thus the certainty equivalence provides a way to measure utility in dollar terms. When $u(\cdot)$ is twice differentiable one can obtain an approximate measure of the certainty equivalence as follows. Let $\bar{x} = E\{X\}$, and consider the Taylor expansion for $u$ :

$$u(x) \approx u(\bar{x}) + u'(\bar{x})(x - \bar{x}) + u''(\bar{x})(x - \bar{x})^2/2. \tag{3.23}$$

Substituting $X$ for $x$ and taking expectations, then when the variance of $X$, $\sigma^2$, is small we have:

$$E\{u(X)\} \approx u(\bar{x}) + u''(\bar{x})\sigma^2/2. \tag{3.24}$$

Using only the first two terms of the Taylor expansion we also have:

$$u(CE\{X\}) \approx u(\bar{x}) + u'(\bar{x})(CE\{X\} - \bar{x}) \tag{3.25}$$

from which we can conclude that:

$$CE(X) \approx \bar{x} - \frac{\sigma^2}{2} r(\bar{x}) \tag{3.26}$$

where $r(\bar{x}) \equiv -\frac{u''(\bar{x})}{u'(\bar{x})}$ is called the *coefficient of absolute risk aversion,* and $\frac{\sigma^2}{2} r(\bar{x})$ is the risk premium, the amount that an individual is willing to give up to avoid the risk in the lottery $X$. Under the hypothesis that the individual strictly prefers more money to less ($u' > 0$), the coefficient of absolute risk aversion is greater than or equal to zero if and only if the individual is risk averse. The amount that an individual is willing to give up to remove all risk is increasing with the degree of risk aversion, as measured by $r(\bar{x})$, and in the variance of the income stream.

An important class of utility functions are those with *constant absolute risk aversion (CARA),* where $r(x) = r$ for every $x$. In this case the Bernoulli utility must solve the differential equation:

$$u''(x) = -ru'(x), \tag{3.27}$$

from which it follows that $u(x) = -ae^{-rx} + b$ if $r \neq 0$, or $u(x) = ax + b$ when $r = 0$, where $a$ and $b$ are indeterminate parameters. In exercise 9 it is shown that von Neumann-Morgenstern preferences are invariant to any affine transformation. Given that $u' > 0$, then without loss of generality it may be assumed that $b = 0$ and $a = 1$ if $r > 0$ or $a = -1$ if $r < 0$. When $r = 0$, then Bernoulli utility can take the form $u(x) = x$. With CARA preferences and a normally distributed income stream $X$, it is possible to obtain a closed form solution for the certainty equivalence $CE(X)$.

**Proposition 3.10 .** *Suppose that an individual has CARA preferences with risk aversion parameter r, then the certainty equivalence of a normally distributed random variable X, with mean m and variance $\sigma^2$ is:*

$$CE(X) = m - \frac{r}{2}\sigma^2. \tag{3.28}$$

*Proof.* By definition we get:

$$\begin{aligned}
-\exp(-rCE(X)) &= E\{-\exp(-rX)\} && (3.29)\\
&= -\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp(-rx)\exp\left(\frac{-(x-m)^2}{2\sigma^2}\right)dx && (3.30)\\
&= -\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(-\frac{x^2-2(m-\sigma^2 r)x+m^2}{2\sigma^2}\right)dx && (3.31)\\
&= -\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left[-\left(\frac{(x-(m-\sigma^2 r))^2}{2\sigma^2}+\frac{2mr-\sigma^2 r^2}{2}\right)\right]dx && (3.32)\\
&= -\exp\left[-r\left(m-\frac{\sigma^2 r}{2}\right)\right] && (3.33)
\end{aligned}$$

□

Notice that this implies that when returns are restricted to be normally distributed, then individuals with $CARA$ preferences can be represented by preferences that are linear in $m$ and $\sigma^2$.

**Example 3.11 .** Consider an individual with CARA preferences with a coefficient of absolute risk aversion given by $r$, and income $I$. Suppose that the individual is able to buy a number of stocks in a company whose future return for each stock, $X$, is normally distributed with mean and variance: $(m, \sigma^2)$. Let the price of each stock be $q$. If the individual buys $n$ stocks then her future income is $I - nq + nX$. To compute the individual's demand as a function of the price of the stock, $n(q)$, one can use her certainty equivalent:

$$CE(I - nq + nX) = I - nq + nm - rn^2\sigma^2/2. \tag{3.34}$$

One concludes that the demand for the stock is:

$$n(q) = \begin{cases} 0, & \text{if } q \geq m, \\ \frac{m-q}{r\sigma^2} & \text{if } q \leq m. \end{cases} \tag{3.35}$$

Notice that demand decreases with the price of the stock and its riskiness. Also, the more risk averse the individual, the less she buys of the stock. When $m > q$, then as risk aversion decreases, $r \to 0$, demand becomes unbounded ($n(q) \to \infty$).

## 3.4 Uncertainty and Beliefs

Expected utility theory explicitly assumes that probabilities of the lotteries represent the objective likelihood of receiving a particular consequence. For example a lottery that pays \$0 and \$100 with equal probability has the explicit interpretation that if one were to accept this lottery repeatedly then one would expect to receive \$100 about half the time. For a risk averse individual, the certainly equivalence of this lottery played only once is likely to be much less than \$50.

On the other hand, insurance companies rely upon the law of large numbers to offer insurance policies that can be priced close to the $EMV$ of a risk. This follows from the law of large numbers. More precisely let $X \in \Delta(C)$ be a lottery, and suppose the lottery is run each period $t$, and the outcome is the random variable $X_t$, then by the law of large numbers one has almost surely:[7]

$$\lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} X_t = EMV(X). \tag{3.36}$$

Examples of such risks might include life expectancy, the probability of a traffic accident in Los Angeles on a particular day, the probability of conceiving a boy or girl, etc. In these cases the notion of probability is well-defined, and can be interpreted as the frequency of occurrence if the risky event is repeated.

Frank Knight (1921) observed that many, if not most, important economic decisions are *idiosyncratic*. For example, if one is trying to decide

7 A sequence of random variables $X_n$ converges to $x$ almost surely if $Pr\{lim_{n\to\infty} X_n = x\} = 1$.

whether or not to open a new restaurant, given that tastes are constantly changing, then it is simply not possible to use previous success rates to determine the likelihood of success for your particular business. Knight uses the term "uncertainty", as opposed to "risk" to describe business decisions such as these. Given that it is impossible to repeat this business experiment, then one cannot interpret the "probability of success" as the number of times such a restaurant would be successful.

Yet, in common day language, it is normal to speak of the likelihood of success. For example, the local bank may believe that the restaurant fills a need in the neighborhood, and hence it has a "good chance" of success. Based upon these beliefs, the bank may decide to lend money to the prospective owner of the business. The difficulty is that these probability assessments cannot be based upon replicable experiments.

This problem is a feature of many, if not most, day to day economic decisions. These include mutual fund investment choices, deciding to go to the beach based upon one's estimate of the weather, or making a career choice. All involve uncertain events whose true probabilities are probably unknowable, yet the individual must nevertheless make a choice. What does one mean by rational choice in these circumstances?

Savage (1954) provides a brilliant solution to the problem that is now widely accepted in economics. His insight was recognizing that even though individuals face uncertainty, the fact that they must make *some* choice can be used to derive a *subjective* assessment of a probability. For example, suppose that an individual has to choose between buying a lottery ticket from one of two different organizations A and B, each of which offers a $50 prize. If the individual chooses the ticket from organization A, then one may conclude that the individual believes that the probability of a win with ticket A is higher than with ticket B.

Savage shows that under the appropriate conditions, one can suppose that individuals act *as if* they are expected utility maximizers, where the probability used to compute the expected utility is derived from revealed preferences, rather than from a series of experiments. Once an individual has constructed her subjective probability assessments, she can then incorporate new information using the standard rules of probability and Bayes formula. A shortcoming of the theory is that it does not explicitly allow for risky events - events for which there is some agreement regarding the probability of the event (for example fair dice or the flip of a coin).[8] The existence of such events allows one to *calibrate* one's beliefs with an objective probability. It turns out that allowing a calibrating device not only allows one to make individual beliefs comparable, but also dramatically simplifies the development of the theory, as shown by Anscombe and Aumann (1963). In this chapter we begin with Savage's notion of a small world, and then use Anscombe and Aumann's method to construct an individual's subjective probability distribution.

### 3.4.1 The Small World Model

The first step in making a choice is to identify the possible outcomes. Savage (1954) observes that making a decision is akin to mentally constructing a small world in which one explores the results of different decisions. In constructing a small world model one acts as if the model captures all the relevant features of the decision at hand. One begins with the set of possible *states* $\Omega$, where each state $\omega \in \Omega$has the interpretation that it provides a complete description of the world (for simplicity suppose that there are $N$ states, though everything can be generalized for infinite sets). In principle the state describes all aspects of the world, though in practice one models only those aspects that are relevant to the current decision. Here are some examples.

**Example 3.12 .** One needs to decide whether or not to bring an umbrella to work, in which case $\Omega = \{\text{R, NR}\}$, where R = rain and NR = no rain.

Consider writing a contract for painting the inside and outside of a house that deals with contingencies for rainy days. All the work can be done in a week if there are at least two clear days to complete the outside work, and further suppose that one is certain to have two clear days during a ten day period. In that case the state describes the days that it rains, and is a sequence such as $\{R, NR, R, .., R\}$ taken from the set $\Omega = \{R, NR\}^{10}$.

Notice that in this example the state describes the outcome for each day. We would like to model behavior as a function of the state, yet one must capture the restriction that it is impossible for individuals to

8 Savage was well aware of the difficulties - see his discussion in chapter 4 on shortcomings of the personalistic view.

condition their behavior upon *future* rain patterns. The *event* that it rains on day one is defined as *all* the states that have rain on the first day:

$$\text{Rain Day One} = E_{R1} = \{\omega = \{\omega_1, \omega_2, ..., \omega_{10}\} \in \Omega | \omega_1 = \ \{\text{rain}\}\} \quad (3.37)$$

The event that it does not rain on day one, denoted $E_{NR1}$ is defined in a similar fashion. Notice that $E_{NR1} \cup E_{R1} = \Omega$, that is the events rain or no rain *partition* the state space. As time proceeds, and we learn more about what has occurred, this corresponds to a finer and finer partition of the state space. Once the weather has been observed for each of the 10 days, then the exact state of the world is known.

In this model, the everyday notion of an event, whether it rains or not, whether the price of a stock goes up or down, of whether there is an earthquake, etc, corresponds formally to subsets of the state space. Learning corresponds to narrowing down the true state of the world to smaller and smaller sets of possible states. The formalism provides a way to model both information and learning as events evolving over time. The information available in any period is represented by a set of events called a *partition:*

**Definition 3.13 .** Let $\mathcal{A}$ be a collection of subsets of $\Omega$. Then $\mathcal{A} = \{E_1, E_2, ..., E_n\}$ is a *partition* if

1. $\bigcup_{i=1}^{n} E_i = \Omega$.
2. If $E', E'' \in \mathcal{A}$, and $E' \neq E''$then $E' \cap E'' = \emptyset$.

In our example the set $\{E_R, E_{NR}\}$ is a partition, where $E_R$ denotes rain on day 1 and $E_{NR}$ denotes no rain. If one learns whether or not it rains on day 2, then this results in a refinement of the partition into four sets: $\{E_{R,NR}, E_{R,R}, E_{NR,R}, E_{NR,NR}\}$. At the end of ten days the partition will be the state space itself. Partitions measure how fine or accurate one's information is regarding the true state of the world.

Another, possibly more natural, example is the measurement of some variable, such as one's height. If a person is six feet tall, then depending upon the precision of the measuring and time of day, this really means that the person's height is in some range, say between 5' 11.5" and 6' 0.5". A person's height is really providing information upon the *set* of possible heights. A finer partition corresponds to a more accurate measurement of an individual's height, and corresponds to having more information. More formally, a partition $\Pi''$ is a *refinement* of $\Pi'$ if for each $E'' \in \Pi''$ and $E' \in \Pi'$ either $E'' \subset E'$ or $E'' \cap E' = \emptyset$.

Within Savage's model, a state conceptually represents a complete description of the world In practice however, it represents a partition corresponding to a description of those events relevant to the question at hand. It is worth highlighting this point because the conclusions of this model depend upon the way the world is modeled, and hence any errors in model formation can result in decision making errors. In particular, if contracting parties are using *different* representations of the world, involving possibly different partitions, then there is always some chance that contract terms and conditions will be interpreted differently.

For example, the homeowner might think in terms of rain or no rain, while the contractor might distinguish between rain, mist, drizzle and so on. Suppose that the original contract states that work does not have to proceed if it rains. The contractor might feel that a drizzle is rain, while the homeowner might believe that this corresponds to no rain. Thus, this apparently simple and clear contract can lead to a dispute because parties are using different models of the world.

We return to this question in more detail when we come to complex exchange. We raise the issue now to highlight the fact that this model is very rich and can be used to think about a wide variety of practical issues. For the remainder of this chapter and the next few chapters we proceed under the hypothesis that all individuals use the same model of the world - namely they agree upon the possible states $\omega \in \Omega$, though they might not have the same information.

The next step in the construction is to associate *consequences* to different realized states. For example, we care about whether it rains or not because we might become wet if we do not have an umbrella. Being wet is a consequence. Bringing an umbrella to work to ensure that we do not become wet is a choice that in Savage's model is called an *Act*.

As in the theory of decision under risk, suppose that there is a finite number of consequences $c$ taken from the set $C$. In addition, it is also assumed that there are risky events, such as real lotteries, for which individuals know and agree upon the true probability distributions. This is the point at which the Aumann and Anscombe theory deviates from Savage's. It is assumed that individuals can, if they wish, implement an objective randomization system that allows them to experience consequences in $C$ with specified probabilities. Hence, we may assume that the set of consequences is the set of lotteries used in the theory of decision making under risk, $\Delta(C)$.

We now model the choice of a person in the morning deciding whether or not to bring an umbrella to work. The consequences are being wet or dry when coming home. Suppose the person considers three possible choices: bring umbrella (U), flip a coin to randomly decide whether or not to bring the umbrella (RU) or don't bring it (NU). Using our notion of an event, we can represent these choices as acts as illustrated in table 3.2:

| **Act** | **Event** | |
|---|---|---|
| | $E_R$ | $E_{NR}$ |
| U | Dry | Dry |
| RU | 50% chance of Dry | Dry |
| NU | Wet | Dry |

**Table 3.2**
Consequences of Umbrella Choices

In this example, each choice corresponds to an act that is given by a mapping from events (states) to consequences. For example, the act RU can be written as a function:

$$f_{RU} : \Omega \to \begin{cases} \{.5, Dry, .5, Wet\}, & if\, \omega \in E_R \\ Dry, & if\, \omega \in E_{NR} \end{cases}.$$

More generally, any choice made by the individual can be viewed as an *act,* which formally is simply a function from the set of states of the world to consequences:

$$f : \Omega \to \Delta(C). \tag{3.38}$$

Let the set of acts be denoted by $\mathcal{F}$. The key point is that the definition of an act does not require the specification of the likelihood of states in $\Omega$. Each act provides a complete description of how different states of the world affect the consequences to be experienced by the individuals. The primitives of the model will be how the individual chooses among various acts. From these one will be able to *derive* the *subjective probability* that is assigned to each event.

Consider now an example from finance. Suppose that an entrepreneur wishes to open a new restaurant in a new location and needs to obtain financing. Suppose that all the relevant decisions, such as the choice of decor, chef, etc. have been made. Given these decisions, one still does not know if the project will be successful. This is modeled by supposing that the restaurant will either be very successful or simply profitable. Let the total income associated with these two outcomes be given by the values: $v_H > v_L > 0$.

Though the entrepreneur has completed a business plan, she still needs to raise $I$ dollars to finance the project, which can be done either with debt financing or in a partnership with a friend. With debt financing she must repay to the bank $I + R$ dollars, while in the case of the partnership she agrees to split revenues equally. The payoff that the entrepreneur receives in each state is given in table 3.3.

| **Act** | **State** | |
|---|---|---|
| | $v_H$ | $v_L$ |
| Debt Financing | $v_H - I - R$ | $v_L - I - R$ |
| Partnership | $v_H/2$ | $v_L/2$ |

**Table 3.3**
Entrepreneur's Payoffs with Debt versus Partnership Financing

Suppose that $v_H - I - R > v_H/2$ and $v_L/2 > v_L - I - R$, then if $v_H$ were to occur debt financing is preferred, while a partnership is preferred in the state $v_L$. This is a classic decision in the face of uncertainty. It is very difficult to assign an objective probability to the states $v_H$ and $v_L$. This problem has both a normative and positive interpretation in Savage's model. The normative approach is intended to help the entrepreneur make a better decision. In this case, after building the small world model, the next step is for the entrepreneur to assign probabilities to $v_H$ and $v_L$, and then choose the method of financing that yields the highest payoffs.

The positive approach is intended to capture the observed behavior of the entrepreneur. In that case we *infer* the probabilities assigned from the entrepreneur's choices. If we let $p$ be the probability of $v_H$ occurring, and we observe the choice of debt financing then we conclude:

$$p\,(v_H - I - R) + (1-p)\,(v_L - I - R) \geq p\,(v_H/2) + (1-p)\,(v_L/2)\,,$$

or that

$$p \geq \frac{2(I+R)-v_L}{(v_H-v_L)}.$$

In other words, by observing how the entrepreneur chooses we can infer the range of probabilities that she assigns to uncertain events. In the next section we make this idea more precise.

### 3.4.2 Preferences in the Face of Uncertainty

Let us now outline Anscombe and Aumann (1963)'s theory of decision making under uncertainty. It is assumed that the individual has preferences over lotteries, $\Delta(C)$, satisfying the axioms of expected utility theory, given by a preference relation $\succsim$. This allows one to construct a person's Bernoulli utility function over consequences in $C$. The insight of Anscombe and Aumann is to use these preferences to *calibrate* subjective probability assessments.

If one can construct a lottery over consequences, then one can also define a lottery over acts. An example is the decision to randomize between bringing an umbrella to work or not. Let $\Delta(\mathcal{F})$ denote the set of lotteries over acts. Anscombe and Aumann call these *horse lotteries* to evoke the image of an individual randomly choosing which of several horses to bet on in a horse race. The choice of a horse can be viewed as an act since one is never sure of the objective probability that a horse will win. Preferences over the horse lotteries, $\Delta(\mathcal{F})$, are denoted by $\succsim^*$. It is assumed that $\succsim^*$ also satisfies the axioms of expected utility theory, where each act $f \in \mathcal{F}$ is viewed as a consequence. Thus preferences on horse lotteries have the expected utility form, where each act $f \in \mathcal{F}$ is assigned a Bernoulli utility $u(f)$.

Thus, the axioms of rational choice theory alone ensure one has utility functions over lotteries and horse lotteries. The final step in the Anscombe and Aumann (1963) construction is to find a way to formally link the two sets of preferences. The first axiom requires that $\succsim^*$ rank constant acts in the same way as $\succsim$.

**Definition 3.14 .** Preferences satisfy the property of *constant acts* if for every $l, l' \in \Delta(C)$, $l \succsim l'$ if and only if $f \succsim^* f'$, where $f(\omega) = l$ and $f'(\omega) = l'$ for all $\omega \in \Omega$.

As a minor abuse of notation let the act $l$ refer to an act with the outcome $l$ in each state $\omega \in \Omega$. The next axiom requires that the agent always prefers to have the outcome for a single state improved upon:

**Definition 3.15 .** The preference relation $\succsim^*$is *monotonic* if for every act $f \in F$, lottery $l' \in \Delta(C)$ such that for some state $\omega' \in \Omega$, $l' \succsim f(\omega')$ then $g \succsim^* f$ , where $g(\omega') = l'$, and for $\omega \neq \omega'$, $g(\omega) = f(\omega)$.

Thus, an individual always prefers (under $\succsim^*$) an act that has an outcome in a particular state improved upon (under $\succsim$). Notice that this ranking cannot be strict because it may be the case that state $\omega$ is believed to be impossible, in which case the payoff in state $\omega$ is irrelevant. As Anscombe and Aumann observe, this axiom corresponds to Savage's *sure thing principle* (see section 2.7 of Savage (1954)).

The next axiom requires that the timing of uncertainty and risk does not affect preferences. Consider the horse lottery $\{\alpha_1, f_1; \alpha_2, f_2; ...; \alpha_n, f_n\} \in \Delta(F)$ with the interpretation that the act $f_i$ is chosen with probability $\alpha_i$. This lottery has two possible interpretations. The first is that act $i$ is chosen with probability $\alpha_i$, the state $\omega$ is realized, resulting in the consequence $f_i(\omega)$. The second is that the state $\omega$ is realized first, resulting in the *lottery* $\{\alpha_1, f_1(\omega); \alpha_2, f_2(\omega); ...; \alpha_n, f_n(\omega)\} \in \Delta(C)$, at which point consequence $f_i(\omega)$ is chosen with probability $\alpha_i$. The next axiom requires individuals to be indifferent to the order in which uncertainty is realized.

**Definition 3.16 .** Preferences satisfy the reversal of order in compound lotteries property if for any set of acts, $\{f_i \in F | i = 1, ..n\}$, and any horse lottery $f' = \{\alpha_1, f_1; \alpha_2, f_2; ...; \alpha_n, f_n\} \in \Delta(F)$, then $f' \sim^* f''$, where $f''(\omega) = \{\alpha_1, f_1(\omega); \alpha_2, f_2(\omega); ...; \alpha_n, f_n(\omega)\}$, $\omega \in \Omega$.

These axioms allow us to bootstrap from expected utility over lotteries to utility over acts. Suppose that there is a least and most preferred consequence, $c_{\min}$ and $c_{\max}$ such that $c_{\max} \succ c_{\min}$ (and hence $c_{\max} \succ^* c_{\min}$). Then without loss of generality we may choose the Bernoulli utility representing $\succsim$ such that $u(c_{\max}) = 1$ and $u(c_{\min}) = 0$, and let $U(l) = \sum_{c \in C} p_c u(c)$ be the corresponding expected utility. Let $u^*(\cdot)$ denote the Bernoulli utility function representing $\succsim^*$also be normalized so that $u^*(c_{\max}) = 1$ and $u^*(c_{\min}) = 0$, where $c_i$ represents the constant act $f(s) = c_i$ for all $s \in \Omega$. We may now state the main result of this section.

**Theorem 3.17 .** *(Anscombe and Aumann) Given that* $\succsim$ *and* $\succsim^*$ *over lotteries and horse lotteries respectively satisfy the von Neumann-Morgenstern axioms for expected utility, and that* $\succsim^*$ *also satisfies constant acts, monotonicity and reversal of order in compound lotteries, then there is a unique set of non-negative numbers* $p_s$ *summing to* 1 *such that for all acts* $f \in F$ *the Bernoulli utility for* $\succsim^*$ *satisfies:*

$$u^*(f) = \sum_{\omega\in\Omega} p_s U(f(\omega)) \tag{3.39}$$

$$= \sum_{\omega\in\Omega}\sum_{c\in C} p_s \alpha(f(\omega), c)\, u(c), \tag{3.40}$$

*where* $U(\cdot)$ *is the utility function representing* $\succsim$, $u(\cdot)$ *the corresponding Bernoulli utility, and* $\alpha(f(s), c)$ *is the probability that* $c$ *is chosen for the lottery* $f(\omega) \in \Delta(C)$.

*Proof.* Let $f$ be an act and set $l_{\omega_i} = f(\omega_i)$. Given that for any pair of lotteries, $l \succsim l'$ if and only if $U(l) \geq U(l')$, then from the axiom of constant acts and monotonicity we can replace $l$ by $U(l)$ and with a slight abuse of notation represent $\succsim^*$ by $u^*(U(l_{\omega_1}), U(l_{\omega_2}), ..., U(l_{\omega_N}))$. In particular $u^*(1, 1, ..., 1) = u^*(c_{\max}) = 1$ and $u^*(0, 0, ..., 0) = u^*(c_{\min}) = 0$. For each $\omega_i \in \Omega$, let $f_{\omega_i}$ denote the act such that $f_{\omega_i}(\omega) = c_{\max}$ if $\omega = \omega_i$ and $c_{\min}$ otherwise, then define $p_{\omega_i} = u^*(f_{\omega_i})$.

Next we make the following observation.

**Lemma 3.18 .** *Take any* $(r_{\omega_1}, r_{\omega_2}, ..., r_{\omega_N})$ *such that* $r_{\omega_i} \in [0, 1]$ *for all* $\omega_i \in \Omega$. *If for some* $k > 0$, *we have* $kr_{\omega_i} \in [0, 1]$*for all* $\omega_i \in \Omega$, *then* $u^*(kr_{\omega_1}, ..., kr_{\omega_N}) = ku^*(r_{\omega_1}, ..., r_{\omega_N})$.

To show this first suppose that $k \leq 1$ and notice by continuity there are lotteries $l_{\omega_i}$, such that $u(l_{\omega_i}) = r_{\omega_i}$. Then $(kr_{\omega_1}, ..., kr_{\omega_N})$ represents the act, $f'$, where for each $\omega_i$ the outcome is the compound lottery $\{k, l_{\omega_i}; (1-k), c_{\min}\}$. Let $f''(\omega_i) = l_{\omega_i}$ for all $\omega_i$,then from the reduction of compound lotteries we have:

$$f' \sim^* \{k, f''; (1-k), c_{\min}\}, \tag{3.41}$$

from which using the expected utility properties of $\succsim^*$ we can conclude $u^*(kr_{\omega_1}, ..., kr_{\omega_N}) = ku^*(r_{\omega_1}, ..., r_{\omega_N}) + (1-k)\, u^*(c_{\min}) = ku^*(r_{\omega_1}, ..., r_{\omega_N})$.

Next, if $k > 1$,then:

$$u^* \left(r_{\omega_1}, ..., r_{\omega_N}\right) = u^* \left(kr_{\omega_1}/k, ..., kr_{\omega_N}/k\right) = (1/k)\, u^* \left(kr_{\omega_1}, ..., kr_{\omega_N}\right), \tag{3.42}$$

and multiplying through by $k$ completes the observation.

Now we fix $\left(r_{\omega_1}, r_{\omega_2}, ..., r_{\omega_N}\right)$ such that $u\left(l_{\omega_i}\right) = r_{\omega_i}$, where $l_{\omega_i} = f\left(\omega_i\right)$.

Let $c = r_{\omega_1} + ... + r_{\omega_N}$. If $c = 0$, then $r_{\omega_i} = 0$, and we are done. If $c > 0$, then $r_{\omega_i}/c$ are non-negative and sum to 1. Let $g$ denote an act such that $u\left(g\left(\omega_i\right)\right) = r_{\omega_i}/c$. But then this implies $g\left(\omega_i\right) \sim \left\{r_{\omega_i}/c, c_{\max}; 1 - r_{\omega_i}/c, c_{\min}\right\} = l_i$, and hence the decision maker is indifferent between $g$ and the act that in state $\omega_i$ plays the lottery $l_i$. Then from the reversal of order of compound lotteries property it follows:

$$g \sim^* \left\{r_{\omega_1}/c, f_{\omega_1}; ...; r_{\omega_N}/c, f_{\omega_N}\right\}. \tag{3.43}$$

Hence by the linearity of the Bernoulli utility and the expected utility property:

$$u^*\left(f\right) = u^*\left(r_{\omega_1}, ..., r_{\omega_N}\right) \tag{3.44}$$

$$= cu^*\left(g\right) \tag{3.45}$$

$$= c\left(\sum_{\omega_i \in \Omega} \frac{r_{\omega_i}}{c} u^*\left(f_{\omega_i}\right)\right) \tag{3.46}$$

$$= \left(\sum_{\omega_i \in \Omega} r_{\omega_i} p_{\omega_i}\right) \tag{3.47}$$

$$= \left(\sum_{\omega_i \in \Omega} U\left(f\left(\omega_i\right)\right) p_{\omega_i}\right). \tag{3.48}$$

The final equality follows from the fact that $U\left(\cdot\right)$ is a von Neumann-Morgenstern utility function. □

Notice that the numbers $p_\omega$ are *derived* from the preference relation $\succsim^*$, and represent the *subjective probability* that the individual attaches to state $\omega$. When the conditions of this theorem hold, preferences are said to satisfy the conditions for *subjective expected utility theory (SEU).* Preferences satisfying *SEU* can be used to derive the subjective probability, $p_E$, of any event $E \subset \Omega$ by calibrating with a lottery. Begin with the act:

$$f_A(\omega) = \begin{cases} c_{\max}, & \text{if } \omega \in E \\ c_{\min}, & \text{if } \omega \in E^c \end{cases} \tag{3.49}$$

Given that preferences are continuous, then there is an $\alpha$ and lottery $\{\alpha, c_{\max}; (1-\alpha), c_{\min}\}$ such that $f_A \sim^* \{\alpha, c_{\max}; (1-\alpha), c_{\min}\}$ from which we conclude:

$$p_E u(c_{\max}) + (1 - p_E) u(c_{\min}) = \alpha u(c_{\max}) + (1-\alpha) u(c_{\min}). \tag{3.50}$$

Thus $p_E = \alpha$ and since $\alpha$ is the objective probability of receiving $c_{\max}$, then $p_E = \alpha$ is the subjective probability of event $E$. In particular, subjective probabilities can be manipulated using the standard rules of probability theory. This is illustrated in the next section on the value of information.

### 3.4.3 The Value of Information

In many situations an individual may delay decision making in order to gather more information. This is particularly important for contract formation where individuals may not wish to enter an agreement until more information is received, or, as in the case of principal-agent models, the contract terms themselves are a function of information received. This section illustrates how to model information and to measure its value.

Consider the financing example above, and observe that there may be a number of distinct events or states that determine whether the restaurant has a high or low value. For example, one may be sure of a high profit if another restaurant opens on the same street attracting more potential clients to the area. Alternatively, the street on which the restaurant is located may be turned into a pedestrian only street, again ensuring high profits. These are two different possible events representing potential pieces of information.

For simplicity suppose that there are two underlying states for each event corresponding to high or low profits: the states $\omega_{1H}$ and $\omega_{2H}$ result in high profits, while the states $\omega_{1L}$ and $\omega_{2L}$ result in low profits. In that case the acts available to the decision maker are given in the following table.

| **Act** | **State** | | | |
|---|---|---|---|---|
| | $\omega_{1L}$ | $\omega_{2L}$ | $\omega_{1H}$ | $\omega_{2H}$ |
| Debt Financing - $f_{DF}$ | $v_L - I - R$ | | $v_H - I - R$ | |
| Partnership - $f_P$ | $v_L/2$ | | $v_H/2$ | |

**Table 3.4**
Entrepreneur's Payoffs with Debt versus Partnership Financing with Four Events

Our decision maker is assumed to satisfy the axioms of decision making under uncertainty and is risk neutral, with Bernoulli utility $u(x) = x$. Therefore the payoffs from the acts $f_{DF}$ and $f_P$ are:

$$U(f_{DF}) = E(u(f_{DF}(\cdot))) \tag{3.51}$$
$$= (p_{1L} + p_{2L})\, v_L + (p_{1H} + p_{2H})(v_H - I - R), \tag{3.52}$$
$$U(f_P) = E(u(f_P(\cdot))) \tag{3.53}$$
$$= (p_{1L} + p_{2L})\, v_L/2 + (p_{1H} + p_{2H})\, v_H/2, \tag{3.54}$$

where $p_{nk}$ is the subjective probability for state $\omega_{nk}$. For purposes of discussion suppose that $U(f_P) > U(f_{DF})$ (and maintain the assumption made in the previous section that in the good state debt financing is preferred and vice versa in the bad state).

The reason for expressing the high and low outcomes as the result of more primitive states is that it provides a way to formally model what we mean by information. For example one may be told whether the true state is in $E_1 = \{s_{1L}, s_{1H}\}$ or $E_2 = \{s_{2L}, s_{2H}\}$. Given that an event $E$ has occurred, then the conditional probability of an event $A$ is:

$$P(A|E) = P(A \cap E)/P(E). \tag{3.55}$$

Hence the probability of high (respectively low) profits conditional upon $E_i$ is $p_{iH}/(p_{iH} + p_{iL})$ (respectively $p_{iL}/(p_{iH} + p_{iL})$).

In this case the information is represented by a partition of the state space, $\Pi = \{E_1, E_2\}$. The conditional expectation of an act $f$ given by the information set $\Pi$ is denoted by $E(u(f(\cdot))|\Pi)$. Formally this represents a *function* from $\Omega$ to $\Re$ with the property that it is sensitive to the information contained in $\Pi$:

$$E(u(f)|\Pi)(s) = \begin{cases} U(f|E_1) \text{ if } s \in E_1, \\ U(f|E_2) \text{ if } s \in E_2, \end{cases} \tag{3.56}$$

which in the case of the present example these payoffs are defined by:

$$U(f_{DF}|E_i) = (p_{iL}v_L + p_{iH}v_H)/(p_{iH} + p_{iL}) - I - R, \tag{3.57}$$
$$U(f_P|E_i) = (p_{iL}v_L + p_{iH}v_H)/2(p_{iH} + p_{iL}), \tag{3.58}$$

Formally, the function $E(u(f(\cdot))|\Pi)$ is *measurable* with respect to the information set $\Pi$. This simply means that my decision can vary as a function of event $E_1$ or $E_2$, but for any two states $\omega_a, \omega_b \in E_i$ then $E(u(f(\cdot))|\Pi)(\omega_a) = E(u(f(\cdot))|\Pi)(\omega_b)$. Essentially, one cannot vary one's actions with information one does not have.[9] More precisely:

**Definition 3.19 .** Given a finite set $\Omega$ and a partition $\Pi$ of $\Omega$, then a function $f : \Omega \to Y$, where $Y$ is any subset of $\Re^n$ is measurable with respect to $\Pi$ if for every $E \in \Pi$ we have:

$f(\omega_a) = f(\omega_b), \forall \omega_a, \omega_b \in E.$

When applied to a choice it requires that one's decision cannot vary with information one does not have. If $U(f_{DF}|E_1) > U(f_P|E_1)$ and event $E_1$ is observed, then the entrepreneur would choose debt financing. Notice that the assumption $U(f_P) > U(f_{DF})$ implies that $U(f_{DF}|E_2) < U(f_P|E_2)$, and hence the entrepreneur chooses a partnership if event $E_2$ occurs. In the absence of any information, the entrepreneur chooses the partnership, but if the information partition $\{E_1, E_2\}$ is available then she would choose debt financing if $E_1$ occurs, and a partnership otherwise.

New information *refines* the set of states, and allows us to update the probabilities of events that might occur in the future. In this example, learning about event $E_1$ is valuable because it would cause us to change our financing plan to debt from a partnership. Let $U^*(\Pi)$ be the maximum utility when the partition $\Pi = \{E_1, E_2\}$ is available. Then if $P(E_i) > 0$, for $i = 1, 2$ we have:

$$U^*(\Pi) = U(f_{DF}|E_1)P(E_1) + U(f_P|E_2)P(E_2), \tag{3.59}$$
$$> U(f_P|E_1)P(E_1) + U(f_P|E_2)P(E_2), \tag{3.60}$$
$$= U(f_p). \tag{3.61}$$

9 For a full discussion of measure theory see any standard book on real analysis, such as Ash (1972). The main point here is that the discussion of measurability in mathematics texts can often seem very abstract, but in fact corresponds rather naturally to the amount of information available.

More generally, we can define the value of information from having access to a partition $\Pi$. For any event $E_i \in \Pi$ payoffs conditional upon $E_i$ are defined by:

$$U(f|E_i) = \sum_{\omega\in\Omega} u(f(\omega)) P(\omega|E_i), \tag{3.62}$$

where $P(\omega|E_i)$is the conditional probability of $\omega$ occurring given $E_i$ and our subjective probability assessments. Suppose that the space of consequences are lotteries over money payoffs, and the Bernoulli utility is continuous with no lower bound, then we denote the *value of information* by $V(\Pi)$, and it is defined as the solution to the following equality:

$$\sum_{E\in\Pi} \max_{f\in F} U(f - V(\Pi)|E) P(E) = \max_{f\in F} U(f) \tag{3.63}$$

This is the maximum amount of money one is willing to pay to receive information $\Pi$. More generally, if $\Pi$ represents one's current information partition, then new information $\Pi'$ corresponds to a partition of $\Pi$ (see section 3.4.1). Information has value because it permits the decision maker to tailor her decisions to finer sets of states, a result that is summarized in the next proposition.

**Proposition 3.20 .** *Suppose that preferences over money satisfy the Anscombe and Aumann axioms of decision making under uncertainty, and that the Bernoulli utility is continuous with no lower bound. If* $\Pi''$ *is a refinement of* $\Pi'$ *then* $V(\Pi'') \geq V(\Pi')$. *Moreover,* $V(\Pi'') > V(\Pi')$ *if and only if there exists an* $E'' \in \Pi''$ *such that* $P(E'') > 0$, *and* $\max_{f\in F} U(f|E'') > U(f'|E'')$, *where*

$$f' \in \arg\max_{f\in F} U(f|E'), E'' \subset E' \in \Pi'.$$

*Proof.* The fact that the Bernoulli utility is continuous with no lower bound ensures the existence of the value of information. Given that $\Pi''$ is a refinement of $\Pi'$, then $\Pi''$ defines a refinement for each $E \in \Pi'$, denoted $\Pi_E$.

$$\begin{aligned}
\sum_{E' \in \Pi'} \max_{f \in F} U\left(f - V\left(\Pi'\right) | E'\right) P\left(E'\right) &= \max_{f \in F} U\left(f\right) \\
&= \sum_{E \in \Pi''} \max_{f \in F} U\left(f - V\left(\Pi''\right) | E\right) P\left(E\right) \\
&= \sum_{E' \in \Pi'} \sum_{E'' \in \Pi_E} \max_{f \in F} U\left(f - V\left(\Pi''\right) | E''\right) P\left(E''\right) \\
&\geq \sum_{E' \in \Pi'} \max_{f \in F} \sum_{E'' \in \Pi_E} U\left(f - V\left(\Pi''\right) | E''\right) P\left(E''\right) \\
&= \sum_{E' \in \Pi'} \max_{f \in F} U\left(f - V\left(\Pi''\right) | E'\right) P\left(E'\right)
\end{aligned}$$

Given the continuity of $U$, then it follows that $V\left(\Pi''\right) \geq V\left(\Pi'\right)$. The second result is a straight forward consequence of the fact that utility is a positive linear combination of $U\left(f|E\right)$, for $E$ in the relevant partition.

□

It is also assumed that once subjective probabilities have been determined, then the decision maker is able to apply the standard tools of probability theory. In particular it is assumed that beliefs are updated using Bayes' theorem. Notice that information is strictly valuable if and only if it allows us to make a strictly better decision for some event $E$ that occurs with positive probability.

The fact that information is valuable only for events that have a positive *prior* probability of occurring creates a number of serious problems not only for decision making, but also for contract formation and game theory. One of the reasons that contract disputes arise in court is because an event occurs that neither party to the contract anticipated, or, in other words, an event with subjective probability of zero has occurred. Unfortunately, when such an event occurs Bayesian theory cannot be used to update beliefs. For example suppose that $P\left(E\right) = 0$, and one wishes to contemplate the possibility that event $A$ occurs after observing event $E$, however:

$$P\left(A|E\right) = \frac{P\left(A \cap E\right)}{P\left(E\right)} = \frac{0}{0} = \text{undefined}. \tag{3.64}$$

Hence in cases that the subjective decision maker has made a mistake, and overlooked an event that actually transpires, the rules of probability theory can no longer be used to derive an individual's *posterior* beliefs. There is no satisfactory resolution of this problem in the literature, though several authors have explored the meaning of probability in these contexts. The interested reader is referred to Jeffrey (1965) and de Finetti (1974) for an extended philosophical discussion. Diaconis and Zabell (1982) reviews a number of alternatives to Bayes' rule for the updating of subjective probabilities.

The standard approach to this problem, advocated by Jeffrey (1965), is to use a diffuse prior and suppose that all events have strictly positive probability. This idea forms the basis for the concept of a perfect equilibrium in game theory, which is discussed in further detail in the next chapter (see also Myerson (1991)'s discussion of conditional probability systems, page 21). However, incomplete contracts often arise because the contracting parties are simply unaware that an event might occur, hence planning for such a state cannot even occur. Savage was well aware of this problem, which is precisely why he used the term "small world". For the remainder of this part and for Part III it is assumed that small world problems do not arise (or more formally we assume the conditions outlined in section 5.5 of Savage (1954) hold). When we come to complex exchange in Part IV we shall return to this issue.

## 3.5 Summary

Before attributing all failures of the theory to our (false) model of rational choice one should keep in mind Charlie Munger's first cause of human misjudgment (from a speech at Harvard Law School, 1995):

"*1. Under-recognition of the power of what psychologists call 'reinforcement' and economists call 'incentives'* - Well I think I've been in the top 5% of my age cohort all my life in understanding the power of incentives, and all my life I've underestimated it. And never a year passes but I get some surprise that pushes my limit a little farther."

When used as a positive model, the goal of rational choice theory is to provide a representation of behavior that is rich enough to capture the main qualitative features of human behavior. It is not a perfect representation of human behavior, but over the years it has proven to be the most successful model of human behavior for studying social phenomena. The recent advances in the psychology of economics have been due to studying deviations from rational choice, rather than attempting to replace rational choice theory with a completely new model.

Even so, the model has a great deal of indeterminacy. The fact that individuals must make choices in the face of uncertainty implies that the individuals must choose even when their knowledge regarding the consequences of their choices is limited. Savage's theory of rational choice in that setting implies that individuals building models of the world necessarily entails the use of subjective evaluations of the likelihood of future events. Given that the power of incentives is intimately linked to how individuals connect current actions to future rewards, this implies that the theory of incentive contracts must address the formation of these subjective beliefs. An area of research that devotes a great deal of attention to this problem is game theory, the topic of Chapter 4 The goal is to understand decision making in environments with many interacting individuals whose behavior satisfies the axioms of rational choice theory.

## 3.6 Exercises

1. Let $n$ be the number of possible choices in set $A$. Suppose that an individual could face any subset of set $A$. How many different behaviors are possible as a function of $n$?
2. Suppose we allow indifference, and further suppose that the individual selects each outcome in her indifferent set with equal probability. How many observations within a choice set $C \subset A$, of size $m$ does one need to be able to identify $D = B(C)$ with probability greater than $1 - \epsilon$? Now suppose that one can do an experiment and observe how the individual will choose for different $C \subset A$. How many experiments does one need to identify the behavior $B(.)$ with probability $(1 - \epsilon)$?
3. Prove Theorem 3.1. Hint: first show that transitivity allows one to completely order the set $A$, such that $x_1 \succsim x_2 \succsim ... \succsim x_N$, $x_1, .., x_N$ are the elements of $A$. Then use this ordering to assign a utility to each element of $A$.
4. Show that the group decision by majority rule may not satisfy the conditions for rational choice.

5. Let $A = \Re^2$, and suppose that $\{x_1, x_2\} \succsim \{y_1, y_2\}$ if $x_1 > y_1$ or $x_1 = y_1$ and $x_2 \geq y_2$. These are called *lexicographical preferences.* Prove that they satisfy the axiom of rationality, but are not continuous.
6. Show that in the Lottery Simplex of Figure 1, indifference curves must be a family of parallel straight lines.
7. Explicitly derive the demand for insurance in example 3.9 when $u(x) = \log(x)$.
8. Given continuity and the independence axioms, prove the existence of a continuous utility function. Hint: begin by showing that we can set $U(l) = \alpha$ for all lotteries of the form $l = \alpha\bar{l} + (1-\alpha)\underline{l}$, where $\bar{l}$ and $\underline{l}$ are the most preferred and least preferred lotteries respectively, and $\bar{l} \succ \underline{l}$. These exist due to the continuity of preferences. Then show how continuity can be used to extend this utility function to arbitrary lotteries. Observe that the problem is trivial when $\bar{l}$ is indifferent to $\underline{l}$.
9. Suppose that an individual has preferences represented by expected utility with Bernoulli utility. Show that these preferences can be represented by another Bernoulli utility function $u'(\cdot)$ if and only if there is a $\beta > 0$ and $\alpha$ such that:

$$u'(c) = \beta u(c) + \alpha, \text{ for all } c \in C. \tag{3.65}$$

10. Fill in the details for the proof of proposition 3.6.
11. A risk-averse individual with initial wealth $w$ can invest his wealth into two assets. One is risk-free, with a gross return of $R$; the other is risky, its gross return is some normal random variable $\tilde{\theta}$ with mean $t$ and variance $\sigma^2$. Let $\alpha \in [0,1]$ denote the fraction of initial wealth invested into the risky asset.

    a. Show that if $t \leq R, \alpha = 0$
    b. Show that if $t \geq R, \alpha > 0$
    c. The utility function of the individual is CARA. How does $\alpha$ vary with the coefficient of absolute risk aversion?
    d. Assume that an agent has constant *relative* risk aversion. How does he modify his portfolio when he gets richer? In particular, how does the share of his wealth invested into the risky asset change?

e. Same question for decreasing relative risk aversion.

12. In this exercise we explore a way to model boundedly rational decision making. Suppose that you are in a company that is choosing to bring to market a software package to evaluate the value of securities. Let $S$ be the set of all possible software packages. Suppose that at a cost $c$ you can sample a package $s \in S$, and determine its value $v_s$. Your profits are simply the value of the package you choose, less the costs spent in choosing a package. How would a rational entrepreneur model her decision problem, and what are the characteristics of the optimal solution? (see MacLeod (2002)).

## References


Anscombe, F. J. and R. J. Aumann (1963). A definition of subjective probability. *Annals of Mathematical Statistics 34*, 199–205.

Arrow, K. J. (1958, January). Utilities, attitudes, choices: A review article. *Econometrica 26*(1), 1–23.

Ash, R. B. (1972). *Real Analysis and Probability.* San Diego: Academic Press.

Becker, G. (1976). *The economic approach to human behavior*. Chicago, Il.: University of Chicago Press.

Bernoulli, D. (1738). Specimen theoriae de mensura sortis. *Commentarii Academiae Scientarum Imperialis Petropolitanae 5*, 175–92.

de Finetti, B. (1974). *Theory of Probability, Vols. I and II*. New York, NY: John Wiley & Sons.

Debreu, G. (1954). Representation of a preference ordering by a numerical function. In R. M. Thrall, C. H. Coombs, and R. L. Davis (Eds.), *Decision Processes*, pp. 159–65. New York, NY: Wiley.

Diaconis, P. and S. L. Zabell (1982). Updating subjective probability. *Journal of American Statistical Association 77*, 822–30.

Epstein, L. G. (1992). Behavior under risk: Recent developements in theory and applications. In J.-J. Laffont (Ed.), *Advances in Economic Theory: Sixth World Congress*, Volume I. Cambridge, UK: Cambridge University Press.

Ferraro, F., J. Pfeffer, and R. I. Sutton (2005). Economics language and assumptions: How theories can become self-fulfilling. *The Academy of Management Review 30*(1), pp.8–24.

Frank, R. H., T. Gilovich, and D. T. Regan (1993). Does studying economics inhibit cooperation? *The Journal of Economic Perspectives 7*(2), pp.159–171.

Harless, D. W. and C. F. Camerer (1994, November). The predictive utility of generalized expected utility theories. *Econometrica 62*(6), 1251–1289.

Hey, J. D. and C. Orme (1994, November). Investigating generalizations of expected utility theory using experimental data. *Econometrica 62*(6), 1291–1326.

Hobbes, T. (1651). *Leviathan*. London, England: Reprinted by Penguin Classics, 1985.

Jeffrey, R. C. (1965). *The Logic of Decision*. New York, NY: McGraw-Hill.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decisions under risk. *Econometrica 47*, 262–91.

Knight, F. H. (1921). *Risk, Uncertainty, and Profit*. New York, NY: Hart, Schaffner, and Marx.

Kuhn, T. S. (1970). *The Stucture of Scientific Revolutions, Second Edition*, Volume 2 of *International Encyclopedia of Unified Science*. Chicago, IL: The University of Chicago Press.

Machina, M. J. (1982). "expected utility" analysis without the independence axiom. *Econometrica 50*(2), pp.277–323.

MacLeod, W. B. (2002). Complexity, bounded rationality and heuristic search. *Contributions to Economic Analysis and Policy 1*(1).

Myerson, R. B. (1991). *Game Theory*. Cambridge, MA: Harvard University Press.

Popper, K. R. (1963). *Conjectures and Refutations: the Growth of Scientific Knowledge*. New York, NY.: Basic Books.

Samuelson, P. A. (1938). A note on the pure theory of consumer behavior. *Economica 5*(17), 61–71.

Savage, L. J. (1972 (first published 1954)). *The Foundations of Statistics*. New York, N.Y.: Dover Publications.

Von Neuman, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.