The following notes complement the textbook for STAT 271. They are intended to provide an overview of decision theory. These notes are not to be reproduced without written permission.
General decision theory concerns human decision making in the following idealized context: a decision problem is characterized by a set of actions available to and under the control of the decision maker, and a set of states of nature that are randomly selected (by a metaphorical nature) and are beyond the control of the decision maker. Each state-act pair determines an outcome, and since the states of nature are beyond the control of the decision maker, so too are the outcomes. The decision maker is endowed with information about the states of nature and has a preference ordering over the set of outcomes. The decision maker resolves the decision problem by inducing a preference ordering over the set of acts and then selecting the most preferred act.

General decision theory covers decision making in a very wide array of circumstances, including decision making under certainty, under uncertainty, under risk, and against a rational opponent. These are characterized as follows: the decision maker faces a problem:

under certainty if each act yields the same outcome for every state of nature;
under uncertainty if at least one act yields a set of outcomes of cardinality 2 or more, and the decision maker's information is not complete enough to formulate a probability distribution over the states of nature;
under risk if at least one act yields a set of outcomes of cardinality 2 or more, and the decision maker's information is complete enough to formulate a probability distribution over the states of nature; and
against a rational opponent if the states of nature are purposefully (and not randomly) selected.

Decision making under certainty is remarkably simple, is a special case of decision making under uncertainty and risk, and is covered in Principles of Economics courses. Decision making against a rational opponent is called game theory, constitutes a discipline unto itself, and is remarkably complex but very powerful. The distinction between decision making under uncertainty and under risk was begun by Frank Knight (1921), and rests heavily on the scope of interpretations of the probability calculus one is willing to accept. With the rise of the subjectivist interpretation,
mostly at the hands of Savage (1954), Knight’s distinction has dissolved. Thus, in what follows I will treat the terms as synonyms, and will refer simply to decision making under uncertainty.

The purpose of these notes is to provide the reader with a systematic introduction to the mathematical and structural aspects of the theory of human decision making. Decision theory provides both a normative and a descriptive account of human decision making. As a normative account it is of interest to economists and philosophers. As a descriptive account it is of interest to psychologists. Economists view decision theory as the natural extension of the certainty model to the more realistic world of uncertainty and incomplete information. Philosophers view decision theory as an account of human rationality, and thereby as both an object of study and an analytical tool. It is an object of study in the general inquiry to human rationality and human action taking. It is a tool in the realm of inference and inductive logic. Psychologists view decision as a testable account of human behavior.

The contributions of economists, philosophers, and psychologists are now so intertwined that many of the original disciplinary threads have been lost. For example, the empirical testing of the theory was once the province of psychologists, but now involves researchers in economics, business, and accounting. Similarly, the construction of the basic theory was once the province of economists and mathematicians, but now involves psychologists and philosophers. This set of notes is composed under the dual realizations that decision theory is deeply embedded in these three disciplines and that no single text covers decision theory for the three disciplines.

These notes cover only the details of the theory. They are not intended to be encyclopedic. The relevant literature is too vast to be summarized in one set of notes. Rather, these notes are a guide that introduces the reader to the mileposts in the three component evolution of the theory.

There are many basic text books in applied decisional analysis (Clemen 1992, Samson 1988, Luce and Raiffa 1967). Similarly, there are many books covering the psychological aspects of the theory (Hogarth 1980, Kahneman and Tversky 1979, 1992.) Interestingly, texts on the
I. BASIC DECISION THEORY

INTRODUCTION
Like all theories, decision theory is an elaborate metaphor. Decision theory provides a prescriptive analysis of the process of human cognitive decision making based upon an unfolding of decision problems. In brief, decision theory treats a decision problem as a system of forked paths and determines the optimal path, and thereby the optimal decision, by assigning a value to each path and selecting the path with the extreme value.

The theory consists of two components: an account of the structure of decision problems, and a computation process leading to the resolution (or solution) of decision problems. The former provides a mechanism for displaying all of the details of a decision problem; the latter provides a mechanism for resolving the problem.

ELEMENTS OF BASIC DECISION THEORY
Decision theory pertains to human decision making in a world of incomplete information and incomplete human control over events. The decision theory metaphor posits two players: a cognitive human and a randomizing nature. The human, called the decision maker, performs analyses, makes calculations, and cognitively decides upon a course of action in an effort to optimize his or her own welfare. The metaphorical Nature is non-cognitive, does not perform analyses or make calculations, and does not choose courses of action in any self-interested way. Rather, nature blithely selects courses of action purely in a random way.

The two fundamental concepts of decision theory are states of nature and acts. States of nature are under the control of nature and beyond the control of the decision maker, and are randomly selected by nature. Acts are under the control of the decision maker and any one of the available acts can be selected by the decision maker. Further, decision theory presumes that the problem is presented to the decision maker, i.e., that the problem itself, like the states of nature, is beyond the control of the decision maker. A decision problem is represented as a pair \(<S,A>\) composed of a set \(S\) of states of nature and a set \(A\) of acts.
The states and acts are descriptions of events. The level of detail of the
descriptions are such that the decision maker can infer the consequence,
or outcome, of every act in each state. The traditional characterization of
certainty, risk, and uncertainty is as follows:

- the decision maker faces a decision problem under certainty if and
  only if there is a unique outcome associated with each act;
- the decision maker faces a decision problem under risk if and only if
  there is a set of outcomes associated with each act, and the decision
  maker knows the probabilities of the outcomes in each set; and
- the decision maker faces a decision problem under uncertainty if and
  only if there is a set of outcomes associated with each act, and the
  decision maker does not know the probabilities of the outcomes in each
  set.

Note that the foregoing definitions can be characterized in terms of
knowledge about the states of nature, as follows: The decision maker
faces a problem under risk (uncertainty) if and only if the probabilities of
the states of nature are known (unknown). These notes concentrate on the
theory of decision making under risk. Thus, we shall presume that the
decision maker faces a decision problem wherein each act is associated
with a set of outcomes, and that the decision maker holds, for each act, a
probability distribution defined over the set of associated outcomes. We
will begin with the simplest case, wherein the decision maker has a
single probability system defined over the set of states of nature.

Decision theory posits that the human decision maker brings to the
resolution of a decision problem beliefs and preferences. Specifically, the
theory presumes that the decision maker possesses a probability system
(or systems) that captures his or her (partial) beliefs about nature's
selection of states of nature, a belief system about the outcomes accruing
to the performance of the acts in the various states of nature, and a
preference structure over the outcomes. Thus, a decision maker is
represented as a triple \(<P,F,U>\) composed of a probability measure \(P\), an
outcome mapping \(F\), and a utility function \(U\). The probability measure \(P\) is
defined over the set of states of nature, and captures the decision maker's
view of the random selection process used by nature. The outcome
mapping \(F\) is defined on the (Cartesian product of the) states of nature and
the acts, and presents the outcome resulting from performing each act in each of the states of nature. Thus, the outcome mapping $F$ generates a new set $O$ of outcomes. The utility function is defined over the set $O$ of outcomes and represents the decision maker's preferences over the outcomes.

For example, suppose the decision maker is considering the purchase of a lottery ticket. The probability that a given ticket will win is, say, one in one million. The payoff to the lottery is five million dollars if the ticket wins and nothing if it loses. The cost of a ticket is $5. The component parts of the problem are as follows:

<table>
<thead>
<tr>
<th>States of nature:</th>
<th>$s_1 = $ticket wins</th>
<th>$s_2 = $ticket loses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acts:</td>
<td>$a_1 = $buy a ticket</td>
<td>$a_2 = $do not buy a ticket</td>
</tr>
<tr>
<td>Probabilities:</td>
<td>$P(s_1) = 10^{-6}$</td>
<td>$P(s_2) = 1 - 10^{-6}$</td>
</tr>
<tr>
<td>Outcomes:</td>
<td>$F(s_1,a_1) = o_{11} = $4,999,995</td>
<td>$F(s_1,a_2) = o_{12} = $0</td>
</tr>
<tr>
<td></td>
<td>$F(s_2,a_1) = o_{21} = -$5</td>
<td>$F(s_2,a_2) = o_{22} = $0</td>
</tr>
</tbody>
</table>

For simplicity, we presume that the utility of money is measured in monetary terms, so that the utility function $U$ is the identity function. That is, the utility function is as follows:

| Utilities:        | $U(o_{11}) = 4,999,995$ | $U(o_{12}) = 0$ |
|                   | $U(o_{21}) = -$5         | $U(o_{22}) = 0$ |

Each of the foregoing components requires further discussion. Probability theory and utility theory have been covered in earlier courses. What follows here is a brief review. Outcome mappings are new, but are particularly simple.
BASIC PROBABILITY THEORY
Probability theory is composed of the formal mathematics underlying the common sense account of random behavior. The formal theory, first presented in its contemporary form by Kolmogorov in 1933 (see Kolmogorov, 1950), is as follows:

A probability system is a triple <S,Q,S,P> where S is the ground set, Q_S is a field of sets containing S, and P is a mapping from Q_S to the real number system. A probability system satisfies the following three axioms:

Axiom 1. \( P(X) > 0 \) for all X in Q_S.
Axiom 2. \( P(S) = 1 \).
Axiom 3. For all X and Y in Q_S, if \( X \cap Y = \emptyset \), then \( P(X \cup Y) = P(X) + P(Y) \).

An example is provided by the familiar coin tossing model. Suppose you are about to toss a coin once. The coin will either show heads (H) or tails (T). The ground set S is the set \{H,T\} and the field Q_S is the set \{\emptyset, \{H\}, \{T\}, \{H,T\}\}. The values of the probability function P depend upon our knowledge and beliefs about the coin. If the coin is believed to be fair, then \( P(\{H\}) = P(\{T\}) = 1/2 \). By the axioms, clearly \( P(S) = P(\{H,T\}) = 1 \) and \( P(\emptyset) = 0 \). Thus, the axioms determine only the assignment of the extreme values of 0 and 1 to the null set and ground set. The assignment of values to all of the other events is determined by the decision maker in accordance with the axioms and algebra of probability theory.

These three simple axioms provide the foundation for a very elaborate and powerful theory. Further concepts can be introduced via definitions and interrelated via theorems. The following concepts will be important in the presentation of basic decision theory.

Events
Formal probability theory is usually stated in terms of sets. So stated, the elements of the field Q_S are sets and are called events. In the foregoing example there are four events, as follows:
- the event "a head" denoted by \{H\},
- the event "a tail" denoted by \{T\},
- the sure event "either a head or a tail" denoted by \{H,T\}, and
the impossible event "neither a head nor a tail" denoted by \( \emptyset \).

Simple Events
The singleton sets in \( Q_S \) are called simple events. In our example, the simple events are heads ({H}) and tails ({T}). The simple events are very important. If probabilities can be assigned to each of the simple events, then probabilities can be calculated for all of the remaining events. This follows from Axiom 3. For example, suppose an event has three distinct elements, say x, y, and z. Then the event is \( \{x,y,z\} \). The event can be constructed as the union of the simple events \( \{x\} \), \( \{y\} \), and \( \{z\} \), i.e., \( \{x,y,z\} = \{x\} \cup \{y\} \cup \{z\} \). Then, by a simple extension of Axiom 3, \( P(\{x,y,z\}) = P(\{x\}) + P(\{y\}) + P(\{z\}) \), since \( \{x\} \), \( \{y\} \), and \( \{z\} \) pairwise disjoint since they have no elements in common.

Further, in what follows we will be interested only in probability systems that contain the simple events so that the probability measure is defined over each state of nature. Therefore, in what follows the field of sets \( Q_S \) will always be the power set of \( S \). The power set of \( S \) is simply the set of all subsets of \( S \), and if the number of elements in \( S \) is \( n \), then the number of elements in the power set is \( 2^n \).

Compound Events
An event is a compound event if it contains two or more elements. The compound event in our example is the sure event \( \{H,T\} \). (Note that the sure event and the ground set are identical.)

General Addition Rule
Axiom 3 is a special addition rule. The antecedent condition of Axiom 3 requires that the events X and Y have no element in common, i.e., be mutually exclusive. The General Addition Rule accounts for events that are not mutually exclusive, and is as follows: For all events X and Y, 
\[
P(X \cup Y) = P(X) + P(Y) - P(X \cap Y).
\]
In order to use the General Addition Rule it is necessary to calculate the probability of the joint event \( X \cap Y \). In general, this is fairly easy. For example, suppose \( X = \{x,y\} \) and \( Y = \{y,z\} \). Then \( X \cap Y = \{y\} \). Since this is a simple event, the probability of \( X \cap Y \) is known. Now, for example, suppose
\( X = \{w,x,y,z\} \) and \( Y = \{x,z\} \). Then, \( X \cap Y = \{x,z\} \) and Axiom 3 is used to calculate \( P(X \cap Y) \) as \( P(\{x\}) + P(\{z\}) \).

The foregoing elements of probability theory will suffice for the presentation of basic decision theory. As was noted above, the theory of probability is very elaborate and goes well beyond the few concepts and rules discussed above.

**ELEMENTS OF BASIC UTILITY THEORY**

Utility theory has evolved over the recent centuries and the contemporary form of the theory was developed in response to the work of Daniel Bernoulli published in early 1738. Therein Daniel Bernoulli solved the puzzle, known as the St. Petersburg Paradox, raised by his cousin Nicholas Bernoulli on 1713. The St. Petersburg Paradox, covered in greater detail in Chapter IV, is as follows:

Consider a simple game played with a fair coin where the coin will be flipped until it comes up tails, whereupon the game terminates. The payoff to the player is $2 if the sequence is HT, $4 if HHT, $8 if HHHT, and so forth. Thus, the player receives $2^n if the sequence is \( n \) Hs followed by a T.

The expected monetary value of the game is infinite, and therefore if the value of a game is its expected value, then rational individuals should be willing to pay very large amounts to play. However, Nicholas Bernoulli observed that most individuals place a small finite value on the game. Daniel Bernoulli concluded that the correct value of a game is not the expected monetary value, but rather the expected utility-of-money value, and that the utility of wealth function is not linear. More specifically, by "arguing that incremental utility is inversely proportional to current fortune (and directly proportional to the increment in fortune), Bernoulli concluded that utility is a linear function of the logarithm of monetary price, and showed that in this case the moral expectation of the game is finite" (Zabell 1990, p. 13) As we shall see, Bernoulli had made a mistake in the details. However, the idea that utility of wealth is not linear in wealth is a major contribution.
Various forms of the utility of wealth function are of interest, including the following:

- the quadratic function: \( U(x) = a + bx - cx^2 \) for \( a, b, c > 0 \);
- the logarithmic function: \( U(x) = \log(x+b) \) for \( b > 1 \);
- the simple exponential function: \( U(x) = -ae^{-bx} \) for \( a, b > 0 \);
- the related exponential function \( U(x) = ab^{-x} \) for \( a > 0, b > 1 \), or \( a < 0, 0 < b < 1 \);
- the normed exponential function \( U(x) = a(1-e^{-bx}) \) for \( a, b > 0 \); and
- the normed logistic function: \( U(x) = \frac{e^{a+bx}}{1+e^{a+bx}} \) for \( a, b > 0 \).

We will find that most are flawed, and that only the latter two make sense with respect to the analysis of risk preference.

Utility theory is well developed, and admits of numerous axiom systems. Ramsey (1931) first outlined a theory of expected utility in 1926. von Neumann and Morgenstern (1947) presented the first formalization that proves the existence of an expected utility function from a system of axioms. Marschak (1950) refined and simplified the axiomatization, and restricted the analysis to the case of finitely many outcomes. Herstein and Milnor (1953) further refined the formalization and extended the analysis to infinitely many outcomes. Savage (1954) axiomatized utility and probability together. Many formal axiom systems now exist. See, for example, Fishburn (1970) for specific systems, and Krantz, Luce, Suppes, and Tversky (1971) for the details of utility theory within the context of general measurement theory. See Eatwell, Milgate, and Newman (1990) for a general overview of probability and utility as they relate to decision theory. The theory has also been extended by Kahneman and Tversky (1979, 1992).

The following system is an adaptation of the original von Neumann-Morgenstern system, and is taken from the textbook *Mathematical Psychology* by Coombs, Dawes, and Tversky, pp. 122-126.
An Axiom System for Utility Theory

Let $O$ be the set of outcomes with typical elements $x$, $y$, $z$, and $w$. A gamble is a mixture of outcomes and probabilities, and is denoted $[x, p, y]$ if outcome $x$ is obtained with probability $p$ and outcome $y$ is obtained with probability $1-p$. $O^*$ is the augmented set of outcomes and contains all of the outcomes contained in $O$ and all of the gambles on those outcomes.

The axioms concern the system $(O^*, \succeq^*)$ where $O^*$ is the augmented set of outcomes and $\succeq^*$ is an ordering on $O^*$. The ordering $\succeq^*$ supports two other orderings, $\equiv^*$ and $\succ^*$, as follows:

$x \equiv^* y$ if and only if $x \succeq^* y$ and $y \succeq^* x$, and
$x \succ^* y$ if and only if $x \succeq^* y$ and not $y \succeq^* x$.

The axioms presented below support two theorems. The representation theorem asserts that the ordering $\succeq^*$ on $O^*$ can be represented by a real valued function; the uniqueness theorem asserts that the real valued representation is an interval scale.

The Axioms

Axiom 1. If $x$ and $y$ are elements of $O$, then for all probabilities $p$, $0 < p < 1$, the gamble $[x, p, y]$ is an element of $O^*$.

Axiom 2. $\succeq^*$ is a weak order on $O^*$, i.e.,
(i) for all $x$ in $O^*$, $x \succeq^* x$,
(ii) for all $x$ and $y$ in $O^*$, $x \succeq^* y$ or $y \succeq^* x$
    or $x \equiv^* y$,
(iii) for all $x$, $y$, and $z$ in $O^*$,
    if $x \succeq^* y$ and $y \succeq^* z$, then $x \succeq^* z$.

Axiom 3. For all $x$ and $y$ in $O^*$ and all probabilities $p$ and $q$, $0 < p, q < 1$,
$[[x, p, y], q, y] \equiv^* [x, pq, y]$.

Axiom 4. For all $x$, $y$, and $z$ in $O^*$ and all probabilities $p$, $0 < p < 1$,
if $x \equiv^* y$, then $[x, p, z] \equiv^* [y, p, z]$.

Axiom 5. For all $x$ and $y$ in $O^*$ and all probabilities $p$, $0 < p < 1$,
if x >* y, then x >* [x, p, y] >* y.

Axiom 6. For all x, y, and z in O*, if x >* y >* z, then there exists a probability p, 0 < p < 1, such that y =* [x, p, z].

The Representation Theorem
If the system \((O^*, \geq^*)\) satisfies axioms 1 through 6, then there exists a function U from \((O^*, \geq^*)\) to \((\text{Reals}, \geq)\) such that
(i) \(x \geq^* y\) if and only if \(U(x) \geq U(y)\), and
(ii) \(U([x, p, y]) = pU(x) + (1-p)U(y)\).

The Uniqueness Theorem
If U and V are two functions satisfying (i) and (ii) of the Representation Theorem, then there exist real numbers a and b such that a > 0 and for all x in O*, \(V(x) = aU(x) + b\).

Proofs of these theorems are available in von Neumann and Morgenstern (1947), Savage (1954), and elsewhere. A novel and somewhat informal proof is available in Arrow (1963), pp. 14-27.

A Note on the Uniqueness Theorem
The uniqueness theorem for utility functions states that if an individual's preference structure can be represented by two utility functions, say \(U(x)\) and \(V(x)\), then \(U\) and \(V\) are related by a positive affine transformation. (Note that \(aU(x) + b\) is not a linear function; \(aU(x)\) is a linear function.) That is, there exist two numbers \(a\) and \(b\), with \(a > 0\), such that for \(x\), \(V(x) = aU(x) + b\). You have seen this kind of a transformation before. It was used to transform temperature scales. Specifically, if \(f(x)\) is the temperature of \(x\) measured on the Fahrenheit scale and \(c(x)\) is the temperature measured on the centigrade scale, then \(f(x) = (9/5)c(x) + 32\).

It is instructive to recall the derivation of the Fahrenheit/centigrade transformation. Consider two thermometers, as shown below, one in the Fahrenheit scale and the other in the centigrade scale.
The procedure for relating the two thermometers is based on the proportionality of differences, as follows:

\[
\frac{f - 32}{212 - 32} = \frac{c - 0}{100 - 0}
\]

which simplifies to

\[
f = \frac{180}{100} c + 32 = \frac{9}{5} c + 32.
\]

The same scheme is used to make interpolations in tables of numbers like the Normal(0,1) table.

Setting Utility Values

If a measurement function is unique up to an affine transformation, as are utility and temperature measurement functions, then the functions themselves admit of two arbitrarily selected values. These values are called the zero and the unit. For example, in measuring temperature we can arbitrarily select the numerical value assigned to freezing water and the numerical value assigned to boiling water. In the Fahrenheit scale these values are 32 and 212, whereas in the centigrade scale they are 0 and 100. The zeros are 32 and 0, and the units are 180 (= 212 - 32) and 100 (= 100 - 0), respectively.

In setting the values of a utility function we assign convenient numerical values to the worst and best outcomes. Then the values assigned to all
other outcomes must fall between these two numbers. Further, we can arrange a series of hypothetical gambles that will reveal the utility assignments for sufficiently many points on the utility function to allow us to fit a curve.

Consider a simple choice between a lottery paying $100 with probability p and $0 with probability 1-p, and a sure payment of $20. Since $0 and $100 are the worst and best outcomes, we assign them convenient utility values. Let \( U(0) = 0 \) and \( U(100) = 10 \). Now, find the value of \( p \) at which the individual is indifferent between the lottery and the $20 for sure.

Suppose that probability is 0.3. Since the individual is indifferent between the lottery and the $20 for sure, clearly the expected utilities of the two must be equal. Thus,

\[
E[U(\text{lottery})] = (0.3)U(100) + (0.7)U(0)
\]

\[
= (0.3)(10) + (0.7)(0)
\]

\[
= E[U(\text{sure thing})]
\]

and \( E[U(\text{sure thing})] = U(20) \). Since these are equal, we have discovered that \( U(20) = 3 \). This procedure can be repeated to determine the values of arbitrarily many points on the utility function.

OUTCOME MAPPINGS

The outcome mapping represents the decision maker's view of what will happen given the performance of each of the acts in each of the states of nature. In short, the outcome mapping is a representation of the decision maker's foresight. Decision theory makes the very strong presumption of nearly perfect foresight on the part of the decision maker. In the lottery ticket example the specification of the outcome mapping was trivial. In many decision problems, the construction of the outcome mapping is very difficult, and may involve the use of the laws from the physical or social sciences, computation procedures from mathematics, or computations based on the equations of Finance or the tautologies of Accounting.

Consider the likely complexity of specifying the outcome mapping for guiding a deep space probe to a distant planet or for deciding whether or not to take over another firm. Oskar Morgenstern has made numerous comments about the role of perfect foresight in Economics and decision making. See, e.g., Morgenstern (1935, 1972). The presumption of perfect foresight constitutes a major flaw in the theory.
THE COMPUTATION PROCEDURE OF DECISION THEORY

It remains to discuss the second component of decision theory, the computation procedure. Consider a small generic decision problem \(<S,A>\) where \(S = \{s_1, s_2\}\) and \(A = \{a_1, a_2\}\). Also, consider a generic decision maker \(<P,F,U>\). The components of the decision maker are displayed as follows.

First, the components of the probability system underlying the probability measure \(P\) must be determined. The ground set of the system is the set \(S\) of states of nature. Let the field of sets \(Q_S\) be the power set of \(S\). That is, let \(Q_S = \{\emptyset, \{s_1\}, \{s_2\}, S\}\). Suppose further that the probability assignments are as follows:

\[
\begin{align*}
P(\emptyset) &= 0 \\
P(\{s_1\}) &= p_1 \\
P(\{s_2\}) &= p_2 \\
P(S) &= 1
\end{align*}
\]

The outcome mapping \(F\) can be presented in tabular form as follows:

<table>
<thead>
<tr>
<th>(s_i, a_j)</th>
<th>(a_1)</th>
<th>(a_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s_1)</td>
<td>(o_{11})</td>
<td>(o_{12})</td>
</tr>
<tr>
<td>(s_2)</td>
<td>(o_{21})</td>
<td>(o_{22})</td>
</tr>
</tbody>
</table>

where \(o_{ij} = F(s_i,a_j)\).

Similarly, the utility function \(U\) can be presented in tabular form as follows:

<table>
<thead>
<tr>
<th>(F(s,a))</th>
<th>(a_1)</th>
<th>(a_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s_1)</td>
<td>(u_{11})</td>
<td>(u_{12})</td>
</tr>
<tr>
<td>(s_2)</td>
<td>(u_{21})</td>
<td>(u_{22})</td>
</tr>
</tbody>
</table>

where \(u_{ij} = U(o_{ij})\).
The computation procedure is based upon the theorem informally established by Bernoulli (1738) and formally established by von Neumann and Morgenstern (1947). The theorem simply states that the decision maker behaves in an optimal way by selecting the act that maximizes the decision maker's expected utility. The expected utility assigned to each act is

\[ E[U(a_1)] = P(s_1)U(o_{11}) + P(s_2)U(o_{21}) \] for act \( a_1 \),

and

\[ E[U(a_2)] = P(s_1)U(o_{12}) + P(s_2)U(o_{22}) \] for act \( a_2 \).

In accordance with the Bernoulli/von Neumann-Morgenstern theorem, the decision maker simply selects the act with the greater expected utility.

THE RAINCOAT EXAMPLE - PART I
Suppose that as you are about to leave for the day the radio informs you that the probability of rain is 20%. Suppose further that you are scheduled for an interview and are wearing your good suit. Should you take your raincoat?

The decision problem is remarkably simple, and is as follows:

Decision problem = \(<S,A>\) where

\[ S = \{ \text{rain, no rain} \} \] and \( A = \{ \text{no raincoat, raincoat} \} \)

As the decision maker, you bring to the problem partial beliefs about the states of nature, a set of outcomes, and a preference structure over the outcomes. The probability measure is very simple, as follows:

\[ P(\text{rain}) = .20 \quad P(\text{no rain}) = .80 \]

The outcome mapping is more complex. If you are caught in the rain without a raincoat, then you will be wet for your interview and you will face the cost of laundering your clothes. Note that the payoff involves a non-monetary component (wet for the interview) and a monetary component (the cost of dry cleaning your clothes). The payoffs are as follows:
Let us suppose that this interview is very important to you and you are very eager to make a good impression. Thus, if it is going to rain, you definitely want to have your raincoat for protection. On the other hand, if it not going to rain, then you do not want to be wandering around carrying a superfluous raincoat.

You assess these outcomes as follows: the worst outcome is getting caught in the rain without the raincoat, and the next worst is carrying a superfluous raincoat. The best outcome is having the raincoat when it rains, and the second best outcome is leaving the raincoat home when it does not rain. The relative strengths of your preferences for these outcomes are such that if the worst is assigned 0 and the best 100, the second best is assigned 80 and the second worst is assigned 20. Thus, the utility table is as follows:

<table>
<thead>
<tr>
<th></th>
<th>no raincoat</th>
<th>raincoat</th>
</tr>
</thead>
<tbody>
<tr>
<td>rain</td>
<td>wet for interview &amp; cost of laundry</td>
<td>dry for interview &amp; no cost of laundry</td>
</tr>
<tr>
<td>no rain</td>
<td>dry for interview &amp; no raincoat to carry</td>
<td>dry for interview &amp; raincoat to carry</td>
</tr>
</tbody>
</table>

The expected utility calculations are then as follows:

\[
E[U(\text{no raincoat})] = (0.20)(0) + (0.80)(80) = 0 + 64 = 64,
\]

and
\[ E[U(\text{raincoat})] = (.20)(100) + (.80)(20) = 20 + 16 = 36. \]

Clearly, the preferred choice is to leave the raincoat at home.

This result may seem somewhat counterintuitive. After all, you argue, carrying a raincoat on a sunny day is a minor inconvenience compared to appearing for an important interview in a wet suit. There is no inconsistency. If this argument appeals to you, then the utility function displayed above is not yours. Your utility function would look more like the following:

<table>
<thead>
<tr>
<th></th>
<th>no raincoat</th>
<th>raincoat</th>
</tr>
</thead>
<tbody>
<tr>
<td>rain</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>no rain</td>
<td>80</td>
<td>60</td>
</tr>
</tbody>
</table>

Now the expected utility calculations are as follows:

\[ E[U(\text{no raincoat})] = (.20)(0) + (.80)(80) = 0 + 64 = 64, \]

and

\[ E[U(\text{raincoat})] = (.20)(100) + (.80)(60) = 20 + 48 = 68, \]

and you would take the raincoat with you.

This little ripple was added to the example to make a simple point. While the states of nature and the acts may be presented to the decision maker, the probabilities, outcomes, and utility assignments are endogenous to and inseparable from the decision maker. Indeed, the decision maker is the triple \(<P,F,U>\).
II. CONDITIONAL DECISION THEORY

INTRODUCTION
Conditional decision theory provides an account of decision making under uncertainty and risk given that the decision maker has access to information about the states of nature. Conditional decision theory constitutes the fundamental theory of decision making under uncertainty and risk.

Consider for the moment the role of information in decision making. Recall that the decision problem consists of a pair \( \langle S, A \rangle \) of sets of states of nature and acts, respectively. The decision maker is at risk because he or she does not know which state is the true state of nature. Thus, the information will have an influence on the decision maker's beliefs about the states of nature. More specifically, the influence of information is to revise the probabilities assigned to the states of nature. The revision process is governed by Bayes's Theorem, commonly giving rise to the misnomer 'Bayesian decision theory'.

MORE PROBABILITY THEORY
Conditional decision theory requires concepts from probability theory that are more sophisticated than the ones introduced in the previous chapter. Specifically, conditional decision theory requires the concept of conditional probability and Bayes's Theorem.

Conditional Probability
The conditional probability of an event \( X \) given an event \( Y \), denoted by \( P(X/Y) \) and read "the probability of \( X \) given \( Y \)," is defined as follows:

\[
P(X/Y) = \frac{P(X \cap Y)}{P(Y)}.
\]

As an example, consider the roll of a balanced die where \( X \) is the event of rolling a 1, i.e., \( X = \{1\} \). Then \( P(X) = 1/6 \). Now, let \( Y \) be the event of rolling an odd number, i.e., \( Y = \{1,3,5\} \). The probability of \( X \) given \( Y \), i.e., the probability of rolling a 1 given that you roll an odd number, is \( 1/3 \), as follows. By the foregoing formula, we have
\[ P(X/Y) = P(X \cap Y) / P(Y) \]
\[ = P\{1\} \cap \{1,3,5\} / P\{1,3,5\} \]
\[ = P\{1\} / P\{1,3,5\} \]
\[ = (1/6) / (1/2) \]
\[ = 1/3. \]

The concept of conditional probability provides the foundation for three familiar results -- the general multiplication rule, the concept of probabilistic independence, and the special multiplication rule.

**The General Multiplication Rule**
Usually treated as a theorem, the general multiplication rule is a rearrangement of the definition of conditional probability, and is as follows:

\[ P(X \cap Y) = P(Y)P(X/Y) = P(X)P(Y/X). \]

**Probabilistic Independence**
Two events X and Y are probabilistically independent if and only if \( P(X/Y) = P(X) \).

**The Special Multiplication Rule**
If two events X and Y are probabilistically independent, then \( P(X \cap Y) = P(X)P(Y) \).

Bayes's theorem, particularly in the extended form, plays a pivotal role in conditional decision theory. The theorem was first established in the simple form by Thomas Bayes (1763), and constitutes a minor extension of the definition of conditional probability.

**Bayes's Theorem - Simple Form**
If events \( X_1, \ldots, X_n \) are pairwise mutually exclusive (i.e., \( X_i \cap X_j = \emptyset \) for \( i \neq j \)) and event \( Y \subseteq X_1 \cup \ldots \cup X_n \), then for each \( X_k \),
\[ P(X_k/Y) = P(X_k)P(Y/X_k)/P(Y), \text{ for } k = 1, \ldots, n. \]

The foregoing theorem is easily extended, as follows:
Bayes's Theorem - Extended Form
If events $X_1, \ldots, X_n$ are pairwise mutually exclusive (i.e., $X_i \cap X_j = \emptyset$ for $i \neq j$) and event $Y \subseteq X_1 \cup \cdots \cup X_n$, then for each $X_k$,
$$P(X_k/Y) = \frac{P(X_k)P(Y/X_k)}{\sum_j P(X_j)P(Y/X_j)}, \text{ for } k = 1, \ldots, n.$$  

Note that the difference between the Simple and Extended forms of Bayes's Theorem rests on the simple result that if events $X_1, \ldots, X_n$ are pairwise mutually exclusive and event $Y \subseteq X_1 \cup \cdots \cup X_n$, then $P(Y) = \sum_j P(X_j)P(Y/X_j)$.

BAYESIAN REVISION OF PROBABILITIES
Note that a leap of sorts has been made in going from the expected utility measures of basic decision theory to those of conditional decision theory. The expected utility calculations of basic decision theory involved the probability measure $P(s)$ of the states, whereas the conditional expected utility measures of conditional decision theory are specified in terms of the conditional probability measures $P(s/z)$. In order to get from the probabilities $P(s_1)$ and $P(s_2)$ to the conditional probabilities $P(s_1/z_j)$ and $P(s_2/z_j)$ for each $z_j$ we simply use the extended form of Bayes's Theorem and the reliability probability measures $P(z_j/s)$ that characterize the quality of the information system. In the first instance
$$P(s_1/z_j) = \frac{P(s_1)P(z_j/s_1)}{[P(s_1)P(z_j/s_1) + P(s_2)P(z_j/s_2)]},$$
and in the second
$$P(s_2/z_j) = \frac{P(s_2)P(z_j/s_2)}{[P(s_1)P(z_j/s_1) + P(s_2)P(z_j/s_2)]}.$$

The probabilities $P(s)$ are often referred to as prior probabilities, and the probabilities $P(s/z)$ as posterior probabilities. The terms are derived from the Latin terms a priori and a posteriori, and refer to the probability assessments made by the decision maker prior to and posterior to the arrival of information. Better terminology replaces the former with initial and the latter with revised.
Jeffrey (1983) provides an alternative to the Bayesian revision of probabilities. Jeffrey's process is known as Probability Kinematics, and is used when there are concept shifts introduced by the information system. Dacey (1987) shows how Probability Kinematics, and not probability theory, provides the logic of strategic deception and manipulation.

THE RAINCOAT EXAMPLE - PART II
Consider the raincoat problem from the previous section, and recall that \( P(\text{rain}) = .20 \) and \( P(\text{no rain}) = .80 \). The weather report covers a very large geographical area and you would like a refined estimate of the probability of rain. You have a barometer affixed to the wall by the front door. Before leaving the house you plan to look at the barometer. How will you react to the barometer's reading?

Suppose that over time you have discovered that the barometer is quite accurate. On 90% of the days where in fact it has rained, the barometer reading was falling. Further, on 70% of the days where in fact it has remained dry, the barometer reading was rising.

You now have all of the details required to determine your rational reaction to the barometer reading. The setting is the following. You have heard the radio report of 20% probability of rain. You have not yet read the barometer. The barometer will display one of two readings. Either the barometric pressure is falling or it is rising. Thus, the possible signals from the barometer are 'falling' and 'rising'.

As noted above, the signal from the barometer will influence the probabilities of the states of nature. Also, as noted above, the influence will be made via Bayes's theorem. The calculations are as follows. On the basis of the radio report we have \( P(\text{rain}) = .20 \) and \( P(\text{no rain}) = .80 \). We need the following four probabilities: \( P(\text{rain/ falling}), P(\text{no rain/ falling}), P(\text{rain/ rising}), \) and \( P(\text{no rain/ rising}) \).

Consider the first. By the extended form of Bayes's theorem we have

\[ P(\text{rain/ falling}) = \]
\[ P(\text{rain})P(\text{falling/rain}) / [P(\text{rain})P(\text{falling/rain}) + P(\text{no rain})P(\text{falling/no rain})]. \]

Substitution yields

\[ P(\text{rain/falling}) = (.2)(.9) / [(.2)(.9) + (.8)(.3)] = .18 / .42. \]

Similarly, the remaining probabilities are as follows:

\[ P(\text{no rain/falling}) = (.8)(.3) / .42 = .24 / .42, \]
\[ P(\text{rain/rising}) = (.2)(.1) / [(.2)(.1) + (.8)(.7)] = .02 / .58, \text{ and} \]
\[ P(\text{no rain/rising}) = (.8)(.7) / .58 = .56 / .58 \]

Given the revised probabilities of the states of nature, we are prepared to determine the proper response to each of the barometer's possible signals. Recall that in the earlier version of this example there were two expected utility calculations, one for each act. Now there will be four, two for each signal form the barometer. Thus, to determine the proper reaction to a falling barometer we need two expected utility calculations, i.e., \( E[U(\text{coat})/\text{falling}] \) and \( E[U(\text{no coat})/\text{falling}] \). These calculations are as follows:

\[ E[U(\text{no coat})/\text{falling}] = (.18 / .42)(0) + (.24 / .42)(80) = 19.2 / .42, \text{ and} \]
\[ E[U(\text{coat})/\text{falling}] = (.18 / .42)(100) + (.24 / .42)(20) = 22.8 / .42. \]

Clearly, the proper response to a falling barometer is to take the coat.

The proper response to a rising barometer is determined in similar fashion. We require two expected utility calculations, specifically, \( E[U(\text{no coat})/\text{rising}] \) and \( E[U(\text{coat})/\text{rising}] \). These calculations are as follows:

\[ E[U(\text{no coat})/\text{rising}] = (.02 / .58)(0) + (.56 / .58)(80) = 44.8 / .58, \text{ and} \]
\[ E[U(\text{coat})/\text{rising}] = (.02 / .58)(100) + (.56 / .58)(20) = 13.2 / .58. \]
Clearly, the proper response to a rising barometer is to leave the coat at home.

Please note that according to the example, the decision maker has yet to examine the barometer. We have formulated a contingency plan, called a Bayes Strategy, with the following provisions: if the barometer is falling, then take the raincoat; if it is rising, then do not take the raincoat. The next step is to go to the barometer, determine the reading, and then act.
INTRODUCTION
The economic theory of information answers the following question -- How much is the information system \( <Z_S, P_{Z/S}> \) worth? The economic theory of information is in fact the economic theory of information systems. The world is awash in information systems. Consumer Reports and the Wall Street Journal are examples of the \( Z_S \) components of information systems. The former provides signals about consumer products, and the latter provides signals about the financial markets.

Similarly, high school transcripts and SAT or ACT scores comprise \( Z_S \) components of information systems about undergraduate student quality, while college transcripts and GMAT scores comprise information systems about graduate student quality. And, clearly, all accounting systems are information systems. All information systems have economic value. The economic theory of information specifies how that value is determined.

The economic theory of information provides an account of the economic evaluation of information systems, and is a extension of conditional decision theory. The contemporary economic theory of information was developed, in large measure, by Jacob Marschak, and has been popularized in the accounting literature by Demski. The economic theory of information is logically preceded by the theory of conditional decision making, as covered in the previous chapter.

ELEMENTS OF THE ECONOMIC THEORY OF INFORMATION
The fundamental calculation in the economic theory of information is very similar to an expected utility calculation from basic decision theory. Suppose, for example, that the Bayes Strategy specifies that if signal \( z_1 \) is received, then the decision maker is to select act \( a_1 \), and if \( z_2 \), then \( a_2 \). This means that the maximum expected value associated with \( z_1 \) is \( E[U(a_1)/z_1] \), and the maximum expected value associated with \( z_2 \) is \( E[U(a_2)/z_2] \). Recall also that the probabilities of \( z_1 \) and \( z_2 \), \( P(z_1) \) and \( P(z_2) \), respectively, are easily obtained as the denominators in the respective Bayes revisions of the probabilities of the states of nature. (To see this,
compare the simple and extended forms of Bayes's Theorem.) Thus, the expected (gross) value of the information system, \(E[Z_S, P_{Z/S}]\), is:

\[
E[Z_S, P_{Z/S}] = P(z_1)E[U(a_1)/z_1] + P(z_2)E[U(a_2)/z_2].
\]

In general, the calculation of the expected (gross) value of the information system is:

\[
E[Z_S, P_{Z/S}] = \sum_z P(z) \max_a E[U(a)/z].
\]

THE RAINCOAT EXAMPLE - PART III
The raincoat problem from the foregoing sections yielded the following Bayes Strategy: if \(z_1\), then \(a_2\), and if \(z_2\), then \(a_1\). The analysis is reproduced here for the reader's convenience.

<table>
<thead>
<tr>
<th></th>
<th>(a_1)</th>
<th>(a_2)</th>
<th>(P(s))</th>
<th>(P(z/s))</th>
<th>(P(s/z))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s_1)</td>
<td>0</td>
<td>100</td>
<td>.2</td>
<td>.9</td>
<td>.1</td>
</tr>
<tr>
<td>(s_2)</td>
<td>80</td>
<td>20</td>
<td>.8</td>
<td>.3</td>
<td>.7</td>
</tr>
</tbody>
</table>

The calculation for the expected (gross) value of the information system is then as follows:

\[
E[Z_S, P_{Z/S}] = (.42)(22.8/.42) + (.58)(44.8/.58) = 22.8 + 44.8 = 67.6.
\]
IV. ANALYSIS OF RISK ATTITUDE

The contemporary analysis of attitude toward risk begins with Daniel Bernoulli's formulation and resolution of the St. Petersburg Paradox. Bernoulli's work was published in 1738, in Latin, in the Papers of the Imperial Academy of Sciences in Petersburg. The title of the original paper is "Specimen Theoriae Novae de Mensura Sortis." An English translation, by Louise Sommer, entitled "Exposition of a New Theory on the Measurement of Risk," was published in 1954. The complete reference to these two papers is in the bibliography. The paper is fundamental to the contemporary view of risk attitude.

The St. Petersburg Paradox, strictly speaking, is not a paradox. It is a puzzle involving a simple game. The game is as follows: A fair coin will be flipped until it shows tails, whereupon the game will be terminated. If the coin shows n heads before showing the terminating tail, then the bettor will receive $2^n$ dollars. Bernoulli posed this game to various individuals, including professional and amateur gamblers. He never received a high bid. The puzzle lies in the absence of a high bid. At the time, the value of a game was presumed to be the expected value of the game. The lack of a high bid is paradoxical because the game has an infinite expected value.

Note that the expected value of the game is

$$EV(\text{game}) = \frac{1}{2}2 + \frac{1}{4}4 + \frac{1}{8}8 + \ldots + \frac{1}{2^n}2^n + \ldots,$$

which is an infinite sum of ones.

Bernoulli attacked the paradox by positing that the value of a game is not its expected value but rather its expected utility. He argued that the utility of wealth is increasing at a decreasing rate. Menger (1934) added the further requirement that the utility function be bounded above. Then $EV(\text{game})$ is replaced with $EU(\text{game})$, where

$$EU(\text{game}) = \frac{1}{2}U(2) + \frac{1}{4}U(4) + \frac{1}{8}U(8) + \ldots + \frac{1}{2^n}U(2^n) + \ldots,$$

which is finite. Bernoulli introduced the utility function $U$ to account for risk attitude. Bernoulli's new theory included the old theory, as follows: the value of a game is its expected value if and only if the decision maker is neutral to risk. Bernoulli surmised that individuals are in fact averse to risk, and thereby discount the value of large sums gained at long odds, and posited that the utility function has the basic shape of the
logarithmic function. As noted, Menger (1934) added the condition that a modest increase to an already large amount of money has very little incremental U-value.

To see the shape of the utility function for a risk averse individual note that such an individual would reject a bet that won or lost a given amount at fair odds. Thus, a risk averse individual with initial wealth \( W \) will choose \( W \) with probability 1 over a gamble paying \( W-h \) and \( W+h \) each with probability 1/2. Thus we have

\[
U(W) > (1/2)U(W-h) + (1/2)U(W+h),
\]

which is equivalent to

\[
U(W) - U(W-h) > U(W+h) - U(W).
\]

Thus, the increments to utility of equal increments to wealth are decreasing as wealth increases.

Risk averse individuals respond to risk by erecting various forms of protection, usually in the form of insurance. Because they are averse to risk, these individuals are willing to pay more than the actuarial value for insurance. To see this, consider a simple house insurance contract, for which the home owner pays a premium, that pays to replace the house in case of fire. Suppose that the premium is \( \rho \), the probability of loss by fire is \( P \), the value of the house is \( H \), and the decision maker's initial wealth is \( W \). Then the decision problem is as follows:

<table>
<thead>
<tr>
<th></th>
<th>payoffs</th>
<th>probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>insurance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no insurance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fire</td>
<td>( W-\rho )</td>
<td>( W-H )</td>
</tr>
<tr>
<td>no fire</td>
<td>( W-\rho )</td>
<td>( W )</td>
</tr>
</tbody>
</table>

The foregoing claim is as follows: If the home owner is risk averse, then there exists a premium \( \rho \) such that (i) \( \rho > PH \), and (ii) insurance is preferred to no insurance.
The expected values for the problem are as follows:

\[ EV(\text{insurance}) = W - \rho, \]
\[ EV(\text{no insurance}) = P(W-H) + (1-P)(W) = W - PH. \]

Thus, the actuarial value of the insurance is PH. The expected utilities for the problem are as follows:

\[ EU(\text{insurance}) = PU(W-\rho) + (1-P)U(W-\rho) = U(W-\rho), \]
\[ EU(\text{no insurance}) = PU(W-H) + (1-P)U(W). \]

These values are located in the figure below.

Note that for any premium \( \rho \) such that \( W^* < W - \rho < W - PH \), we have (i) \( \rho > PH \), and (ii) \( EU(\text{insurance}) > EU(\text{no insurance}) \). Thus, the foregoing claim that risk averse individuals will pay more than the actuarial value for insurance is established.

The three basic risk attitudes are risk aversion, risk neutrality, and risk preference. Risk aversion is viewed as middle class normalcy. Risk neutrality is a reasonable presumption if the wealth interval under consideration is short, so that the relevant segment of the utility function is nearly straight. Risk preference is usually presumed to be very rare, and typified by professional gamblers.
There are four relevant attitudes toward risk: (i) an individual is risk neutral if and only if the individual’s utility function is increasing at a constant rate; (ii) an individual is risk averse if and only if the individual’s utility function is increasing at a decreasing rate; (iii) an individual is risk preferring if and only if the individual’s utility function is increasing at an increasing rate; and (iv) an individual is risk averse/risk preferring if and only if the individual is risk averse over increases to wealth and risk preferring over decreases to wealth.

Risk aversion is generally viewed as the essence of middle class normalcy, and risk preference is viewed as an abnormality (Fishburn and Kochenberger 1979). Risk aversion and risk preferring behavior are regularly seen together, and various attempts have been made to explain their joint appearance. The principal analyses of hybrid risk attitudes are Battalio, Kagel, and Jiranyakul (1990), Battalio, Kagel, and MacDonald (1985), Camerer (1989), Fishburn and Kochenberger (1979), Friedman and Savage (1948), Kagel, MacDonald, and Battalio (1990), Kahneman and Tversky (1979). In particular, Battalio, Kagel, and Jiranyakul (1990) and Kahneman and Tversky (1979) show that approximately 80% of human experimental subjects are risk averse over gains and risk preferring over losses, i.e., risk averse/risk preferring. For most individuals, \( U(0) = 0 \) and the slope of the utility function is greater at \(-h\) than at \(h\), for all \(h > 0\) (Fishburn and Kochenberger 1979). An extended discussion of these findings is given in Neilson (1991). The graph is as follows:

The S-shaped utility function, and the risk attitude it represents, plays a major role in adapting basic utility theory to the findings of empirical
research in the area of human decision making (Kahneman and Tversky (1979), Tversky and Kahneman (1992)), and in explaining various forms of human behavior (Dacey (1991,1994)).

ARROW-PRATT ANALYSIS
Measures of risk aversion have been devised by Pratt (1964) and Arrow (1965, 1971) that allow a useful categorization of utility functions. The two basic measures are for absolute and relative risk aversion. The measures are as follows:

\[ R_{\text{absolute}} = -\frac{U''(x)}{U'(x)}, \]

\[ R_{\text{relative}} = -\frac{xU''(x)}{U'(x)}. \]

The Arrow-Pratt theory of risk aversion specifies that \( R_{\text{absolute}} \) is a decreasing function of \( x \), whereas \( R_{\text{relative}} \) is an increasing function of \( x \), and if \( R_{\text{relative}} \) is constant, then it is unity. (Arrow, 1965, pp. 35-37)

Note that if the individual is risk neutral, then \( U \) is a linear function of \( x \) and \( U''(x) = 0 \) for all \( x \), so that \( R_{\text{absolute}} = R_{\text{relative}} = 0 \). Further, \( R_{\text{absolute}} \) and \( R_{\text{relative}} \) are greater (less) than 0 if and only if the individual is risk averse (preferring). The class of constant absolute risk aversion utility functions is represented by the functions \( U(x) = ae^{-bx} \) and \( ab^{-x} \), where \( R_{\text{absolute}} = b \) and \( \ln b \), respectively. The class of constant relative risk aversion utility functions is represented by the function \( U(x) = a\log(x) \). Then, \( R_{\text{relative}} = 1 \). Note that the logarithmic functions, while favored by Bernoulli, are not bounded above.

The class of S-shaped utility functions can be composed via piecewise functions, as follows:

\[ U(x) = a(1-e^{-bx}) \text{ for } x > 0 \]

\[ U(x) = -c(1-e^{dx}) \text{ for } x < 0, \]

or

\[ U(x) = a(1-b^{-x}) \text{ for } x > 0 \]
\[ U(x) = -c(1-d^x) \] for \( x < 0 \),

where \( a, b, c, d > 0 \). Note that these functions are bounded above and below at \( a \) and \(-c\), respectively, and that \( U(0) = 0 \).

Another S-shaped utility function is provided by the logistic function. To retain the property that \( U(0) = 0 \), the functional form is

\[ U(x) = \frac{a[e^{bx} - 1]}{e^{bx} + 1}, \]

where \( a,b > 0 \). The function is bounded above and below at \( a \) and \(-a\), respectively. Furthermore,

\[
R_{\text{absolute}} = b\frac{e^{bx} - 1}{e^{bx} + 1} = abU(x),
\]

\[
R_{\text{relative}} = abxU(x).
\]

Therefore, both \( R_{\text{absolute}} \) and \( R_{\text{relative}} \) both are increasing in \( x \). The former makes the logistic function unacceptable.

A convenient constant absolute risk aversion utility function is

\[ U(x) = a(1-e^{-bx}) \] for \( x > 0 \),

and the piecewise S-shaped utility function

\[ U(x) = a(1-e^{-bx}) \] for \( x > 0 \]

\[ U(x) = -c(1-e^{dx}) \] for \( x < 0 \),

where \( a, b, c, d > 0 \) for illustrations. The former has been used by Cozzolino (1974) for illustration. For simplicity, in using the latter I will presume \( d = b \) and \( c > a \). These presumptions (i) set the level of risk aversion equal to the level of risk preference, and (ii) set the slope of the risk preferring segment at \(-x\) to a value that is greater than the value of the slope of the risk averse segment at \( x \), for all \( x \). The first presumption
is for convenience and simplicity, while the second reflects the findings of Fishburn and Kochenberger (1978).

BIBLIOGRAPHY


