Chapter 6

Choice Under Uncertainty

Up until now, we have been concerned with choice under certainty. A consumer chooses which commodity bundle to consume. A producer chooses how much output to produce using which mix of inputs. In either case, there is no uncertainty about the outcome of the choice.

We now turn to considering choice under uncertainty, where the objects of choice are not certainties, but distributions over outcomes. For example, suppose that you have a choice between two alternatives. Under alternative A, you roll a six-sided die. If the die comes up 1, 2, or 3, you get \$1000. If it comes up 4, 5, or 6, you lose \$300. Under alternative B, you choose a card from a standard 52 card deck. If the card you choose is black, you pay me \$200. If it is a heart, you get a free trip to Bermuda. If it is a diamond, you have to shovel the snow off of my driveway all winter.

If I were to ask you whether you preferred alternative A or alternative B, you could probably tell me. Indeed, if I were to write down any two random situations, call them L_1 and L_2, you could probably tell me which one you prefer. And, there is even the possibility that your preferences would be complete, transitive (i.e., rational), and continuous. If this is true then I can come up with a utility function representing your preferences over random situations, call it $U(L)$, such that L_1 is strictly preferred to L_2 if and only if $U(L_1) > U(L_2)$. Thus, without too much effort, we can extend our standard utility theory to utility under uncertainty. All we need is for the consumer to have well defined preferences over uncertain alternatives.

Now, recall that I said that much of what we do from a modeling perspective is add structure to people's preferences in order to be able to say more about how they behave. In this situation, what we would like to be able to do is say that a person's preferences over uncertain alternatives should be able to be expressed in terms of the utility the person would assign to the outcome if it were certain to occur, and the probability of that outcome occurring. For example, suppose we are considering two different uncertain alternatives, each of which offers a different distribution over three outcomes: I buy you a trip to Bermuda, you pay me \$500, or you paint my house. The probability of each outcome under alternatives A and B are given in the following table:

	Bermuda	**-\$500**	**Paint my house**
A	.3	.4	.3
B	.2	.7	.1

What we would like to be able to do is express your utility for these two alternatives in terms of the utility you assign to each individual outcome and the probability that they occur. For example, suppose you assign value u_B to the trip to Bermuda, u_m to paying me the money, and u_p to painting my house. It would be very nice if we could express your utility for each alternative by multiplying each of these numbers by the probability of the outcome occurring, and summing. That is:

$$\begin{aligned} U(A) &= 0.3u_B + 0.4u_m + 0.3u_p \\ U(B) &= 0.2u_B + 0.7u_m + 0.1u_p. \end{aligned}$$

Note that if this were the case, we could express the utility of any distribution over these outcomes in the same way. If the probabilities of Bermuda, paying me the money, and painting my house are p_B, p_m, and p_p, respectively, then the expected utility of the alternative is

$$p_B u_B + p_m u_m + p_p u_p.$$

This would be very useful, since it would allow us to base our study of choice under uncertainty on a study of choice over certain outcomes, extended in a simple way.

However, while the preceding equation, known as an **expected utility form**, is useful, it is not necessarily the case that a consumer with rational preferences over uncertain alternatives will be such that those alternatives can be represented in this form. Thus the question we turn to first is what additional structure we have to place on preferences in order to ensure that a person's preferences can be represented by a utility function that takes the expected utility form. After identifying these conditions, we will go on to show how utility functions of the expected utility form can be used to study behavior under uncertainty, and draw testable implications about people's behavior that are not implied by the standard approach.

6.1 Lotteries

In our study of consumer theory, the object of choice was a commodity bundle, x. In producer theory, the object of choice was a net input vector, y. In studying choice under uncertainty, the basic object of choice will be a **lottery**. A lottery is a probability distribution over a set of possible outcomes.

Suppose that there are N possible outcomes, denoted by $a_1, ..., a_N$. Let $A = \{a_1, ..., a_N\}$ denote the set of all possible outcomes. A **simple lottery** consists of an assignment of a probability to each outcome. Thus a **simple lottery** is a vector $L = (p_1, ..., p_N)$ such that $p_n \geq 0$ for $n = 1, ..., N$, and $\sum_n p_n = 1$.

A **compound lottery** is a lottery whose prizes are other lotteries. For example, suppose that I ask you to flip a coin. If it comes up heads, you roll a die, and I pay you the number of dollars that it shows. If the die comes up tails, you draw a random number between 1 and 10 and I pay you that amount of dollars. The set of outcomes here is $A = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10)$. The coin flip is then a lottery whose prizes are the lotteries $\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, 0, 0, 0, 0\right)$ and $\left(\frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}\right)$. Thus the coin flip represents a compound lottery. Notice that since the coin comes up heads or tails with probability $\frac{1}{2}$ each, the compound lottery can be reduced to a simple lottery:

$$
\begin{aligned}
& \frac{1}{2}\left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, 0, 0, 0, 0\right) + \frac{1}{2}\left(\frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}, \frac{1}{10}\right) \\
= \ & \left(\frac{2}{15}, \frac{2}{15}, \frac{2}{15}, \frac{2}{15}, \frac{2}{15}, \frac{2}{15}, \frac{1}{20}, \frac{1}{20}, \frac{1}{20}, \frac{1}{20}\right),
\end{aligned}
$$

where the final vector gives the probability of each outcome **before the coin is flipped.**

Generally, a compound lottery can be represented as follows. Suppose that there are K lotteries, denoted by $L_1, ..., L_K$. Let Q be a lottery whose prizes are the lotteries $L_1, ..., L_K$. That is, suppose that lottery Q awards the prize L_k with probability q_k. So, you can think of a compound lottery as a two-stage lottery. In the first stage, which lottery L_k you play in the second stage is determined. In the second stage, you play lottery L_k.

So, we call Q a compound lottery. It assigns probability q_k to L_k, where $q_k \geq 0$ and $\sum_k q_k = 1$. If p_n^k is the probability that L_k assigns to outcome n, this compound lottery can be reduced to a simple lottery where

$$p_n = \sum_k q_k p_n^k$$

is the probability of outcome n occurring. That is, p_n gives the probability of outcome n being the final outcome of the compound lottery before any of the randomizations have occurred.

If L and L' are lotteries, a compound lottery over these two lotteries can be represented as $aL + (1-a)L'$, where $0 \leq a \leq 1$ is the probability of lottery L occurring.

6.1.1 Preferences Over Lotteries

We begin by building up a theory of rational preferences over lotteries. Once we do that, we'll know that there is a utility function that represents those preferences (under certain conditions). We'll then go on to ask whether those preferences can be represented by a utility function of the expected utility form.

Let $\tilde{L}$ be the set of all possible lotteries. Thus $\tilde{L}$ is like X from consumer theory, the set of all possible consumption bundles. We want our consumer to have rational preferences over lotteries. So, suppose that the relation $\succsim$ represents strict preferences over lotteries, and suppose that these preferences are rational, i.e., complete and transitive.

We will also assume that the consumer's preferences are **consequentialist.** Basically, this means that consumers care only about the distribution over final outcomes, not whether this distribution comes about as a result of a simple lottery, or a compound lottery. In other words, the consumer is indifferent between any two compound lotteries that can be reduced to the same simple lottery. This property is often called **reduction of compound lotteries**.[1] Because of the reduction property, we can confine our attention to the set of all simple lotteries, and from now on we will let $\tilde{L}$ be the set of all simple lotteries.

The other requirement we needed for preferences to be representable by a utility function was continuity. In consumer theory, continuity meant that you could not pass from the lower level set of a consumer's utility function to the upper level set without passing through the indifference set. Something similar is involved with continuity here, but what we are interested in is continuity in probabilities.

Suppose that $L \succ L'$. The essence of continuity is that adding a sufficiently small probability of some other lottery, L'', to L should not reverse this preferences. That is:

$$\text{if } L \succ L', \text{ then } L \succ (1-a)L' + aL'' \text{ for some } a > 0.$$

[1] Another way to think about the reduction property is that we're assuming there is no process-oriented utility. Consumers do not enjoy the process of the gamble, only the outcome, eliminating the "fun of the gamble" in settings like casinos.

Formally, $\succsim$ on $\tilde{L}$ is continuous if for any L, L', and L'', and any $a \in (0,1)$, the sets:

$$\left\{a \in (0,1) \,|\, L \succsim (1-a)\,L' + aL''\right\}$$

and

$$\left\{a \in (0,1) \,|\, (1-a)\,L' + aL'' \succsim L\right\}$$

are closed.

Continuity is mostly a technical assumption needed to get the existence of a utility function. But, notice that its validity is less compelling than it was in the certainty case. For example, suppose L is a trip to Bermuda and L' is \$3000, and that you prefer L' to L. Now suppose we introduce L'', which is violent, painful, death. If preferences are continuous, then the trip to Bermuda should also be less preferred than \$3000 with probability $1-a$ and violent painful death with probability a, provided that a is sufficiently small. For many people, there is no probability $a > 0$ such that this would be the case, even though when $a = 0$, L' is preferred to L.

If the consumer has rational, continuous preferences over $\tilde{L}$, we know that there is a utility function $U()$ such that $U()$ represents those preferences. In order to get a utility function of the expected utility form that represents those preferences, the consumer's preferences must also satisfy the **independence axiom**.

The preferences relation $\succsim$ on $\tilde{L}$ satisfies the **independence axiom** if for all L, L', and L'' and $a \in (0,1)$, $L \succsim L'$ if and only if $aL + (1-a)\,L'' \succsim aL' + (1-a)\,L''$.

The essence of the independence axiom is that outcomes that occur with the same probability under two alternatives should not affect your preferences over those alternatives. For example, suppose that I offer you the choice between the following two alternatives:

$$\begin{array}{rcl} L & : & \$5 \text{ with probability } \frac{1}{5},\ 0 \text{ with probability } \frac{4}{5} \\ L' & : & \$12 \text{ with probability } \frac{1}{10},\ 0 \text{ with probability } \frac{9}{10}. \end{array}$$

Suppose you prefer L to L'. Now consider the following alternative. I flip a coin. If it comes up heads, I offer you the choice between L and L'. If it comes up tails, you get nothing. What the independence axiom says is that if I ask you to choose either L or L' before I flip the coin, your preference should be the same as it was when I didn't flip the coin. That is, if you prefer L to L', you should also prefer $\frac{1}{2}L + \frac{1}{2}0 \succsim \frac{1}{2}L' + \frac{1}{2}0$, where 0 is the lottery that gives you 0 with probability 1.

The previous example illustrates why the independence axiom is frequently called "independence of irrelevant alternatives." The irrelevant alternative is the event that occurs regardless of your choice. Thus, the independence axiom says that alternatives that occur regardless of what you choose should not affect your preferences.

Although the independence axiom seems straightforward, it is actually quite controversial. To illustrate, consider the following example, known as the Allais Paradox.

Consider a lottery with three possible outcomes: \$2.5 million, \$0.5 million, and \$0. Now, consider the following two lotteries (the numbers in the table are the probabilities of the outcome in the column occurring under the lottery in the row):

	$\$2.5M$	$\$0.5M$	$\$0$
L_1	0	1	0
$L_1^{'}$	.1	.89	.01

That is, L_1 offer \$500,000 for sure. $L_1^{'}$ offers \$2.5M with probability 0.1, \$500,000 with probability 0.89, and 0 with probability 0.01. Now, before going on, decide whether you would choose L_1 or L_1'.

Next, consider the following two alternative lotteries over the same prizes.

	$\$2.5M$	$\$0.5M$	0
L_2	0	0.11	0.89
$L_2^{'}$	.1	0	.9

It is not unusual for people to prefer L_1 to $L_1^{'}$, but $L_2^{'}$ to L_2. However, such behavior is a violation of the independence axiom. To see why, define lotteries $L_A = (0,1,0)\,, L_B = \left(\frac{10}{11}, 0, \frac{1}{11}\right)$, and $L_C = (0,0,1)$. Notice that

$$\begin{aligned} L_1 &= 0.89L_A + 0.11L_A \\ L_1^{'} &= 0.89L_A + 0.11L_B. \end{aligned}$$

Thus preferences over L_1 should be preferred to $L_1^{'}$ if and only if L_A is preferred to L_B.

Similarly, consider that L_2 and $L_2^{'}$ can be written as:

$$\begin{aligned} L_2 &= 0.11L_A + 0.89L_C \\ L_2^{'} &= 0.11L_B + 0.89L_C. \end{aligned}$$

Thus if this person satisfies the independence axiom, L_2 should be preferred to $L_2^{'}$ whenever L_A is preferred to L_B, which is the same as in the L_1 vs. $L_1^{'}$ case above. Hence if L_1 is preferred to $L_1^{'}$, then L_2 should also be preferred to $L_2^{'}$.

Usually, about half of the people prefer L_1 to $L_1^{'}$ but $L_2^{'}$ to L_2. Does this mean that they are irrational? Not really. What it means is that they do not satisfy the independence axiom. Whether or not such preferences are irrational has been a subject of debate in economics. Some people think yes. Others think no. Some people would argue that if your preferences don't satisfy the independence axiom, it's only because you don't understand the problem. And, once the nature of your failure has been explained to you, you will agree that your behavior should satisfy the independence axiom and that you must have been mistaken or crazy when it didn't. Others think this is complete nonsense. Basically, the independence axiom is a source of great controversy in economics. This is especially true because the independence axiom leads to a great number of paradoxes like the Allais paradox mentioned earlier.

In the end, the usefulness of the expected utility framework that we are about to develop usually justifies its use, even though it is not perfect. A lot of the research that is currently going on is trying to determine how you can have an expected utility theory without the independence axiom.

6.1.2 The Expected Utility Theorem

We now return to the question of when there is a utility function of the expected utility form that represents the consumer's preferences. Recall the definition:

Definition 9 *A utility function $U(L)$ has an expected utility form if there are real numbers $u_1, ..., u_N$ such that for every simple lottery $L = (p_1, ..., p_N)$,*

$$U(L) = \sum_n p_n u_n.$$

The reduction property and the independence axiom combine to show that **utility function $U(L)$ has the expected utility form if and only if it is linear**, meaning it satisfies the property:

$$U\left(\sum_{k=1}^{K} t_k L_k\right) = \sum_{k=1}^{K} t_k U(L_k) \tag{6.1}$$

for any K lotteries. To see why, note that we need to show this in "two directions" - first, that the expected utility form implies linearity; then, that linearity implies the expected utility form.

1. Suppose that $U(L)$ has the expected utility form. Consider the compound lottery $\sum_{k=1}^{K} t_k L_k$.
$$U\left(\sum_{k=1}^{K} t_k L_k\right) = \sum_n u_n \left(\sum_{k=1}^{K} t_k p_n^k\right) = \sum_{k=1}^{K} t_k \left(\sum_n u_n p_n^k\right) = \sum_{k=1}^{K} t_k U(L_k).$$
So, it is linear.

2. Suppose that $U(L)$ is linear. Let L^n be the lottery that awards outcome a_n with probability 1. Then
$$U(L) = U\left(\sum_n p_n L^n\right) = \sum_n p_n U(L^n) = \sum_n p_n u_n.$$
So, it has the expected utility form.

Thus proving that a utility function has the expected utility form is equivalent to proving it is linear. We will use this fact momentarily.

The **expected utility theorem** says that if a consumer's preferences over simple lotteries are rational, continuous, and exhibit the reduction and independence properties, then there is a utility function of the expected utility form that represents those preferences. The argument is by construction. To make things simple, suppose that there is a best prize, a_B, and a worst prize, a_W, among the prizes. Let L^B be the "degenerate lottery" that puts probability 1 on a_B occurring, and L^W be the degenerate lottery that puts probability 1 on a_W. Now, consider some lottery, L, such that $L^B \succ L \succ L^W$. By continuity, there exists some number, a_L, such that
$$a_L L^B + (1 - a_L) L^W \sim L.$$
We will define the consumer's utility function as $U(L) = a_L$ as defined above, and note that $U(L^B) = 1$ and $U(L^W) = 0$. Thus the utility assigned to a lottery is equal to the probability put on the best prize in a lottery between the best and worst prizes such that the consumer is indifferent between L and $a_L L^B + (1 - a_L) L^W$.

In order to show that $U(L)$ takes the expected utility form, we must show that:
$$U\left(tL + (1-t) L'\right) = t a_L + (1-t) a_{L'}.$$
If this is so, then $U()$ is linear, and thus we know that it can be represented in the expected utility form. Now, $L \sim a_L L^B + (1 - a_L) L^W$, and $L' \sim a_{L'} L^B + (1 - a_{L'}) L^W$. Thus:
$$\begin{aligned} & U\left(tL + (1-t) L'\right) \\ = \; & U\left(t\left(a_L L^B + (1-a_L) L^W\right) + (1-t)\left(a_{L'} L^B + (1-a_{L'}) L^W\right)\right) \end{aligned}$$

by the independence property. By the reduction property:

$$= U\left(\left(ta_L + (1-t)\,a_{L'}\right)L^B + \left(1 - \left(\left(ta_L + (1-t)\,a_{L'}\right)\right)\right)L^W\right)$$

and by the definition of the utility function:

$$= ta_L + (1-t)\,a_{L'}.$$

This proves that $U()$ is linear, and so we know that $U()$ can be written in the expected utility form.[2]

I'm not going to write out the complete proof. However, I am going to write out the expected utility theorem.

Expected Utility Theorem: *Suppose that rational preference relation $\succsim$ is continuous and satisfies the reduction and independence axioms on the space of simple lotteries $\tilde{L}$. Then $\succsim$ admits a utility function $U(L)$ of the expected utility form. That is, there are numbers $u_1, ..., u_N$ such that $U(L) = \sum_{n=1}^{N} p_n^L u_n$, and for any two lotteries,*

$$L \succsim L' \text{ if and only if } U(L) \geq U(L').$$

Note the following about the expected utility theorem:

1. The expected utility theorem says that under these conditions, there is one utility function, call it $U(L)$, of the expected utility form that represents these preferences.

2. However, there may be other utility functions that also represent these preferences.

3. In fact, any monotone transformation of $U()$ will also represent these preferences. That is, let $V()$ be a monotone transformation, then $V(U(L)) = V\left(\sum_n p_n^L u_n\right)$ also represents these preferences.

4. However, it is not necessarily the case that $V(U(L))$ can be written in the expected utility form. For example, $v = e^u$ is a monotone transformation, but there is no way to write $V(U(L)) = \exp\left(\sum_n p_n^L u_n\right)$ in the expected utility form.

5. But, there are some types of transformations that can be applied to $U()$ such that $V(U())$ also has the expected utility form. It can be shown that the transformed utility function also has the expected utility form if and only if $V()$ is linear. This is summarized as follows:

[2] This is not a formal proof, but it captures the general idea. There are some technical details that must be addressed in the formal proof, but you can read about these in your favorite micro text.

The Expected Utility Form is preserved only by positive linear transformations. *If $U()$ and $V()$ are utility functions representing $\succsim$, and $U()$ has the expected utility form, then $V()$ also has the expected utility form if and only if there are numbers $a > 0$ and b such that:*

$$U(L) = aV(L) + b.$$

In other words, the expected utility property is preserved by positive linear (affine) transformations, but any other transformation of $U()$ does not preserve this property.

MWG calls the utility function of the expected utility form a **von-Neumann-Morgenstern (vNM)** utility function, and I'll adopt this as well. That said, it is important that we do not confuse the vNM utility function, $U(L)$, with the numbers $u_1, ..., u_N$ associated with it.[3]

An important consequence of the fact that the expected utility form is preserved only by positive linear transformations is that a vNM utility function imposes cardinal significance on utility. To see why, consider the utility associated with four prizes, u_1, u_2, u_3, and u_4, and suppose that

$$u_1 - u_2 > u_3 - u_4.$$

Suppose we apply a positive linear transformation to these numbers:

$$v_n = au_n + b.$$

Then

$$\begin{aligned} v_1 - v_2 &= au_1 + b - (au_2 + b) = a(u_1 - u_2) \\ &> a(u_3 - u_4) = au_3 + b - (au_4 + b) = v_3 - v_4. \end{aligned}$$

Thus $v_1 - v_2 > v_3 - v_4$ if and only if $u_1 - u_2 > u_3 - u_4$. And, since any utility function of the expected utility form that represents the same preferences will exhibit this property, differences in utility numbers are meaningful. The numbers assigned by the vNM utility function have cardinal significance. This will become important when we turn to our study of utility for money and risk aversion, which we do next.

6.1.3 Constructing a vNM utility function

Let $A = \{a_1, ..., a_N\}$ denote the set of prizes. Suppose that a_N is the best prize and a_1 is the worst prize. We are going to show how to construct numbers $u_1, ..., u_N$ that make up a utility function with the expected utility form that represents preferences over these prizes.

[3] Later, when we allow for a continuum of prizes (such as monetary outcomes), the numbers $u_1, ..., u_N$ become the function $u(x)$, and we'll call the lowercase $u(x)$ function the Bernoulli utility function.

First, we are going to arbitrarily choose $u_N = 1$ and $u_1 = 0$. Why? Because we can. Remember, the point here is to construct a utility function that has the expected utility form. We could just as easily do it for arbitrary specifications of u_N and u_1, but this is notationally a bit simpler.

Now, for each prize a_i, define u_i to be the probability such that the decision maker is indifferent between prize a_i for sure and a lottery that offers a_N with probability u_i and a_1 with probability $1-u_i$. Let's refer to the lottery that offers prize a_N with probability u_i and prize a_1 with probability $(1-u_i)$ lottery S_i. So, u_i's are between 0 and 1. Notice that if we specify the numbers in this way it makes sense that $u_N = 1$ and $u_1 = 0$, since the decision maker should be indifferent between the best prize for sure and a lottery that offers the best prize with probability $u_1 = 1$, etc.

This gives us a way of defining numbers u_i. Now, we want to argue that this way of defining the u_i's, combined with consequentialism (reduction of compound lotteries) and Independence of Irrelevant alternatives yields a utility function that looks like $U(L) = \sum p_i u_i$.

So, consider lottery $L = (p_1, ..., p_N)$. This lottery offers prize a_i with probability p_i. But, we know that the decision maker is indifferent between a_i for sure and a lottery that offers prize a_N with probability u_i and price a_1 with probability $1-u_i$. Thus, using IIA, we know that the decision maker is indifferent between lottery L and a compound lotter in which, with probability p_i, the decision maker faces another lottery: u_N with probability u_i and u_1 with probability $1-u_i$. This lottery is depicted as L' in the following diagram. Note that L' only has two distinct prizes: a_N and a_1. By reduction of compound lotteries, we can combine the total probability of each outcome, making an equivalent simple lottery, L''. The utility for lottery L'' is $\left(\sum p_i u_i\right) u_N + \left(1 - \sum p_i u_i\right) u_1$. Since $u_N = 1$ and $u_1 = 0$, this gives that $U(L) = U(L') = U(L'') = \sum p_i u_i$, which is what we wanted to show. Defining utility in this way gives us a representation with the expected utility form.

6.2 Utility for Money and Risk Aversion

The theory of choice under uncertainty is most frequently applied to lotteries over monetary outcomes. The easiest way to treat monetary outcomes here is to let x be a continuous variable representing the amount of money received. With a finite number of outcomes, assign a number u_n to each of the N outcomes. We could also do this with the continuous variable, x, just by letting u_x be the number assigned to the lottery that assigns utility x with probability 1. In this

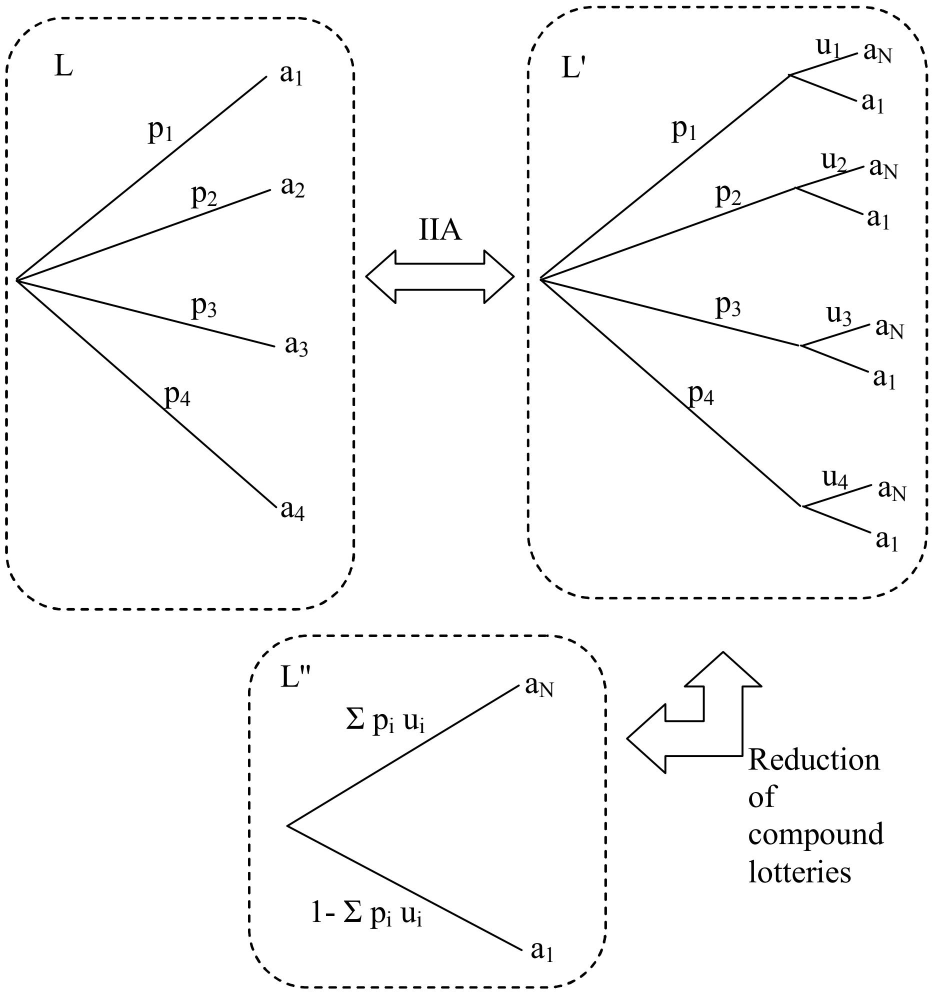
L
p1
a1
p2
a2
p3
a3
p4
a4
IIA
L'
u1
aN
a1
u2
aN
a1
u3
aN
a1
u4
aN
a1
L"
Σ pi ui
aN
1- Σ pi ui
a1
Reduction of compound lotteries

Figure 6.1:

case, there would be one value of u_x for each real number, x. But, this is just what it means to be a function. So, we'll let the function $u(x)$ play the role that u_n did in the finite outcome case. Thus $u(x)$ represents the utility associated with the lottery that awards the consumer x dollars for sure.

Since there is a continuum of outcomes, we need to use a more general probability structure as well. With a discrete number of outcomes, we represented a lottery in terms of a vector $(p_1, ..., p_N)$, where p_n represents the probability of outcome n. When there is a continuum of outcomes, we will represent a lottery as a distribution over the outcomes. One concept that you are probably familiar with is using a probability density function $f(x)$. When we had a finite number of outcomes, we denoted the probability of any particular outcome by p_n. The analogue to this when there are a continuous number of outcomes is to use a probability density function (pdf). The pdf is defined such that:

$$\Pr(a \leq x \leq b) = \int_a^b f(x)\,dx.$$

Recall that when a distribution can be represented by a pdf, it has no atoms (discrete values of x with strictly positive probability of occurring). Thus the probability of any particular value of x being drawn is zero. The expected utility of a distribution $f()$ is given by:

$$U(f) = \int_{-\infty}^{+\infty} u(x) f(x)\,dx,$$

which is just the continuous version of $U(L) = \sum_n p_n u_n$. In order to keep things straight, we will call $u(x)$ the Bernoulli utility function, while we will continue to refer to $U(f)$ as the vNM utility function.

It will also be convenient to write a lottery in terms of its cumulative distribution function (cdf) rather than its pdf. The cdf of a random variable is given by:

$$F(b) = \int_{-\infty}^{b} f(x)\,dx.$$

When we use the cdf to represent the lottery, we'll write the expected utility of F as:

$$\int_{-\infty}^{+\infty} u(x)\,dF(x).$$

Mathematically, the latter formulation lets us deal with cases where the distribution has atoms, but we aren't going to worry too much about the distinction between the two.

The Bernoulli utility function provides a convenient way to think about a decision maker's attitude toward risk. For example, consider a gamble that offers \$100 with probability $\frac{1}{2}$ and 0 with probability $\frac{1}{2}$. Now, if I were to offer you the choice between this lottery and c dollars for sure, how small would c have to be before you are willing to accept the gamble?

The expected value of the gamble is $\frac{1}{2}100 + \frac{1}{2}0 = 50$. However, if offered the choice between 50 for sure and the lottery above, most people would choose the sure thing. It is not until c is somewhat lower than 50 that many people find themselves willing to accept the lottery. For me, I think the smallest c for which I am willing to accept the gamble is 40. The fact that $40 < 50$ captures the idea that I am **risk averse**. My expected utility from the lottery is less than the utility I would receive from getting the expected value of the gamble for sure. The minimum amount c such that I would accept the gamble instead of the sure thing is known as the **certainty equivalent** of the gamble, since it equals the certain amount of money that offers the same utility as the lottery. The difference between the expected value of the lottery, 50, and my certainty equivalent, 40, is known as my **risk premium**, since I would in effect be willing to pay somebody 10 to take the risk away from me (i.e. replace the gamble with its expected value).

Formally, let's define the certainty equivalent. Let $c(F, u)$ be the certainty equivalent for a person with Bernoulli utility function u facing lottery F, defined according to:

$$u(c(F, u)) = \int u(x)\, dF(x)$$

Although generally speaking people are risk averse, this is a behavioral postulate rather than an assumption or implication of our model. But, people need not be risk averse. In fact, we can divide utility functions into four classes:

1. Risk averse. A decision maker is risk averse if the expected utility of any lottery, F, is not more than the utility of the getting the expected value of the lottery for sure. That is, if:

$$\int u(x)\, dF(x) \leq u\left(\int x dF(x)\right) \text{ for all } F.$$

   (a) If the previous inequality is strict, we call the decision maker strictly risk averse.

   (b) Note also that since $u(c(F, u)) = \int u(x)\, dF(x)$ and $u()$ is strictly increasing, an equivalent definition of risk aversion is that the certainty equivalent $c(F, u)$ is no larger than the expected value of the lottery, $\int x dF(x)$ for any lottery F.

2. Risk loving. A decision maker is risk loving if the expected utility of any lottery is not less than the utility of getting the expected value of the lottery for sure:

$$\int u(x)\, dF(x) \geq u\left(\int x dF(x)\right).$$

(a) Strictly risk loving is when the previous inequality is strict.

(b) An equivalent definition is that $c(F,u) \geq \int x dF(x)$ for all F.

3. Risk neutral. A decision maker is risk neutral if the expected utility of any lottery is the same as the utility of getting the expected value of the lottery for sure:

$$\int u(x)\, dF(x) = u\left(\int x dF(x)\right).$$

(a) An equivalent definition is $c(F,u) = \int x dF(x)$.

4. None of the above. Many utility functions will not fit into any of the cases above. They'll be risk averse, risk loving, or risk neutral depending on the lottery involved.

Although many utility functions will fit into the "none of the above" category, risk aversion is by far the most natural way to think about actual people behaving, with the limiting case of risk neutrality. So, most of our attention will be focused on the cases of risk neutrality and risk aversion. Risk loving behavior does arise, but generally speaking people are risk averse, and so we start our study there.

Consider again the definition of risk aversion:

$$\int u(x)\, dF(x) \leq u\left(\int x dF(x)\right).$$

It turns out that this inequality is a version of Jensen's Inequality, which says that $h()$ is a concave function if and only if

$$\int h(x)\, dF(x) \leq h\left(\int x dF(x)\right).$$

for all distributions $F()$. Thus, risk aversion on the part of the decision maker is equivalent to having a concave Bernoulli utility function. Strict risk aversion is equivalent to having a strictly concave Bernoulli utility function.

Similarly, (strict) risk loving is equivalent to having a (strictly) convex Bernoulli utility function, and risk neutrality is equivalent to having a linear Bernoulli utility function.

The utility functions of risk averse and risk neutral decision makers are illustrated in MWG Figure 6.C.2. (panels a and b). In panel a, a risk averse consumer is diagrammed. Notice that with a strictly concave utility function, the expected utility of the lottery that offers 3 or 1 with equal probability is $\frac{1}{2}u(1) + \frac{1}{2}u(3) < u\left(\frac{1}{2}1 + \frac{1}{2}3\right)$. On the other hand, in panel b, $\frac{1}{2}u(1) + \frac{1}{2}u(3) = u\left(\frac{1}{2}1 + \frac{1}{2}3\right)$; the consumer is indifferent between the gamble and the sure thing. Thus the manifestation of risk aversion in panel a is in the fact that the dotted line between $(1, u(1))$ and $(3, u(3))$ lies everywhere below the utility function.

To see if you understand, draw the diagram for a risk-loving decision maker, and convince yourself that $\frac{1}{2}u(1) + \frac{1}{2}u(3) > u\left(\frac{1}{2}1 + \frac{1}{2}3\right)$.

6.2.1 Measuring Risk Aversion: Coefficients of Absolute and Relative Risk Aversion

As we said, risk aversion is equivalent to concavity of the utility function. Thus one would expect that one utility function is "more risk averse" than another if it is "more concave." While this is true, it turns out that measuring the risk aversion is more complicated than you might think (isn't everything in this course?). Actually, it is only slightly more complicated.

You might be tempted to think that a good measure of risk aversion is that Bernoulli utility function $u_1()$ is more risk averse than Bernoulli utility function $u_2()$ if $|u_1''()| > |u_2''()|$ for all x. However, there is a problem with this measure, in that it is not invariant to positive linear transformations of the utility function. To see why, consider utility function $u_1(x)$, and apply the linear transformation $u_2() = au_1() + b$, where $a > 1$. We know that such a transformation leaves the consumer's attitudes toward risk unchanged. However, $u_2''() = au_1''() > u_1''()$. Thus if we use the second derivative of the Bernoulli utility function as our measure of risk aversion, we find that it is possible for a utility function to be more risk averse than another, even though it represents the exact same preferences. Clearly, then, this is not a good measure of risk aversion.

The way around the problem identified in the previous paragraph is to normalize the second derivative of the utility function by the first derivative. Using $u_2()$ from the previous paragraph, we then get that:

$$\frac{u_2''()}{u_2'()} = \frac{au_1''()}{au_1'()} = \frac{u_1''()}{u_1'()}.$$

Thus this measure of risk aversion is invariant to linear transformations of the utility function. And, it's almost the measure we will use. Because $u'' < 0$ for a concave function, we'll multiply by -1 so that the risk aversion number is non-negative for a risk-averse consumer. This gives us the following definition:

Definition 10 *Given a twice-differentiable Bernoulli utility function $u()$, the Arrow-Pratt measure of absolute risk aversion is given by:*

$$r_A(x) = -\frac{u''(x)}{u'(x)}.$$

Note the following about the Arrow-Pratt (AP) measure:

1. $r_A(x)$ is positive for a risk-averse decision maker, 0 for a risk-neutral decision maker, and negative for a risk-loving decision maker.

2. $r_A(x)$ is a function of x, where x can be thought of as the consumer's current level of wealth. Thus we can admit the situation where the consumer is risk averse, risk loving, or risk neutral for different levels of initial wealth.

3. We can also think about how the decision maker's risk aversion changes with her wealth. How do you think this should go? Do you become more or less likely to accept a gamble that offers 100 with probability $\frac{1}{2}$ and -50 with probability $\frac{1}{2}$ as your wealth increases? Hopefully, you answered more. This means that you become less risk averse as wealth increases, and this is how we usually think of people, as having non-increasing absolute risk aversion.

4. The AP measure is called a measure of absolute risk aversion because it says how you feel about lotteries that are defined over absolute numbers of dollars. A gamble that offers to increase or decrease your wealth by a certain percentage is a relative lottery, since its prizes are defined relative to your current level of wealth. We also have a measure of **relative risk aversion**,
$$r_R(x) = -\frac{xu''(x)}{u'(x)}.$$
But, we'll come back to that later.

6.2.2 Comparing Risk Aversion

Frequently it is useful to know when one utility function is more risk averse than another. For example, risk aversion is important in the study of insurance, and a natural question to ask is how a person's desire for insurance changes as he becomes more risk averse. Fortunately, we already have the machinery in place for our comparisons. We say that utility function $u_2()$ is at least as risk averse as $u_1()$ if any of the following hold (in fact, they are all equivalent):

1. $c(F, u_2) \leq c(F, u_1)$ for all F.

2. $r_A(x, u_2) \geq r_A(x, u_1)$

3. $u_2()$ can be derived from $u_1()$ by applying an increasing, concave transformation, i.e., $u_2() = g(u_1(x))$, where $g()$ is increasing and concave. Note, this is what I meant when I said being

more risk averse is like being more concave. However, as you can see, this is not the most useful of the definitions we have come up with.

4. Starting from any initial wealth position, x_0, any lottery F that would be at least as good as x_0 for certain to a person with utility function $u_2()$ would also be acceptable to a person with utility function $u_1()$. That is,

$$\text{if } u_2(x_0) \leq \int u_2(x)\, dF(x)\text{, then } u_1(x_0) \leq \int u_1(x)\, dF(x)\,.$$

Note that in MWG, they give definitions for "more risk averse" rather than "at least as risk averse." Usually, you can go from what I say is "at least as risk averse" to something that is "more risk averse" by simply making the inequality strict for some value. That is, $u_2()$ is more risk averse than $u_1()$ if:

1. $c(F, u_2) \leq c(F, u_1)$ for all F, with strict inequality for some F.

2. $r_A(x, u_2) \geq r_A(x, u_1)$ for all x, with strict inequality for some x.

3. $u_2()$ can be derived from $u_1()$ by applying an increasing, **strictly** concave transformation, i.e., $u_2() = g(u_1(x))$, where $g()$ is increasing and concave.

4. Starting from any initial wealth position, x_0, any lottery F that would be at least as good as x_0 for certain to a person with utility function $u_2()$ would be strictly preferred to x_0 for certain by a person with utility function $u_1()$. That is,

$$\text{if } u_2(x_0) \leq \int u_2(x)\, dF(x)\text{, then } u_1(x_0) < \int u_1(x)\, dF(x)\,.$$

As usual, which definition is most useful will depend on the circumstances you are in. Practically, speaking, I think that number 3 is the least likely to come up, although it is useful in certain technical proofs. Note that it need not be the case that any two utility functions $u_2()$ and $u_1()$ are such that one is necessarily at least as risk averse as the other. In fact, the usual case is that you won't be able to rank them. However, most often we will be interested in finding out what happens to a particular person who becomes more risk averse, rather than actually comparing the risk aversion of two people.

In addition to knowing what happens when a person becomes more risk averse, we are also frequently interested in what happens to a person's risk aversion when her wealth changes. As I mentioned earlier, the natural assumption to make (since it corresponds with how people actually seem to behave) is that people becomes less risk averse as they become wealthier. In terms of the measures we have for risk aversion, we say that a person exhibits non-increasing absolute risk aversion whenever $r_A(x)$ is non-increasing in x. In MWG Proposition 6.C.3, there are some alternate definitions. Figuring them out would be a useful exercise. Of particular interest, I think, is part iii), which says that having non-increasing (or decreasing) absolute risk aversion means that as your wealth increases, the amount you are willing to pay to get rid of a risk decreases. What does this say about insurance? Basically, it means that the wealthy will be willing to pay less for insurance and will receive less benefit from being insured. Formalizing this, let z be a random variable with distribution F and a mean of 0. Thus z is the prize of a lottery with distribution F. Let c_x (the certainty equivalent) be defined as:

$$u(c_x) = \int u(x+z)\,dF(z).$$

If the utility function exhibits decreasing absolute risk aversion, then $x - c_x$ (corresponding to the premium the person is willing to pay to get rid of the uncertainty) will be decreasing in x.

As before, it is natural to think of people as exhibiting nonincreasing relative risk aversion. That is, they are more likely to accept a proportional gamble as their initial wealth increases. Although the concept of relative risk aversion is useful in a variety of contexts, we will primarily be concerned with absolute risk aversion. One reason for this that many of the techniques we develop for studying absolute risk aversion translate readily to the case of relative risk aversion.

6.2.3 A Note on Comparing Distributions: Stochastic Dominance

We aren't going to spend a lot of time talking about comparing different distributions in terms of their risk and return because these are concepts that involve slightly more knowledge of probability and will most likely be developed in the course of any applications you see that use them. However, I will briefly mention them.

Suppose we are interested in knowing whether one distribution offers higher returns than another. There is some ambiguity as to what this means. Does it mean higher average monetary return (i.e., the mean of F), or does it mean higher expected utility? In fact, when a consumer is risk averse, a distribution with a higher mean may offer a lower expected utility if it is riskier. For example, a sufficiently risk averse consumer will prefer $x = 1.9$ for sure to a 50-50 lottery over 1 and 3. This is true even though the mean of the lottery, 2, is higher than the mean of the sure thing, 1.9. Thus if we are concerned with figuring out which of two lotteries offers higher utility than another, simply comparing the means is not enough.

It turns out that the right concept to use when comparing the expected utility of two distributions is called **first-order stochastic dominance (FOSD).** Consider two distribution functions, $F()$ and $G()$. We say that $F()$ first-order stochastically dominates $G()$ if $F(x) \leq G(x)$ for all x. That is, $F()$ FOSD $G()$ if the graph of $F()$ lies everywhere below the graph of $G()$. What is the meaning of this? Recall that $F(y)$ gives the probability that the lottery offers a prize that is less than or equal to y. Thus if $F(x) \leq G(x)$ for all x, this means that for any prize, y, the probability that $G()$'s prize is less than or equal to y is greater than the probability that $F()$'s prize is less than or equal to y. And, if it is the case that $F()$ FOSD $G()$, it can be shown that any consumer with a strictly increasing utility function $u()$ will prefer $F()$ to $G()$. That is, as long as you prefer more money to less, you will prefer lottery $F()$ to $G()$.

Now, it's important to point out that most of the time you will not be able to rank distributions in terms of FOSD. It will need not be the case that either $F()$ FOSD $G()$ or $G()$ FOSD $F()$. In particular, the example from two paragraphs ago (1.9 for sure vs. 1 or 3 with equal probability) cannot be ranked. As in the case of ranking the risk aversion of two utility functions, the primary use of this concept is in figuring out (in theory) how a decision maker would react when the distribution of prizes "gets higher." FOSD is what we use to capture the idea of "gets higher." And, knowing an initial distribution $F()$, FOSD gives us a good guide to what it means for the distribution to get higher: The new distribution function must lay everywhere below the old one.

So, FOSD helps us formalize the idea of a distribution "getting higher." In many circumstances it is also useful to have a concept of "getting riskier." The concept we use for "getting riskier" is called **second-order stochastic dominance (SOSD).** One way to understand SOSD is in terms of mean preserving spreads. Let X be a random variable with distribution function $F()$. Now, for each value of x, add a new random variable z_x, where z_x has mean zero. Thus z_x can be thought of as a noise term, where the distribution of the noise depends on x but always has a mean of zero. Now consider the random variable $y = x + z_x$. Y will have the same mean as X, but it will be riskier because of all of the noise terms we have added in. And, we say that for any Y than has the same mean as X and can be derived from X by adding noise, Y is riskier than X. Thus, we say that X second-order stochastically dominates Y.

Let me make two comments at this point. First, as usual, it won't be the case that for any two random variables (or distributions), one must SOSD the other. In most cases, neither will SOSD.

Second, if you do have two distributions with the same mean, and one, say $F()$, SOSD the other, $G()$, then you can say that any risk averse decision maker will prefer $F()$ to $G()$. Intuitively, this is because $G()$ is just a noisy and therefore riskier version of $F()$, and risk-averse decision makers dislike risk.

OK. At this point let me apologize. Clearly I haven't said enough for you to understand FOSD and SOSD completely. But, I think that what we have done at this point is a good compromise. If you ever need to use these concepts, you'll know where to look. But, not everybody will have to use them, and using them properly involves knowing the terminology of probability theory, which not all of you know. So, at this point I think it's best just to put the definitions out there and leave you to learn more about them in the future if you ever have to.

6.3 Some Applications

6.3.1 Insurance

Consider a simple model of risk and insurance. A consumer has initial wealth w. With probability π, the consumer suffers damage of D. With probability $1-\pi$, no damage occurs, and the consumer's wealth remains w. Thus, in the absence of insurance, the consumer's final wealth is $w-D$ with probability π, and w with probability $1-\pi$.

Now, suppose that we allow the consumer to purchase insurance against the damage. Each unit of insurance costs q, and pays 1 dollar in the event of a loss. Let a be the number of units of insurance that the consumer purchases. In this case, the consumer's final wealth is $w-D+a-qa$ with probability π and $w-qa$ with probability $1-\pi$. Thus the benefit of insurance is that is repays a dollars of the loss when a loss occurs. The cost is that the consumer must give up qa dollars regardless of whether the loss occurs. Insurance amounts to transferring wealth from the state where no loss occurs to the state where a loss occurs.

The consumer's utility maximization problem can then be written as:

$$\max_{a} \pi u(w-D+(1-q)a)+(1-\pi)u(w-aq).$$

The first-order condition for an interior solution to this problem is:

$$\pi(1-q)u'(w-D+(1-q)a^*)-(1-\pi)qu'(w-a^*q)=0$$

Let's assume for the moment that the insurance is fairly priced. That means that the cost to the consumer of 1 dollar of insurance is just the expected cost of providing that coverage; in insurance jargon, this is called "actuarially fair" coverage. If the insurer must pay 1 dollar with probability π, then the fair price of insurance is $\pi * 1 = \pi$. Thus for the moment, let $q = \pi$, and the first-order condition becomes:

$$u'(w - D + (1-\pi)a^*) = u'(w - \pi a^*).$$

Now, if the consumer is strictly risk averse, then $u'()$ is strictly decreasing, which means that in order for the previous equation to hold, it must be that:

$$w - D + (1-\pi)a^* = w - \pi a^*.$$

This means that the consumer should equalize wealth in the two states of the world. Solving further,

$$D = a^*.$$

Thus, a utility-maximizing consumer will purchase insurance that covers the full extent of the risk - "full insurance" - if it is priced fairly.

What happens if the insurance is priced unfairly? That is, if $q > \pi$? In this case, the first-order condition becomes

$$\begin{aligned}
\pi(1-q)u'(w - D + (1-q)a^*) - (1-\pi)qu'(w - qa^*) &= 0 \text{ if } D > a^* > 0, \\
\pi(1-q)u'(w-D) - (1-\pi)qu'(w) &\leq 0 \text{ if } a^* = 0, \\
\pi(1-q)u'(w-qD) - (1-\pi)qu'(w-qD) &\geq 0 \text{ if } a^* = D.
\end{aligned}$$

Now, if we consider the case where $a^* = D$, we derive the optimality condition for purchasing full insurance Then,

$$u'(w-qD)(\pi(1-q) - (1-\pi)q) = u'(w-qD)(\pi - q) \geq 0.$$

Thus if the consumer is able to choose how much insurance she wants, she will never choose to fully insure when the price of insurance is actuarially unfair (since the above condition only holds if $\pi \geq q$, but by definition, unfair pricing means $q > \pi$).

There is another way to see that if insurance is priced fairly the consumer will want to fully insure. If it is actuarially fairly, the price of full insurance is πD. Thus if the consumer purchases full insurance, her final wealth is $w - \pi D$ with probability 1, and her expected utility is $u(w - \pi D)$. If she purchases no insurance, her expected utility is:

$$\pi u(w-D) + (1-\pi)u(w),$$

which, by risk aversion, is less than the utility of the expected outcome, $\pi(w-D)+(1-\pi)w = w-\pi D$. Thus

$$u(w-\pi D) > \pi u(w-D)+(1-\pi)u(w).$$

So, any risk-averse consumer, if offered the chance to buy full insurance at the actuarially fair rate, would choose to do so.

What is the largest amount of money that the consumer is willing to pay for full insurance, if the only other option is to remain without any insurance? This is found by looking at the consumer's certainty equivalent. Recall that the certainty equivalent of this risk, call it ce, solves the equation:

$$u(ce) = \pi u(w-D)+(1-\pi)u(w).$$

Thus ce represents the smallest sure amount of wealth that the consumer would prefer to the lottery. From this, we can compute the maximum price she would be willing to pay as:

$$ce = w - p_{\max}.$$

The last two results in this example may be a bit confusing, so let me summarize. First, if the consumer is able to choose how much insurance she wants, and insurance is priced fairly, she will choose full insurance. But, if the consumer is able to choose how much insurance she wants and insurance is priced unfairly, she will choose to purchase less than full insurance. However, if the consumer is given the choice only between full insurance or no insurance, she will be willing to pay up to $p_{\max} = w - ce$ for insurance.

6.3.2 Investing in a Risky Asset: The Portfolio Problem

Suppose the consumer has utility function $u()$ and initial wealth w. The consumer must decide how much of her wealth to invest in a riskless asset and how much to invest in a risky asset that pays 0 dollars with probability π and r dollars with probability $1-\pi$. Let x be the number of dollars she invests in the risky asset. Note, for future reference, that the expected value of a dollar invested in the risky asset is $r(1-\pi)$. The riskless asset yields no interest or dividend - its worth is simply \$1 per unit. The consumer's optimization problem is:

$$\max_x \pi u(w-x)+(1-\pi)u(w+(r-1)x).$$

The first-order condition for this problem is:

$$-\pi u'(w-x)+(1-\pi)(r-1)u'(w+(r-1)x) \begin{array}{ll} \leq 0 & \text{if } x^*=0 \\ =0 & \text{if } 0<x^*<w \\ \geq 0 & \text{if } x^*=w. \end{array}$$

The question we want to ask is when will it be the case that the consumer does not invest in the risky asset. That is, when will $x^*=0$? Substituting x^* into the first-order condition yields:

$$\begin{aligned} -\pi u'(w)+(1-\pi)(r-1)u'(w) &\leq 0 \\ u'(w)(-\pi+(1-\pi)(r-1)) &\leq 0. \end{aligned}$$

But, since $u'(w)>0$,[4] for this condition to hold, it must be that:

$$-\pi+(1-\pi)(r-1)\leq 0$$

or

$$(1-\pi)r\leq 1.$$

Thus the only time it is optimal for the consumer not to invest in the risky asset at all is when $(1-\pi)r\leq 1$. But, note that $(1-\pi)r$ is just the expected return on the risky asset and 1 is the return on the safe asset. Hence only when the expected return on the risky asset is less than the return on the safe asset will the consumer choose not to invest at all in the risky asset. Put another way, whenever the expected return on the risky asset is greater than the expected return on the safe asset (i.e., it is actuarially favorable), the consumer will choose to invest at least some of her wealth in the risky asset. In fact, you can also show that if the consumer chooses $x^*>0$, then $(1-\pi)r>1$.

Assuming that the consumer chooses to invest $0<x^*<w$ in the risky asset, which implies that $(1-\pi)r>1$, we can ask what happens to investment in the risky asset when the consumer's wealth increases. Let $x(w)$ solve the following identity:

$$-\pi u'(w-x(w))+(1-\pi)(r-1)u'(w+(r-1)x(w))=0$$

Differentiate with respect to $w:$

$$-\pi u''(w-x(w))\left(1-x'(w)\right)+(1-\pi)(r-1)u''(w+(r-1)x(w))\left(1+(r-1)x'(w)\right)=0$$

[4] $u'()>0$ by assumption, since more money is better than less money (i.e. the marginal utility of money is always positive).

and to keep things simple, let $a = w - x(w)$ and $b = w + (r-1)x(w)$. Solve for $x'(w)$

$$x'(w) = \frac{\pi u''(a) - (1-\pi)(r-1)u''(b)}{\pi u''(a) + (1-\pi)(r-1)^2 u''(b)}.$$

By concavity of $u()$, the denominator is negative. Hence the sign of $x'(w)$ is opposite the sign of the numerator. The numerator will be negative whenever:

$$\begin{aligned}
\pi u''(a) &< (1-\pi)(r-1)u''(b) \\
\frac{u''(a)}{u''(b)} &> \frac{(1-\pi)(r-1)}{\pi} \\
\frac{u''(a)}{u''(b)}\frac{u'(b)}{u'(a)} &> \frac{(1-\pi)(r-1)}{\pi}\frac{u'(b)}{u'(a)} = 1
\end{aligned}$$

where the inequality flip between the first and second lines follows from $u'' < 0$, and the equality in the last line follows from the first-order condition, $-\pi u'(a) + (1-\pi)(r-1)u'(b) = 0$. Thus this inequality holds whenever

$$\frac{u''(a)}{u'(a)} < \frac{u''(b)}{u'(b)},$$

(with another inequality flip because $\frac{u''}{u'} < 0$) or, multiplying by -1 (and flipping the inequality again),

$$r_A(w - x(w)) > r_A(w - x(w) + rx(w)).$$

So, the consumer having decreasing absolute risk aversion is sufficient for the numerator to be negative, and thus for $x'(w) > 0$. Thus the outcome of all of this is that whenever the consumer has decreasing absolute risk aversion, an increase in wealth will lead her to purchase more of the risky asset. Or, risky assets are normal goods for decision makers with decreasing absolute risk aversion. This means that the consumer will purchase more of the risky asset in absolute terms (total dollars spent), but not necessarily relative terms (percent of total wealth).

The algebra is a bit involved here, but I think it is a good illustration of the kinds of conclusions that can be drawn using expected utility theory. Notice that the conclusion is phrased, "If the decision maker exhibits decreasing absolute risk aversion...", which is a behavioral postulate, not an implication of the theory of choice under uncertainty. Nevertheless, we believe that people do exhibit decreasing absolute risk aversion, and so we are willing to make this assumption.

6.4 Ex Ante vs. Ex Post Risk Management

Consider the following simple model, which develops a "separation" result in the choice under uncertainty model. Suppose, as before, the consumer has initial wealth w, and there are two states of the world. With probability π, the consumer loses D dollars and has final wealth $w - D$. We call this the loss state. With probability $(1-\pi)$, no loss occurs, and the consumer has final wealth w. Suppose there are perfect insurance markets, meaning that insurance is priced fairly: insuring against a loss of \$1 that occurs with probability π costs \$$\pi$ (we showed this result earlier). Thus, by paying π dollars in both states, the consumer increases wealth in the loss state by 1 dollar. The net change is therefore that by giving up π dollars of consumption in the no-loss state, the consumer can gain $1-\pi$ dollars of consumption in the loss state. Notice, however, that any such transfer (for any level of insurance, $0 \leq a \leq D$) keeps expected wealth constant:

$$\begin{aligned} & \pi \left(w - D + (1-\pi) a\right) + (1-\pi)(w - \pi a) \\ = \quad & w - \pi D \end{aligned}$$

So, another way to think of the consumer's problem is choosing how to allocate consumption between the loss and no-loss states while keeping expected expenditure constant at $w - \pi D$. That is, another way to write the consumer's problem is:

$$\begin{aligned} & \max_{c_L, c_N} \pi u(c_L) + (1-\pi) u(c_N) \\ s.t. \quad : \quad & \pi c_L + (1-\pi) c_N \leq w - \pi D. \end{aligned}$$

This is where the separation result comes in. Notice that expected initial wealth only enters into this consumer's problem on the right-hand side (just as present value of lifetime income only entered into the right-hand side of the intertemporal budget constraint in the Fisher theorem, and profit from the farm only entered into the right-hand side of the farm's budget constraint in the agricultural household separation result).

So, suppose the consumer needs to choose between different vectors of state dependent income. That is, suppose $w = (w_L, w_N)$ and $w' = (w'_L, w'_N)$. How should the consumer choose between the two? The answer is that, if insurance is fairly priced, the consumer should choose the income vector with the highest expected value (i.e., w if $\pi w_L + (1-\pi) w_N > \pi w'_L + (1-\pi) w'_N$, w' if the inequality is reversed, either one if expected wealth is equal in the two states).

What happens if insurance markets are not perfect? Let's think about the situation where there are no insurance possibilities at all. In this case, the only way the consumer can reduce risk is to choose projects (i.e., state-contingent wealth vectors) that are, themselves, less risky. The result is that the consumer may choose a low-risk project that offers lower expected return, even though, if the consumer could access insurance markets, she would choose the risky project and reduce risk through insurance.

There are a number of applications of this idea. Consider, for example, a farmer who must decide whether to produce a crop that may not come in but earns a very high profit if it does, or a safe crop that always earns a modest return. It may be that if the farmer can access insurance markets, the optimal thing for him to do is to take the risky project and then reduce his risk by buying insurance. However, if insurance is not available, the farmer may choose the safer crop. The result is that the farmer consistently chooses the lower-value crop and earns lower profit in the long run than if insurance were available. A similar story can be told where the risky project is a business venture. In the absence of insurance, small businesses may choose safe strategies that ensure that the company stays in business but do significantly worse over the long run.

The distinction raised in the previous paragraph is one of ex ante vs. ex post risk reduction. Ex post risk reduction means buying insurance against risk. Ex ante risk reduction means engaging in activities up front that reduce risk (perhaps through choosing safer projects). In a very general way, if there are not good mechanisms for ex post risk reduction (either private or government insurance), then individuals will engage in excessive ex ante risk reduction. The result is that individuals are worse off than if there were good insurance markets, and society as a whole generates less output.