1 Lecture 1

1.1 The Standard Model of Choice Under Risk

Up until now, we have thought of the objects between which our consumers are choosing as being physical items - chairs, tables, apples, brandy etc. We pretty much know what will happen when we buy such things. However, we can also think of cases where the outcomes of the choices we make are uncertain - we don’t know exactly what will happen when we buy a particular object. Think of the following examples:

- You are deciding whether or not to buy a share in AIG
- You are deciding whether or not to put your student loan on black at the roulette table
- You are deciding whether or not to buy a house that straddles the San Andreas fault line

In each case, while you may understand exactly what it is that you are buying, (or choosing between), the outcomes, in terms of the things that you care, about are uncertain. Here we are going to think about how to model a consumer who is making such choices.

Economists tend to differentiate between two different types of ways in which we may not know for certain what will happen in the future: risk and uncertainty (sometimes called ambiguity). The difference between the two is that, in the former case, the probabilities of different outcomes are known, while in the latter case they are not. Sometimes the difference is illustrated by thinking
about the difference between a horse race and a roulette wheel. The idea being that, for a roulette wheel, we may not know which number is going to come up, but we know how likely each number is to come up. In contrast, in a horse race, we may not even know that: reasonable people may disagree about how likely it is for different horses to win. We will begin by discussing models of choice under risk, then move on to choice under uncertainty.

In order to be concrete, let’s think about a specific example. You are in a fairground, and come across a (very boring) game of chance. For an amount of money $x$, you can flip a coin. If it comes down as heads, you get $10. If it comes down tails, you get nothing (let’s assume that you get to choose the coin, so you are pretty sure that there is a 50% chance of a head and a 50% chance of tails). The question is, for what price $x$ would you choose to play the game. In other words, you have a choice between the following two options.

1. Not play the game and get nothing
2. Play the game, and get $-x$ for sure, plus a 50% chance of getting $10.

How would you make a decision like this? The earliest thinkers on the subject suggested the following strategy: Figure out the expected value (or average pay-out) of playing the game, and see if it is bigger than 0. If it is, then play the game, if not, then don’t.

So what is the expected value of the game? With a 50% chance you will get $10 - x$, while with a 50% chance you will get $-x$. Thus, the average payoff is going to be

$$0.5(10 - x) + 0.5(-x)$$

$$= 5 - x$$

Thus the value of the game is $5 - x$. In other words, following this strategy, you should play the game if the cost of playing is less than $5.

Does this sound sensible? People thought so until Daniel Bernoulli (the Dutch-Swiss maths superstar) came up with the following example:

**Example 1 (The St. Petersburg Paradox)** Imagine that the fairground guy offers you a different game. Now, you first of all flip the coin. If it comes down heads, then you get $2. If it comes
down tails, you flip again. If you get heads on that go, you get $4, otherwise you flip again. If it
comes down heads then you get $8, otherwise you flip again, and so on.

What is the expected value of this game? Well, there is a \( \frac{1}{2} \) probability that you will get heads
on the first trial, and so get $2. But there is a \( \frac{1}{2} \) chance that you will get tails and flip again. There
is then a \( \frac{1}{2} \) chance that you will get heads on that go, and so get $4. For that to happen, you would
have to get tails on the first go (probability \( \frac{1}{2} \)) and heads on the second go (probability \( \frac{1}{2} \)). Thus,
there is a \( \frac{1}{4} \) probability that you will get $4. Using the same logic, there is a \( \frac{1}{8} \) chance you will get
$8 and so on. The expected value of the game is therefore

\[
\frac{1}{2} \times 2 + \frac{1}{4} \times 4 + \frac{1}{8} \times 8 + \frac{1}{16} \times 16 + \ldots \\
= 1 + 1 + 1 + 1 + \ldots
= \infty
\]

The expected value of the game is \( \infty \), and so that is how much you should be willing to pay. In
other words, however much the fairground guy is prepared to charge you, you should be willing to
pay it.

Assuming that you are not one of the people that is prepared to pay \( \infty \) to play this game, what
has gone wrong? Bernoulli suggested one solution: Perhaps the difference in ‘happiness’ brought
about by getting extra money decreases as the amount of money you have increases. In other words
getting $1 extra if you only have $1 means a lot more than getting $1 extra if you have $1 million.
This is what we would (these days) call the decreasing marginal utility of wealth.

**Example 2** Say a pauper finds a magic lottery ticket, that has a 50% chance of $1 million and
a 50% chance of nothing. A rich person offers to buy the ticket off him for $499,999 for sure.
According to our ‘expected value’ method, the pauper should refuse the rich person’s offer!

Bernoulli argued that this is ridiculous. For the pauper, the difference in quality of life between
getting nothing and $499,999 is massive, while the difference between $499,999 and $1 million is
relatively small. Thus, by turning down the rich person’s offer, they are gaining relatively little (a
50% chance of getting $1 million rather than $499,999) and loosing an awful lot (a 50% chance of
getting 0 rather than $499,999). Moreover, Bernoulli argued, if this is the case, what we should be
maximizing is expected utility, rather than expected value. In other words, if \( u(x) \) is the utility of getting an amount \( x \) then, the pauper should choose to accept the rich guy’s offer if

\[
\frac{1}{2} u(\$1,000,000) + \frac{1}{2} u(\$0) < u(\$499,999)
\]

The idea is that the utility gap between 0 and \$499,999 is larger than the gap between \$499,999 and \$1,000,000. For example, it could be that

\[
\begin{align*}
    u(\$0) &= 0 \\
    u(\$499,999) &= 10 \\
    u(\$1,000,000) &= 16
\end{align*}
\]

If this were the case, then Bernoulli suggests that the pauper should accept the offer, as the expected utility of the lottery ticket is 8, while the expected utility of the rich man’s offer is 10. In fact he proposed that the utility of getting an amount \( x \) could be approximated by the function \( \ln(x) \) (note that this exhibits decreasing marginal utility). If this is right, then the most that you should pay for the St. Petersburg game is about \$60.

By and large, most of modern economics agrees with Bernoulli’s assessment of how choice under uncertainty should work. However, if you remember back to choice under certainty, we in general don’t like the idea of utility functions coming out of nowhere. When we were talking about choice under certainty, we were very careful to ask the question: what has to be true about a person’s preferences for us to be able to represent them with a utility function? The answer was that preferences had to be complete and transitive. Here we want to be equally careful: Under what circumstances can people’s preferences be represented by an expected utility function? Do they just have to be complete and transitive? Or do we need something more?

To answer this question, we need to be a little bit more formal about what we are doing here. Before we were thinking about people’s preferences over objects (chairs, lamps, Miley Cyrus CD’s etc.). Now we are thinking about peoples preferences over what we will call lotteries. What do we mean by a lottery?. In order to fix ideas, lets think of games of chance in which you can win one of four prizes: \( a, b, c \) and \( d \). A lottery is just a list of four numbers indicating the probability of winning each of the prizes. so for example
is a lottery. This gives a 15% chance that you win a, a 35% chance you win b, a 50% chance that you win the c and a 0% chance you win d. Another lottery is

\[
p = \begin{pmatrix}
    p_a \\
p_b \\
p_c \\
p_d
\end{pmatrix}
\]

The only thing that has to be true for such a list to be a lottery is (a) each of the numbers has to be greater than or equal to zero (you can’t have a negative probability) and (b) they have to add to 1 (the probability of winning one of the prizes is 100%). In general, we will write

Note that here we are making an implicit assumption that we can see the lotteries that people have preferences are over - i.e. when we see people choosing between objects, we are happy to identify the probability distribution that they are choosing.

We will think about a consumer that has complete and transitive preference over such lotteries. In other words,

1. **Completeness**: For any two lotteries \( p \) and \( q \) one and exactly one of the following is true: \( p \succ q \) or \( q \succ p \) or \( p \sim q \)

2. **Transitivity**: For any three lotteries \( p, q \) and \( r \)

   (a) If \( p \succ q \) and \( q \succ r \) then \( p \succ r \)
(b) If $p \sim q$ and $q \sim r$ then $p \sim r$

Is this enough to guarantee an expected utility representation? In other words, can we find utility numbers for each of the prizes $u(a)$, $u(b)$ $u(c)$ and $u(d)$ such that

$$p > q$$

if and only if

$$p_a u(a) + p_b u(b) + p_c u(c) + p_d u(d) > q_a u(a) + q_b u(b) + q_c u(c) + q_d u(d)$$

The answer is no: we need two more axioms. The first (and most important) is called the independence axiom. To understand this axiom, think of the following question:

**Question:** Think of two different lotteries, $p$ and $q$. Just for concreteness, let’s say that $p$ is a 25% chance of winning the apple and a 75% chance of winning the banana, while $q$ is a 75% chance of winning the apple and a 25% chance of winning the banana. Say you prefer the lottery $p$ to the lottery $q$. Now I offer you the following choice between option 1 and 2

1. I flip a coin. If it comes up heads, then you get $p$. Otherwise you get the lottery that gives you the cockatiel for sure
2. I flip a coin. If it comes up heads, you get $q$. Otherwise you get the lottery that gives you the cockatiel for sure

Which do you prefer?

The independence axiom basically states that if you prefer $p$ to $q$, then you have to prefer option 1 to option 2. This seems intuitively plausible. After all, in the choice between 1 and 2, then if the coin comes up tails, then you get the same thing in both cases. If it comes up heads then for 1 your get $p$ and for 2 you get $q$. If you prefer $p$ to $q$, then it seems natural that you should prefer 1 to 2. In fact, the independence axiom says slightly more than this:

**Axiom 1 (The Independence Axiom)** Say a consumer prefers lottery $p$ to lottery $q$. Then, for any other lottery $r$ and number $0 < \alpha \leq 1$ they must prefer

$$\alpha p + (1 - \alpha)r$$
The independence axiom is both beautiful and intuitive. Of course, we can find circumstances in which it doesn’t work well (which we will discuss in the next lecture), but for now the important thing is that the independence axiom is necessary for an expected utility representation (you will prove this for homework).

The second axiom that we need is called the Archimedean axiom. This states the following:

**Axiom 2 (The Archimedean Axiom)** For all lotteries \( p, q \) and \( r \) such that \( p \succ q \succ r \), there must exist an \( a \) and \( b \) in \((0, 1)\) such that

\[
apq + (1-a)r
\]

This essentially tells us that there is no lottery which is infinitely good or infinitely bad. Again, you will prove that expected utility implies the Archimedean axiom for homework.

Perhaps more interestingly, the Independence and Archimedean axioms (along with completeness and transitivity) are also sufficient for an expected utility representation - an idea captured in the following theorem:

Let \( X \) be a finite set of prizes, \( \Delta(X) \) be the set of lotteries on \( X \). Let \( \succeq \) be a binary relation on \( \Delta(X) \). Then \( \succeq \) is complete, reflexive, transitive and satisfies the independence and Archimedean axioms if and only if there exists a \( u : X \to \mathbb{R} \) such that, for any \( p, q \in \Delta(X) \),

\[
\text{if and only if } \sum_{x \in X} p(x)u(x) \geq \sum_{x \in X} q(x)u(x)
\]

A rigorous proof of this lies beyond the scope of this course, but we can sketch what happens (see the relevant chapter in notes on the theory of choice for details).

1. Show that, if \( \delta_x \succ \delta_y \) (i.e. the lottery that gives prize \( x \) for sure is preferred to the lottery that gives prize \( y \) for sure) and \( a > b \), then

\[
a\delta_x + (1-a)\delta_y \succ b\delta_x + (1-b)\delta_y
\]

(this is intuitively obvious, but needs to be proved from the independence axiom)
2. Find the best prize - in other words the prize such that getting that prize for sure is preferred to all other lotteries. Give that prize utility 1 (for convenience, let's say that \( a \) is the best prize).

3. Find the worst prize - in other words the prize such that all lotteries are preferred to getting that prize for sure. Give that prize utility 0 (for convenience, let's say that \( d \) is the worse prize) (it is not, in fact, completely obvious that such best and worst prizes exist, but they do).

4. For other prizes (e.g. \( b \)), find the probability \( \lambda \) such that the consumer is indifferent between getting apples with probability \( \lambda \) and nothing with probability \( (1 - \lambda) \), and bananas for sure.

Let \( u(b) = \lambda \). i.e.

\[
\begin{pmatrix}
0 \\
1 \\
0 \\
0 \\
\end{pmatrix} \sim u(b) \begin{pmatrix}
1 \\
0 \\
0 \\
0 \\
\end{pmatrix} + (1 - u(b)) \begin{pmatrix}
0 \\
0 \\
0 \\
1 \\
\end{pmatrix}
\]

(for us to know such a \( \lambda \) exists requires the Archimedean axiom)

5. Do the same for \( c \), so

\[
\begin{pmatrix}
0 \\
0 \\
1 \\
0 \\
\end{pmatrix} \sim u(c) \begin{pmatrix}
1 \\
0 \\
0 \\
0 \\
\end{pmatrix} + (1 - u(c)) \begin{pmatrix}
0 \\
0 \\
0 \\
1 \\
\end{pmatrix}
\]

6. Now, we need to show that

\[
p \succ q
\]

if and only if

\[
p_a u(a) + p_b u(b) + p_c u(c) + p_d u(d) \\
> q_a u(a) + q_b u(b) + q_c u(c) + q_d u(d)
\]

To see why this is true, let's think of a simple example: again let's say that \( p \) is a 25% chance of winning \( b \) and a 75% chance of winning \( c \) while \( q \) is a 75% chance of winning \( b \) and a 25%
chance of winning the $c$. Now, note that

$$p = \begin{pmatrix}
0 \\
0.25 \\
0.75 \\
0
\end{pmatrix}
\begin{pmatrix}
0 \\
1 \\
0 \\
0
\end{pmatrix} + 0.75 \begin{pmatrix}
0 \\
0 \\
1 \\
0
\end{pmatrix}

But we know that

$$\begin{pmatrix}
0 \\
1 \\
0 \\
0
\end{pmatrix} \sim u(b) \begin{pmatrix}
1 \\
0 \\
0 \\
0
\end{pmatrix} + (1 - u(b)) \begin{pmatrix}
0 \\
0 \\
0 \\
1
\end{pmatrix}

and

$$\begin{pmatrix}
0 \\
0 \\
1 \\
0
\end{pmatrix} \sim u(c) \begin{pmatrix}
1 \\
0 \\
0 \\
0
\end{pmatrix} + (1 - u(c)) \begin{pmatrix}
0 \\
0 \\
0 \\
1
\end{pmatrix}$$
and so, by the independence axiom

\[ p \sim 0.25 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} u(b) + (1 - u(b)) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} + 0.75 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} u(c) + (1 - u(c)) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \]

\[ = (0.25u(b) + 0.75u(c)) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + (1 - 0.25u(b) - 0.75u(c)) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \]

or \( p \) is indifferent to a mixture between the best and the worst lotteries, with the weight on the best lottery given by \( 0.25u(b) + 0.75u(c) \). But this is the expected utility of \( p \). Similarly, \( q \) is indifferent to such a mixture, with the weight on the best lottery equal to the expected utility of \( q \). In other words

\[ q \sim (0.75u(b) + 0.25u(c)) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + (1 - 0.75u(b) + 0.25u(c)) \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \]

Thus, if the expected utility of \( p \) is higher than the expected utility of \( q \), this tells us that the lottery to which \( p \) is indifferent puts a higher weight on the best prize than does the lottery to which \( q \) is indifferent. Thus (by step 1), \( p \) must be preferred to \( q \).

One interesting thing to note is that, now we are dealing with lotteries, utility numbers mean more than they did in the case where we were dealing with non-risky objects. Previously, we said that utility was unique only up to a strictly monotone transformation - i.e. any utility function that kept the order the same would work. This is not the case here. For example, lets say that the
utility function

\[
\begin{align*}
    u(a) & = 1 \\
    u(b) & = 2 \\
    u(c) & = 3 \\
    u(d) & = 4
\end{align*}
\]

represented preferences. Now consider the following (strictly monotone transformation)

\[
\begin{align*}
    v(a) & = 1 \\
    v(b) & = 4 \\
    v(c) & = 9 \\
    v(d) & = 16
\end{align*}
\]

Will these represent the same preferences? The answer is clearly not: In the first case, \( \frac{1}{2}u(a) + \frac{1}{2}u(c) = u(b) \), and so \( \frac{1}{2}a + \frac{1}{2}c \sim b \). However, \( \frac{1}{2}v(a) + \frac{1}{2}v(c) > v(b) \), indicating the fact that \( \frac{1}{2}a + \frac{1}{2}c \succ b \), so clearly these cannot be representing the same preferences.

This means that these utility numbers (sometimes called von Neumann Morgensten utility numbers, after the pair that first proved this theorem) contain more information than just the order. In fact, you can see this from the ‘proof’ of the theorem: the utility number of some prize \( b \) represents the probability that must be placed on the best prize to make a lottery between the best prize and the worst prize indifferent to \( b \). This should suggest that we cannot just arbitrarily change these numbers. The question remains: how much can we change the utility numbers and still represent the same preferences (equivalently, how seriously should we take vNM utility numbers).

The answer is that these numbers are unique up to a positive affine transformation:

**Theorem 1** Let \( \succeq \) be a set of preferences on \( \Delta(X) \) and \( u : X \to \mathbb{R} \) form an expected utility representation of \( \succeq \). Then \( v : X \to \mathbb{R} \) also forms an expected utility representation of \( \succeq \) if and only if

\[
v(x) = ax + b \quad \forall x \in X
\]

for some \( a \in \mathbb{R}_{++}, b \in \mathbb{R} \)

**Proof.** Homework \( \blacksquare \)
1.1.1 Risk Aversion

How does our expected utility representation relate back to the St. Petersburg paradox and the pauper? Well, what we were basically saying there is that it seemed sensible for the pauper to turn down a ‘fair gamble’: Even though the lottery ticket gave a higher average payoff than $499,999 for sure, they would still prefer the sure thing. We call such people risk averse.

Definition 1 A person is (strictly) risk averse if they always (strictly) prefer an amount $x$ for sure to a lottery that has an expected value of $x$.

Is there anything in what we have done so far that says that people have to be risk averse? The answer is no. Consider someone who had a utility function over money described by the function $u(x) = x^2$. For them, the utility of a 50/50 gamble between 10 and 0 is $\frac{1}{2} \cdot 10^2 + \frac{1}{2} \cdot 0^2 = 50 > 25 = 5^2$. Thus such a person would prefer the 50/50 gamble to its expected value for sure. Thus, if we want risk aversion, we have to impose it as an additional condition.

One question we might want to ask is: what utility functions lead to risk aversion. We have a clue in the examples that we have already used: we showed that a subject with log utility is risk averse, while one with a squared utility function is risk loving. The former is an example of a concave utility function, while the latter is an example of a convex utility function. And in fact, this is the defining characteristic of a risk averse agent.

Proposition 1 A decision maker whose preferences can be represented by an expected utility function over money is (strictly) risk averse if and only if their utility function is (strictly) concave.

Proof. Homework ■

One natural question that we might want to ask is: how does risk aversion change with wealth? For example, is someone with a wealth level of $1000 more likely to accept a 50/50 gamble between -$100 and $110 than someone who has a wealth level of $100000? We characterize this facet of behavior in the following way:
**Definition 2** We define someone as exhibiting decreasing absolute risk aversion if

\[ \delta_w + q \geq \delta_{w'} + \delta_z \]

\[ \Rightarrow \delta_{w'} + q \geq \delta_{w'} + \delta_z \]

for \( \forall w, w', z \in \mathbb{R}, w' > w \) and \( q \in P_z \), where we interpret \( \delta_w + q \) as the lottery that gives \( w \) plus whatever the outcome of \( q \) is.

Note that we are now interpreting the gamble \( q \) as adding and subtracting from the wealth level \( w \), rather than a gamble over final wealth levels as we have been doing up to now.

So we define someone as having decreasing absolute risk aversion if, for every gamble that they are prepared to take over a sure thing at wealth level \( w \), they will also take the gamble over the sure thing at any higher wealth level \( w' \). (Do you think that this is a sensible property?)

We might be interested in what type of utility function leads to decreasing absolute risk aversion. We can answer this question (if we assume that \( u \) is twice continuously differentiable) by defining the Arrow-Pratt measure of risk aversion.

**Definition 3** For an expected utility maximizer with \( u : Z \rightarrow \mathbb{R} \), the Arrow-Pratt measure of risk aversion is defined by

\[ \lambda(z) = \frac{-u''(z)}{u'(z)} \]

What does this mean? Well, first note, that, if the DM is risk averse and prefers more money to less, then \( \lambda(z) \) is positive (as the second derivative of a concave function is negative). Notice also, that, in some sense, if the utility function is ‘more’ concave then \( \lambda(z) \) will increase. And in fact this intuition is correct in the sense that in many instances we can use the Arrow-Pratt measure to compare levels of risk aversion. For example:

**Proposition 2** An expected utility decision maker exhibits decreasing absolute risk aversion if and only if \( \lambda(z) \) is nondecreasing in \( Z \).

Another use of this measure is to compare the risk aversion of different decision makers:
Definition 4  Decision maker A is at least as risk averse as decision maker B if, for every lottery $p$ and amount $z$, if decision maker A chooses $p$ over $\delta_z$, then B also chooses $p$ over $\delta_z$.

Expected utility decision makers who can be ranked in this way will also be ranked according to their Arrow Pratt measures of risk aversion:

Proposition 3  If decision maker A is at least as risk averse as decision maker B, then for every $z$ the Arrow Pratt measure of risk aversion of A will be higher than that of B.