Choice Under Risk

Mark Dean

Lecture Notes for Fall 2015 PhD Class ECON 2090 - Brown University

1 Introduction

We are now going to enrich our model of decision making by endowing our objects of choice with some extra structure. Up until now, we have thought about objects as just objects - they had no defining characteristics whatever. We are now going to think of these objects as representing some sort of uncertain prospect: when we choose one of these prospects, we are not sure what the actual outcome will be. Note that, intuitively, this requires a bit of a shift in our thinking. Now the things that we are choosing over is not directly the thing that we care about. Instead, the objects we choose over will lead to the things we do care about it with some uncertainty. Note that it can be quite subtle where this uncertainty may lie. For example, if we buy a lawnmower, we may know very well that it is a lawnmower, but not know if it is a good or a bad lawnmower. So what we really care about is not owning a lawnmower, but whether we have a good or bad lawnmower. Or maybe what we really care about is getting our grass cut. Taking this a bit further, it may be that we buy a tuna sandwich for dinner, and we know exactly what is in that sandwich, but we are unsure whether we really fancy tuna for dinner. Here, we might want to think about what we really care about is our own internal state (for example whether we are satisfied or not) and the uncertainty is over the effect a product will have on that state. Up to now, we have not had to worry about these subtleties. However, as will become clear, we will now be treating some of these relationships as observable, so we do have to worry about it.

There are broadly speaking two ways in which economists model uncertainty. We will term these 'risk' and 'ambiguity' (ambiguity is more often called uncertainty, but if we do that we lose an umbrella term for both together). Risk is also sometimes termed objective uncertainty, while ambiguity is called subjective uncertainty. The difference between the two is often demonstrated by thinking about the difference between a roulette wheel and a horse race. When we bet on a roulette wheel, we do not know what the outcome will be, but there is a sense in which we can all agree on what the *probability* of each outcome is. In contrast, with a horse race, not only do we know what the outcome is, we may also disagree (or not know) the probability of each outcome. The former is considered a case of risk, while the latter a case of ambiguity. While many decision theoretic models are written in these terms, I am not clear that the dichotomy really makes sense. More usefully we might want to think of a continuum. Think of making a bet about the color of a ball drawn from a bag filled with 40 red and black balls (a device that we will come back to). First, imagine that I tell you nothing about the number of each color of balls in the bag. This is usually considered a situation of ambiguity. Now imagine that I tell you that there are 30 red and 10 black balls in the bag. This is generally considered a situation of risk. What about cases where you have some information about the number of balls in the bag.? It seems sensible to treat risk and ambiguity as extreme points on this distribution rather than completely dichotomous cases.

This chapter is going to involve representation theorems that relate *preference relations* to *utility representations.* We are not going to talk about choice very much. However, this should not concern you - the last chapter described how we can get from choices to preferences. Therefore, in order to turn these condition on preferences to conditions on choices we need to demand the OWC axiom, then place the conditions directly on preferences derived from binary comparisons. Of more concern is the extent to which we need to observe a complete set of choices in order to test these axioms. This we will come back to later.

We are going to begin by thinking about choice under risk (i.e. cases where an uncertain object is characterized by known probabilities). The canonical model for this situation is, of course, that of expected utility maximization.

Note that the first two sections here are stolen pretty much verbatim from Notes on the Theory of Choice Chapter 5, so you can always just go and read that if you prefer.

2 Choice Under Risk: Finite Prize Space

In order to begin our discussion of choice under risk we need to begin by defining what our observables are. The main change is that the objects of choice are now going to be *lotteries.* A lottery is an object that tells you with what probability you are going to get one of a set of different outcomes. Thus, in order to define what I mean by a lottery, I need to start by describing the *prize space,* or set of things that you could potentially win. Here, we are going to start by being agnostic. The prize space is just going to be a set Z containing objects that have no defining characteristics. To begin with, we are going to assume that Z is finite. Things become technically rather more complicated if we drop this assumption.

Now that we have our prize space, I can define the set of lotteries:

Definition 1 *For a finite prize space Z, the set of lotteries is defined as*

$$P = Z \rightarrow [0, 1]$$

such that $\sum_{z \in Z} p(z) = 1$

In other words, it is the set of probability distributions over Z. Note that if we define addition and scalar multiplication in the standard way, then P is convex. (we will abuse notation a little by writing $\alpha p + (1 - \alpha)q$ to refer to the lottery that gives probability $\alpha p(z) + (1 - \alpha)q(z)$ to every z)

We are going to take as our observables as a preference relation $\succeq$ on P. Note that this is quite an ambitious thing to do, even assuming we can observe preferences. It means that we are conceptualizing the prize space, and the probability distribution that our DM has over this prize space for every option. This may be easy to do in the casino. Outside of that environment it may be considerably more difficult.

Our basic model of decision making is that people make choices to maximize expected utility:

Definition 2 *A preference relation $\succeq$ on a finite prize space has an expected utility representation if there exists $u : Z \rightarrow \mathbb{R}$ such that*

$$\begin{aligned} p &\succeq q \text{ if and only if} \\ \sum_{z \in Z} p(z)u(z) &\geq \sum_{z \in Z} q(z)u(z) \end{aligned}$$

We will come back to discuss whether we like this model in a little while. For now, one question to ask ourselves is, why don't we think that this is a sensible requirement for choices over bundles of goods?

So, what conditions do we need on $\succeq$ in order to guarantee an expected utility representation. Well, first, an expected utility representation is also a utility representation, so we know our first conditions

Axiom 1 $\succeq$ *is a complete preference relation*

Is this enough to guarantee that $\succeq$ has a *utility* representation? No! The space of lotteries is uncountable, even in the case of a finite prize space, so we need something else. In the last chapter we discussed how continuity of preferences is necessary and sufficient for continuous preference relation. As an expected utility representation is continuous, we can safely demand this condition as well

Axiom 2 $\succeq$ *is continuous.*

Great, so now we know that there is a continuous function $U : P \to \mathbb{R}$ such that $U(p) \geq U(q)$ iff $p \succeq q$. Is this enough to guarantee an expected utility representation? In other words, is it the case that every preference relation that has a continuous utility representation has an expected expected utility representation? Hopefully you know that the answer to this question is no (though you may want to try coming up with an example of a complete, continuous preference relation that does not have an expected utility representation, just to make sure that you understand what is going on.)

In arguably the most famous decision theoretic work of all time, Von Neumann and Morgnesten provided the necessary condition. Essentially they noticed that, to be represented this way, preferences have to have an affine structure. The resulting axiom is elegant, powerful and intuitive

Axiom 3 (Independence) *For any* $p, q, r \in P$, $p \succeq q$ *if and only if* $\alpha p + (1-\alpha)r \succeq \alpha q + (1-\alpha)r$ *for* $\forall\ \alpha \in (0, 1]$

This axiom, also known as the substitution axiom, is the real implication of expected utility theory. It says that the preferences between p and q survive mixing with some other lottery r. This axiom usually strikes people as pretty plausible when they first read it. It has also been justified on normative grounds, usually in the following way: think of $\alpha p + (1-\alpha)r$ as a compound lottery. You are going to flip a (weighted) coin, and depending on how it comes down you will get to play either lottery p or lottery r. $\alpha q + (1-\alpha)r$ is exactly the same except with p replaced with q. So, if the coin comes down 'heads' then in both cases you get r. The only difference is if the coin comes down 'tails', in which case you will either get p or q. Thus your preferences should really depend on your preferences between p and q.

Keen as we are on intuitive plausibility as a concept, what we really want to know is whether or not it holds empirically. We will come back to that later on, but for now, you should think about why you might not expect this axiom to hold if (for example) we were dealing with choice between bundles of goods

It turns out that these three properties are both necessary and sufficient for an expected utility representation. In order to prove this, it is going to be hand to prove 4 intermediate lemma

Lemma 1 *For $p \succ q$ and $0 \leq a < b \leq 1$*

$$bp + (1-b)q \succ ap + (1-a)q$$

Proof. *The case of $a = 0$ is direct from the independence axiom, so assume that $a \neq 0$. Let $r = bp + (1-b)q$. Note that $r \succ q$, and*

$$\begin{aligned} & r \\ = & \left(1 - \frac{a}{b}\right) r + \frac{a}{b} r \\ \succ & \left(1 - \frac{a}{b}\right) q + \frac{a}{b} r \\ = & \left(1 - \frac{a}{b}\right) q + \frac{a}{b} \left(bp + (1-b)q\right) \\ = & ap + (1-a)q \end{aligned}$$

■

Lemma 2 *For any p, q such that $p \succ q$ and r such that $p \succeq r \succeq q$, there exists an α such that*

$$ap + (1-\alpha)q \sim r$$

This α is unique

Proof. *Let*

$$\alpha = \sup \{a \in [0,1] | r \succeq ap + (1-\alpha)q\}$$

Our claim is that $ap + (1-\alpha)q \sim r$. To see this, assume not, then either $ap + (1-\alpha)q \succ r$, in which case, by continuity and lemma 1 there exists an ε such that

$$(a-\varepsilon)p + (1-\alpha+\varepsilon)q \succ r$$

and it must be the case that $bp + (1-b)q \succ r$ for all $b > (a-\varepsilon)$. Therefore α is not the sup of the defined set. Alternatively, it could be that $ap + (1-\alpha)q \prec r$, but in this case we know that there is an ε such that

$$(a+\varepsilon)\, p + (1-\alpha-\varepsilon)q \prec r$$

so α would not be an upper bound

Uniqueness comes directly from lemma 1 ■

Lemma 3 *$p \sim q$ implies that, for all $\alpha \in [0,1]$, $\alpha p + (1-\alpha)r \sim \alpha q + (1-\alpha)r$*

Exercise 1 *Exercise*

Proof. Let $\delta(z)$ be the lottery that gives prize z for sure for any $z \in Z$. There exists a z^* z_* such that

$$\begin{aligned} \delta(z^*) &\succeq p \;\forall\; p \in P \\ \delta(z_*) &\preceq p \;\forall\; p \in P \end{aligned}$$

Lemma 4 *Exercise*

■

We are now ready to prove our theorem

Theorem 2 *A preference relation on P admits an expected utility representation if and only if it is complete, continuous, and satisfies the independence axiom*

Proof. *The proof that the representation implies the axioms is trivial, so we will ignore it. We will also ignore the case where $p \sim q$ for all $p, q \in P$, as this is trivial. Therefore we know that $\delta(z^*) \succ \delta(z_*)$. Define the function $f : P \to [0,1]$ as*

$$f(p) = \{\alpha \in [0,1] | p \sim \alpha\delta(z^*) + (1-\alpha)\delta(z_*)\}$$

which we know from the above lemma is well defined. Moreover

$$\begin{aligned} p &\succeq q \\ &\Leftrightarrow f(p)\delta(z^*) + (1 - f(p))\delta(z_*) \\ &\geq f(q)\delta(z^*) + (1 - f(q))\delta(z_*) \\ &\Longleftrightarrow f(p) \geq f(q) \end{aligned}$$

\ ■

Proof. *so f is a utility representation of $\succeq$,*

Note also that

$$\begin{aligned} & \alpha p + (1-\alpha)q \\ \sim\ & \alpha\left(f(p)\delta(z^*) + (1-f(p))\delta(z_*)\right) \\ & +(1-\alpha)\left(f(q)\delta(z^*) + (1-f(q))\delta(z_*)\right) \\ =\ & \left(\alpha f(p) + (1-\alpha)f(q)\right)\delta(z^*) \\ & +\left(1 - \left(\alpha f(p) + (1-\alpha)f(q)\right)\right)\delta(z_*) \end{aligned}$$

so

$$f(\alpha p + (1-\alpha)q) = \alpha f(p) + (1-\alpha)f(q)$$

so f is affine.

Now we are ready to define our utility function. Let $u : Z \to \mathbb{R}$ be defined as

$$u(z) = f(\delta(z))$$

All we need to show is that $f(p) = \sum p(z)u(p)$. As we are in a finite prize space, we can do this by inducting on the support of the lottery p. It is trivially true if the support is 1. Assume it *to be true for lotteries of support $n-1$. Let p be a lottery of support n, and z' be in the support of p. Define the lottery q as*

$$\begin{aligned} q(z) &= 0 \text{ if } z = z' \\ &= \frac{p(z)}{1-p(z')} \text{ otherwise.} \end{aligned}$$

q has support of $n-1$, and therefore it is true that $f(q) = \sum q(z)u(z)$. Now note that

$$p = p(z')\delta(z') + (1 - p(z'))q$$

so

$$\begin{aligned} f(p) &= f(p(z')\delta(z') + (1-p(z'))q) \\ &= p(z')f(\delta(z')) + (1 - p(z')f(q) \\ &= p(z')u(z') + (1-p(z'))\sum q(z)u(z) \\ &= p(z')u(z') + (1-p(z'))\sum_{z\neq z'} \frac{p(z)}{1-p(z')}u(z) \\ &= \sum p(z)u(z) \end{aligned}$$

■

As with our standard ordinal utility theory, the next result we would like is to say something about uniqueness. In the case of ordinal utility, uniqueness was up to any strictly positive transformation. In this case, this is not true: any strictly positive transformation of a set of utility numbers u will not give us another expected utility representation. Rather, in this case we have uniqueness up to a positive *affine* transformation

Theorem 3 *Let $u(z)$ form an expected utility representation for a set of preferences $\succeq$ on P. Then $v(z)$ also forms an expected utility representation of $\succeq$ if and only if*

$$v(z) = \alpha u(z) + \beta \ \forall \ z \in Z$$

where $\alpha \in \mathbb{R}_{++}$ and $\beta \in \mathbb{R}$

You will prove this fact for homework, but note that this means that we have to take these utility numbers a little more seriously than we did the utility numbers from ordinal utility. Specifically, it means that the ratio of the gaps between utilities are invariant to the utility numbers we choose i.e. if v and u form utility representations of $\succeq$, then, for any $x, y, z, w \in Z$

$$\begin{aligned}
\frac{u(x)-u(y)}{u(z)-u(w)} &= \frac{\alpha u(x)-\beta-(\alpha u(y)-\beta)}{\alpha u(z)-\beta-(\alpha u(w)-\beta)} \\
&= \frac{v(x)-v(y)}{v(z)-v(w)}
\end{aligned}$$

So in this case there is a meaningful sense in which the utility gap between x and y is bigger (or smaller) than that between z and w: this gap will be invariant to the representation we choose. You should try and figure out what this gap 'means'.

3 General Prize Spaces

We are now going to somewhat briefly consider more general prize spaces (i.e. where the set of prizes in non-finite) Broadly speaking, the message is going to be the same, but we are going to require some technical bells and whistles (which we will touch on, rather than go nuts about)

We can think of the proof of the finite prize space as having two stages

1. Showing that $\succeq$ has an affine representation (i.e. there is a function $F : P \rightarrow \mathbb{R}$ such that $p \succeq q$ if and only if $F(p) \geq F(q)$, and such that $F(\lambda p + (1 - \lambda)q) = \lambda F(p) + (1 - \lambda)F(q))$ $\forall$ $\lambda \in [0, 1]$

2. Showing that an affine function on P can be rewritten as $F(p) = \sum_{z \in Z} u(z)p(z)$ (or $= \int_{z \in Z} u(z)f_p(z)d(z))$ for some $u : Z \rightarrow \mathbb{R}$

Proving an expected utility representation for more general prize spaces also requires these two stages. We will have a look at each of these stages in turn

3.1 Representing Preferences with Affine Functions[1]

So our first task is to understand whether or not our current axioms are enough to guarantee an affine representation of preferences on more general probability spaces. We are going to take two approaches to this.

First,. we are going to invoke what is known as the 'mixture space' theorem. To do this, we are going to have to understand what a mixture space is. Basically a mixture space is a set of objects, and a rule for combining these objects together. The reason that we want to deal with mixture spaces is (if you think about) the independence axiom is really a condition about what happens when we mix lotteries together. Therefore, if we can prove something generically for mixture spaces, then we can apply this to all sorts of different lottey spaces, as long as they obey the mixture space axioms.

So let's start by defining a mixture space:

[1] For more details of most of the stuff in this section, see Kreps chapter 5.

Definition 3 *A mixture space is a set of objects* Π *and a family of functions* $f_\alpha : \Pi \times \Pi \to \Pi$ *for* $\alpha \in [0, 1]$ *such that*

1. $f_0(\pi, \rho) = \pi$
2. $f_a(\pi, \rho) = f_{1-\alpha}(\pi, \rho)$
3. $f_\alpha(f_\beta(\pi, \rho), \rho) = f_{\alpha\beta}(\pi, \rho)$

The most obvious mixture space is defined by $\Pi = \mathbb{R}$ and $f_\alpha(a, b) = \alpha a + (1 - \alpha)b$. In fact, you can see the definition of the mixture space as attempting to identify other spaces that have the same stucture as this type of mixing on the real line.

Here are some other examples of mixture spaces

Example 1 *Let* Z *be a countable set, and* Π *be the set of functions* $p : Z \to \mathbb{R}$ *such that* $\sum_{z \in Z} p(z) = 1$, $p(z) \geq 0 \ \forall \ z \in Z$ *(i.e the set of all probability distributions on* Z*). Define mixing point by point: i.e.* $f_\alpha(p, q) = r$ *where* $r : Z \to \mathbb{R}$ *defined by* $r(z) = \alpha p(z) + (1 - \alpha)q(z)$*. Then* Π*,* $\{f_\alpha\}$ *is a mixture space (what do you need to check to ensure that this is true?)*

Example 2 *Let* Z *be an arbitrary set, and* Π *be the set of simple probability distributions on* Z *(i.e. the set of probability distributions with finite support). Define* $f_\alpha(p, q) = r$ *as* $r : supp(p) \cup supp(q) \to \mathbb{R}$ *with* $r(z) = \alpha p(z) + (1 - \alpha)q(z)$ *(where we assume that* $p(z) = 0$ *for* z *not in its support and the same for* q*). Again, this is a mixture space*

Example 3 *Let* Z *be a metric space and* Π *be the set of borel-measurable probability distributions on* $\mathbb{R}$*. define* $f_\alpha(p, q) = r$ *setwise, as* $r(S) = \alpha p(S) + (1 - \alpha)q(S)$ *for every borel-measurable set* $S \subset \mathbb{R}$*. Again, this is a mixture space.*

So the mixture space structure is relatively flexible.

Handily, it turns out that assumptions akin to the ones that we have already made are enough to guarantee an affine representation on a mixture space

Theorem 4 (The Mixture Space Theorem) *Let* Π*,* $\{f_\alpha\}$ *be a mixture space, and* $\succeq$ *be a complete preference relation such that*

1. *$\pi \succeq \rho$ implies that $f_\alpha(\pi,\mu) \succeq f_\alpha(\rho,\mu)$ $\forall$ $\alpha \in (0,1]$ and $\mu \in \Pi$ (independence)*

2. *for any $\pi,\rho,\mu \in \Pi$ such that $\pi \succ \rho \succ \mu$, there exists and $\alpha,\beta \in [0,1]$ such that $f_\alpha(\pi,\mu) \succ \rho \succ f_\beta(\pi,\mu)$ (the Archemedian Axiom)*

Then there exists a function $v : \Pi \to \mathbb{R}$ that represents $\succeq$ and is affine (i.e. $v(f_\alpha(\pi,\rho) = \alpha v(\pi) + (1-\alpha)\rho)$

This theorem should look familiar. Basically, the assumptions look similar to those we used in the finite prize space case, and the results show the existence of an affine utility representation, which was an intermediate step in proving the existence of an expected utility function.

The assumptions are not identical to the ones that we made in the simple lottery case - note that we have replaced continuity with this Archemedian axiom. In fact, in the simple lottery case, the two are interchangable - we could have changed axiom 2 for the archemedian axiom in section 2 and everything would have gone through fine. In this case, we are using the Archimedian axiom because we *cannot* as yet use the continuity axiom, as we have not put a topology on Π - thus we cannot define continuity. We will address this problem below.

We will not prove the mixture space theorem here, but roughly speaking it follows along similar lines to the first bit of the proof in section 2. Namely

1. Prove equivalents of lemmata 1, 2 and 3. There is an issue here. In the proofs of these lemma have implicitly used the fact that we end up mixing between real numbers, so (for example) we can do the following

$$\begin{aligned}
& f_a(p, f_b(r,s)) \\
= \;& ap + (1-a)br + (1-a)(1-b)s \\
= \;& (a+b-ab)\left(\frac{a}{a+b-ab}p + \frac{(1-a)b}{a+b-ab}r\right) + (1-(a+b-ab))s \\
= \;& h_{(a+b-ab)}(h_{a/(a+b-ab)}(p,r), s)
\end{aligned}$$

   However, this property is not true for general mixture spaces. There are two ways round this problem. One is to add this assumption (it is true in all the examples we used above, for example) or we can modify the proofs not to need this assumption (they hold anyway - see Fishburn [1970])

2. Use these assumptions to construct an affine function. In section 2 we did this by identifying the best and worst lotteries $\delta(z^*)$ and $\delta(z_*)$, and identifying $v(p)$ as the α that make $p \sim f_\alpha(\delta(z^*),\delta(z_*))$. Obviously we can't do this here, as there may not be a best and worst prize in the set (think of the case of the mixture space formed of lotteries over a countable prize space Z. Would we expect this to have a best and worst lottery in general). Luckily, this doesn't turn out to be that much of a problem. just pick any $\pi, \rho \in \Pi$ such that $\pi \succ \rho$ (if we can't find two such elements, then we are in a trivial world). Then define v in the following way for any $\mu \in \Pi$

    (a) if $\pi \succeq \mu \succeq \rho$, then let $v(\mu) = \alpha$ such that $\mu \sim f_\alpha(\pi,\rho)$ (this is just our standard case - we know that such an α exists by the equivalent to lemma 2.)

    (b) If $\mu \succ \pi$, then let $v(\mu) = \frac{1}{\alpha}$ where $\pi \sim f_\alpha(\mu,\rho)$ (again, we know that such an α exists by the equivalent to lemma 2.)

    (c) if $\rho \succ \mu$, then let $v(\mu) = \frac{\alpha}{\alpha-1}$ where $\rho \sim f_\alpha(\pi,\mu)$

   The proof that v is affine is tedious (it is true though, you should try a bit of it). However, this should feel right. First note that the function does at least represent $\succeq$. For $\mu \succ \pi$, $v(\mu) > 1$ (and so greater that $v(\omega)$ for $\pi \succeq \omega$.) Moreover, as μ gets 'better', then the weight on it that maks the mixture of that and ρ indifferent to π goes down, and so $v(\mu)$ goes up. Also, it is clear that, if the function is to be affine, then is *has* to look like this To see this, note that, for $\mu \succ \pi$

$$\begin{aligned} \pi &\sim f_\alpha(\mu, \rho) \\ &\Rightarrow v(\pi) = v\left(f_\alpha(\mu, \rho)\right) \\ &= \alpha v(\mu) + (1-\alpha)v(\rho) \end{aligned}$$

   Moreover, it is clear that $v(\pi) = 1$ and $v(\rho) = 0$, so

$$\begin{aligned} 1 &= \alpha v(\mu) \\ &\Rightarrow \frac{1}{\alpha} = v(\mu) \end{aligned}$$

similarly, for $\rho \succ \mu$

$$\begin{aligned}
\rho &\sim f_\alpha(\pi,\mu) \\
&\Rightarrow v(\rho) = v(f_\alpha(\pi,\mu)) \\
&= \alpha v(\pi) + (1-\alpha)v(\mu) \\
&\Rightarrow \alpha + (1-\alpha)v(\mu) = 0 \\
&\Rightarrow v(\mu) = \frac{\alpha}{\alpha - 1}
\end{aligned}$$

Of course, this is not the same as showing that the function is affine, but at least we know we are dealing with the only candidate

Rather than using the Archemedian axiom, another approach is to apply a topology to the space of lotteries, and demand that preferences satisfy continuity. We will pursue this approach as well, as it will buy us something later on. The topology we are going to use is called the *weak topology,* which we were introduced to . Now we have a topology, we can demand that our preferences be continuous. It turns out that replacing the Archimedian axiom with continuity not only still allows us to have an affine representation, but also guarantees the existence of a continuous reprsentation

Theorem 5 *Let Z be a separable metric space, and $\Delta(Z)$ be the set of borel probability measures on Z. Let $\succeq$ be a complete preference relation on $\Delta(Z)$ that satisfisfies independence and is continuous in the weak topology. Then there exists a $v : \Delta(Z) \to \mathbb{R}$ that is continuous, affine, and represents $\succeq$*

Again, we will not go into the details of the proof here, but essentially this result follows directly from the mixture space theorem. As previously discussed, we can think of $\Delta(Z)$ as forming a mixture space if we define mixtures in the natural way, so the existence of an affine representation requires only that continuity implies the Archemedian axiom. The proof that the resulting representation is continuous is slightly more technical, but can be found in Ok 'Probability Theory with Economic Applications" chapter F.

3.2 From Affine Functions to Expected Utility Representations

The next thing that we have to do is determine whether the existence of an affine function is enough to get us an expected utility representation. If you look back at the proof in section 2, you will see that this step was accomplished in the case of a finite prize space using induction on the size of the support of the lottery. This provides us with at least one more case in which affinity is enough: The case of simple probability distributions on an arbitrary prize space. Simple probability distributions definitionally have finite support, so the proof from section 2 goes through verbatum.

What about cases in which lotteries may not have finite support? Obviously we cannot prove our desired result by induction, but is it still true? In other words, is the mixture theorem enough to buy us an expected utility representation. Unfortunately the answer is no (See Kreps Chapter 5 for a counter example). Some additional assumptions are needed. One possibility is the 'sure thing' principle, which states that, if everything in the support of lottery p is preferred as a sure thing to lottery q, then p must be preferred to q. However, another nice way to proceed is through continuity. It turns out (though we will not prove it), that a continuous, affine function on a separable metric space has an expected utility representation:

Theorem 6 *Let Z be a separable metric space and $\Delta(Z)$ be the set of borel probability measures on Z. if $v : \Delta(Z) \to \mathbb{R}$ is continuous and affine, then there exists a function $u : Z \to \mathbb{R}$ such that $v(p) = \int_Z u(z)dp(z) \; \forall \; p \in \Delta(Z)$*

This theorem, together with theorem 5 tell us that, for a separable metric space, a continuous, complete preference relation on the set of borel probability distributions that satisfies independence will have an expected utiliy representation

4 Risk Aversion

We are now going to have a breif discussion of risk aversion. Presumably we all have an intuitive notion of risk aversion, and presumably this is something to do with rejecting a lottery in favor of receiveing the expected value of the lottery for sure. Note that, for this to make sense, the term 'expected value' has to make sense. Therefore, we need these lotteries to be over numbers, or specifically monetary amounts. For this discussion, we will consider the behavior of someone who has preferences over simple lotteries on $\mathbb{R}$ (a set we will define as P_S) that satisfy the mixture space axioms. We will also assume that people prefer more money to less - in other words

$$\delta(z) \succeq \delta(z')\text{if and only if } z \geq z'$$

Note that this is enough both to ensure that the expected utility function $u : \mathbb{R} \to \mathbb{R}$ is strictly increasing, and also that the preferences obey first order stochastic dominance.

We will let $e(p)$ refer to the the expected value of the lottery p and $v(p)$ refer to its variance.

We will define risk aversion in the intuitive way:

Definition 4 *A decision maker is risk averse if $\delta_{e(p)} \succeq p \; \forall \; p \in P_S$. A desision maker is strictly risk averse if $\delta_{e(p)} \succ p \; \forall \; p \in P_S$. such that $v(p) > 0$. Risk loving is defined analogously. Risk neutrality implies that $\delta_{e(p)} \sim p \; \forall \; p \in P_S$*

One basic result that we can derive is that global risk aversion is equivalant to the claim that the utility function u is concave.we can see this by applying Jensen's inequality, which essentially states that the expectation of a concave function is below its expected value. We will prove this for functions on a finite support, but it is in fact true more generally

Theorem 7 (Jensen's Inequality) *Let $f : \mathbb{R} \to \mathbb{R}$. Then f is concave if and only if*

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \geq \sum_{i=1}^{m} \lambda_i f(x_i)$$

for any $\lambda \in \mathbb{R}_+^m$ such that $\sum_{i=1}^{m} \lambda_i = 1$

Proof. *One direction is obvious. The other direction we can prove using induction. Assume it is true for $m - 1$, then note that*

$$\begin{aligned} & \sum_{i=1}^{m} \lambda_i x_i \\ = \quad & \lambda_1 x_1 + (1 - \lambda_1) \sum_{i=2}^{m} \frac{\lambda_i}{(1 - \lambda_1)} x_i \end{aligned}$$

so we know by the concavity of f that

$$f\left(\lambda_1 + (1 - \lambda_1) \sum_{i=2}^{m} \frac{\lambda_i}{(1 - \lambda_1)} x_i\right) \geq \lambda_1 f(x_1) + (1 - \lambda_1) f\left(\sum_{i=2}^{m} \frac{\lambda_i}{(1 - \lambda_1)} x_i\right)$$

but, by the inductive assumption

$$f\left(\sum_{i=2}^{m}\frac{\lambda_i}{(1-\lambda_1)}x_i\right)\geq\sum_{i=2}^{m}\frac{\lambda_i}{(1-\lambda_1)}f(x_i)$$

■

A direct corrolary of this is the following theorem

Theorem 8 *An expected utility decision maker is risk averse if and only if their utility function is concave, and risk loving if and only if it is convex.*

The next job is to define the concept of certainty equivalence. This is the amount of money with certainty that the DM considers indifferent to a lottery p.

Definition 5 *The certainty equivalence of a lottery p is denoted as $C(p)$ and is defined as*

$$C(p)=\{z\in\mathbb{R}|\delta_z\sim p\}$$

We would like this to be a well defined function. In general, this is not the case, but with some more assumptions, we can ensure that it is

Proposition 1 *For an expected utility decision maker*

1. *If the decsion maker prefers more money to less, then $C(p)$ contains at most 1 element for every p.*
2. *If u is continuous, then $C(p)$ is non-empty*
3. *If the decision maker is globally risk averse, then $C(p)$ is non-empty*

This proposition is pretty trivial to prove, especially when one realises that concave functions are continuous on their interior.

Next we have to define the risk premium of a lottery. This is just the difference between its expected value and its certainty equivalence

Definition 6 *The risk premium of a lottery p is denoted as $r(p)$ and is defined as*

$$r(p)=e(p)-C(p)$$

4.1 Risk Aversion and Wealth

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 7 *We define someone as exhibiting decreasing absolute risk aversion if*

$$\begin{aligned} \delta_w + q &\succeq \delta_w + \delta_z \\ &\Rightarrow \delta_{w'} + q \succeq \delta_{w'} + \delta_z \end{aligned}$$

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 8 *For an expected utility maximizer with* $u : Z \to \mathbb{R}$*, 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 9 *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 choses p over δ_z, then B also chooses p over δ_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.*

5 Evidence on Expected Utility Theory

We are now going to examine some of the experimental evidence regarding expected utility theory. When we talked about evidence with respect to standard utility maximization, we discussed two different types of test. The first was to take a set of 'typical' choices, and see if they were close to satisfying the relevant axioms according to some particular metric. The second was to look at particular types of choice that are designed to highlight failures of the axiom set. In this section we are going to concentrate largly on the latter type of evidence.[2] In other words, we are going to look at particular cases in which the EU axioms (and in particular the independence axiom) are violated. We will begin by discussion some of the observed violations of the predictions of EU theory, then discuss some of the suggested remedies.

5.1 The Machina Triangle

In order to demonstrate these failures graphically, we are going to introduce a handy tool called the Machina triangle.[3] This is a way of graphically representing preferences and choices over lotteries when there are three possible prizes. (prize A, B and C) In this case, any lottery can be reprented by two numbers: the probability of winning prize A, and probability of winning prize B, as this uniquely determines the probability of winning prize C. Thus, the probability simplex over these three prizes can be represented by a triangle in two dimentions, of the type demonstrated in figure 1. In this figure, the point p represents the lottery that gives prize A with probability 0.4, prize B with probability 0.1. and prize C with probability 0.5.

We can represents preferences on this space of lotteries using indifference curves. Furthermore, different decision making models will put different constraints on what these indifference curves looks like. Let's think about the case of an expected utility maximizer. Such a person has preferences that can be represented by a utility function of the form

$$U(p) = u(a)p_a + u(b)p_b + u(c)p_c$$

where $u(x)$ is the utility of prize x and p_x is the probability attached to prize x by lottery p To make things interesting, lets assume that the degenerate lotteries are strictly ranked, then without

[2]Evidence of the first type does exist. See for example Choi et al [2007]

[3]Named after Mark Machina, the famous theorist at UCSD who came up with the idea.

loss of generality, let's assume that $\delta_a \succ \delta_c \succ \delta_b$ We can then normalize $u(a) = 1$ and $u(c) = 0$. $u(b)$ will therefore be strictly negative. The utility of a lottery is therefore given by

$$U(p) = p_a + u(b)p_b$$

An indifference curve is identified by a set of lotteries that all have the same utility, say $\bar{u}$. Thus, one particular indifference curve is identified by the equation

$$p_a + u(b)p_b = \bar{u}$$

or, rearranging

$$p_a = \bar{u} - u(b)p_b$$

This tells us a lot about what the indifference curve of an expected utility maximizer must look like in the Machina triangle: First, they must be straight lines (as $u(b)$ is a constrant). Second, they must be parallel (as the slope of the curve, $-u(b)$ is independent of of the actual indifference curve (defined by $\bar{u}$). Third, they must be upward sloping (as $u(b)$ is negative). Thus, *any* expected utility maximizer must have indifference curves that are of the form shown in Figure 2.

We will now use this tool to discuss some of the empirical problems with expected utility, and some of the attempts that have been made to solve them

5.2 Violations of EU Theory

5.2.1 Allais Paradox, Common Consequence and Common Ratio Effects

Possibly the most famous experiment in economics was suggested by Maurice Allias, a French economist who took a dislike to the assumptions of "l'ecole Américaine". The experiment involves two binary choices, each between two lotteries. The first choice is between:

- Lottery a_1: 100% chance of \$1 million[4]
- Lottery a_2: 10% chance of \$5 million, 89% chance of \$1 million 1% chance of \$0

[4] This is how the experiment was originally suggested as a thought experiment. The astute amongst you may have guessed that it is not usually run as an actual experiment with these amounts.

The second choice is between

- Lottery b_1: 10% chance of \$5 million, 90% chance of \$0
- Lottery b_2: 11% chance of \$1 million, 89% chance of \$0

The modal response (REF) is for subjects to choose a_1 over a_2 and b_1 over b_2. However, this is clearly a violation of the independence axiom, as

$$\begin{aligned}
a_1 &= 0.89\{\$1m\} + 0.11\{\$1m\} \\
a_2 &= 0.89\{\$1m\} + 0.11\left\{\frac{10}{11}\$5m + \frac{1}{11}\$0\right\} \\
b_1 &= 0.89\{\$0\} + 0.11\left\{\frac{10}{11}\$5m + \frac{1}{11}\$0\right\} \\
b_1 &= 0.89\{\$0\} + 0.11\{\$1m\}
\end{aligned}$$

In terms of the Machina triangle, it is clear from figures 3 and 4 that these choices cannot be explained by parallell in difference curves, letting prize a by \$5m, c be \$1m and b be \$0.

Subsequently, the Allais paradox has been seen as an example of a type of violation of the independence axiom known as the common consequence effect. The common consequence effect has the following structure: first a choice between a_1 and a_2, where a_1 and a_2 have the following structure

- $a_1 : \alpha$ chance of x and $(1-\alpha)$ chance of p^{**}
- $a_2 : \alpha$ chance of p and $(1-\alpha)$ chance of p^{**}

Then a choice between b_1 and b_2 which have the following structure:

- $b_1 : \alpha$ chance of p and $(1-\alpha)$ chance of p^{*}
- $b_2 : \alpha$ chance of x and $(1-\alpha)$ chance of p^{*}

Where:

1. p has prizes above and below x

2. p^{**} first order stochastically dominates p^*

The commonly observed violation of independence is that a_1 is chosen over a_2 but b_1 is chosen over b_2. Thus people seem to find the lottery p more attractive relative to the sure thing x when they are paired with the inferior lottery p^* than when paired with the superior lottery p^{**}. The Allais paradox is an example of the common consequence effect with $p = \left\{\frac{10}{11}\$5m + \frac{1}{11}\$0\right\}$, $x = \{\$1m\}$, $p^* = \{\$0\}$ and $p^{**} = \{\$1m\}$.

It is worth noting that the common consequence effect often goes away when it is pointed out to subjects, in the sense that the lotteries are explicitly described as two stage lotteries as they are above.

A second type of commonly observed violation is illustrated by the following pair of choices. The first choice is between:

- Lottery a_1: 100% chance of \$3000
- Lottery a_2: 80% chance of \$4000, 20% chance of \$0

The second choice is between

- Lottery b_1: 25% chance of \$3000, 75% chance of \$0
- Lottery b_2: 20% chance of \$4000, 80% chance of \$0

Here, the modal choice is of a_1 and b_2. Again, this is a violation of the independence axiom, as

$$b_1 = 0.25a_1 + 0.75\{\$0\}$$

and

$$b_2 = 0.25a_2 + 0.75\{\$0\}$$

This is again an example of a broader class of violations, known as the common ratio effect. These have the following structure :first a choice between a_1 and a_2,

- $a_1 : p$ chance of x and $(1 - p)$ chance of 0

- a_2 : q chance of y and $(1-q)$ chance of 0

Then a choice between b_1 and b_2 which have the following structure:

- b_1 : αp chance of x and $(1-\alpha p)$ chance of 0
- b_2 : αq chance of y and $(1-\alpha q)$ chance of 0

Where:

1. $p > q$
2. $y > x > 0$