Last updated: October 16, 2017

PhD Micro (Part 1)
Expected Utility

Motaz Al-Chanati

October 3, 2017

1 Lotteries

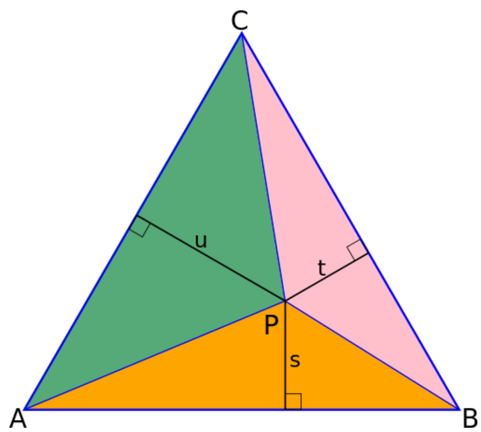
We're now going to switch back to thinking about consumers and their decisions. Before, we considered choices over bundles of objects. There was no uncertainty there: if I chose x over y, then I would definitely be getting x. Now, we will introduce uncertainty in the form of *lotteries*. A lottery consists of two elements:

1. Prizes: a bundle x that the consumer could win from the lottery. This prize comes from a *finite* set of prizes X (i.e. $x \in X$)
2. Probabilities: the consumer wins a prize x with probability $p(x) \in [0, 1]$. The vector of probabilities for all bundles p comes from the set of all possible probability distributions over X, which we denote as ΔX (i.e. $p \in \Delta X$). In class, we called this set $\mathscr{L}$. It must also be that the probabilities add up to 1, i.e. $\sum_{x \in X} p(x) = 1, \forall p \in \Delta X$

Let's have an example so we're comfortable with the terminology. Say that we have a lottery where you could win either an apple, a banana, or a car. Our set of prizes is then $X = \{\text{apple, banana, car}\}$. What are the possible probabilities? There are infinitely many! All we need is a 3×1 vector $p = (p(a), p(b), p(c))$ such that $p(a) + p(b) + p(c) = 1$. For example, we could have $p = (0.5, 0.2, 0.3)$, which says that you win the apple with 50% probability, the banana with 20% probability, and the car with 30% probability. So p is a possible probability distribution, i.e. $(0.5, 0.2, 0.3) \in \Delta X$. However, so is $(0.1, 0.1, 0.8) \in \Delta X$ or $(0, 1, 0) \in \Delta X$. As always X can be any arbitrary set, but for the most part, we're going to use $X \subset \mathbb{R}$ so that we interpret the prizes as money. This makes it easier to conceptualize these situations, but as our example above shows, the prizes can be over absolutely anything.

We can also represent the probability vector as a function $p : X \to [0, 1]$, i.e. $p(x)$ takes an element $x \in X$ and tells us what probability $p(x) \in [0, 1]$ the lottery assigns it (with the condition $\sum_{x \in X} p(x) = 1$). Another way to write it is to say that there K possible prizes, i.e. $|X| = K$ (remember that X is finite), and then we can express the probability vector as $p = (p_1, ..., p_K)$ where $p_k = p(x_k)$. A special case is a lottery where you win one prize with probability one. We call these degenerate lotteries, and denote it as δ_x, where $\delta_x(y) = 1$ if $y = x$ and 0 if $y \neq x$. If we used the indexed notation, then $\delta_k(x_j) = 1$ if $x_j = x_k$ and 0 otherwise.

Let's bring in a graphical interpretation. Start with the simplest case of two prizes: x_1 and x_2 with corresponding probabilities p_1 and p_2. Say I wanted to express the possible probabilities. Since $p_2 = 1 - p_1$, then we could just draw a line from 0 to 1 representing p_1. Plotting a point on the line tells us the value of p_1, and the part 'left over' represents the value of p_2. As our point moves closer to 1, we are putting a higher weighting on p_1. So by exploiting the "summing to 1" property, we can represent a 2 dimensional probability vector on a 1 dimensional graph. Similarly, we can represent a 3 dimensional probability vector on a 2 dimensional graph. This is called the *simplex*. We represent this as an equilateral triangle with length 1. See the figure below.[1] Any point within the triangle represents a probability distribution p. A vertex represents a degenerate lottery (so in the figure, the vertices represent - from left to right - $\delta_A, \delta_C, \delta_B$). Since this is an equilateral triangle, we use the property that the sum of perpendicular distances to each edge add to the length of the triangle. In our case, $s + t + u = 1$. This means that the probability for a prize is equal to the distance from the point to the side opposite its corresponding vertex. So in our example: $p = (p_A, p_B, p_C) = (t, u, s)$. Just keep in mind that the closer the point is to a vertex, the larger the weight on the corresponding prize



Finally, you should get comfortable with taking the convex combination of two lotteries. For some $\alpha \in [0, 1]$, and two lotteries p, p', we can have a compound lottery of the form: $\alpha p + (1 - \alpha)p'$. You can interpret this in two ways. It can be seen as a two stage lottery: first you a play a lottery to see which lottery you will play in the second stage (p or p'). You play lottery p with probability α and you play lottery p' with probability $1 - \alpha$. After a lottery is chosen in the first stage, you play the chosen lottery with its given probability values. Alternatively, it can be seen as its own lottery, which has a probability distribution p^* such that $p^*(x) = \alpha p(x) + (1 - \alpha)p'(x)$. For example, if $p = (0.5, 0.5), p' = (0.25, 0.75), \alpha = 0.4$, then you should convince yourself that in either way, you will win x_1 with probability 0.35 and x_2 with probability 0.65. We can generalize compound lotteries such that when we have N lotteries p^n and N weights α^n where $\sum_{n=1}^{N} \alpha^n = 1$, we have a lottery defined by $p^* = \sum_{n=1}^{N} \alpha^n p^n$

2 Preferences

Since we are studying a consumer making decisions, we need to also include preferences. Note that we are going to keep the set of prizes X fixed. All we are doing is choosing between different lotteries, which in effect, means that we are choosing over different probability distributions. From now on we will call our arbitrary lottery p, which tells us the probability weighting and thus fully characterizes the lottery. In the first part of the course, we were comparing bundles out of a consumption set like $x, x' \in X$. Now we are choosing lotteries $p, p' \in \Delta X$. Just as we could have a preference over bundles

[1]Source: https://en.wikipedia.org/wiki/Viviani%27s_theorem

($x \succsim x'$), we will also have preferences over lotteries ($p \succsim p'$). As before we will make the standard rationality assumptions about the preferences:

- **Complete:** For any $p, p' \in \Delta X$, either $p' \succsim p$ or $p \succsim p'$ (or both)
- **Transitive:** For any $p, p', p'' \in \Delta X$, if $p \succsim p'$ and $p' \succsim p''$, then $p \succsim p''$

Our preferences $\succsim$ will also have the usual standard induced strict $\succ$ and indifference $\sim$ preferences. We're going to introduce two new lottery-specific axioms:

Axioms for Preferences over Lotteries

- *Continuity:* For any $p, p', p'' \in \Delta X$, the sets $\{\alpha \in [0,1] : \alpha p + (1-\alpha)p'' \succsim p'\}$ and $\{\alpha \in [0,1] : p' \succsim \alpha p + (1-\alpha)p''\}$ are closed
- *Independence:* If $p \succsim p'$, then for any $p'' \in \Delta X$ and $\alpha \in [0,1]$, it must be that $\alpha p + (1-\alpha)p'' \succsim \alpha p' + (1-\alpha)p''$

These definitions have the following implications:

- (Continuity): If $p \succsim p' \succsim p''$, then there exists $\alpha \in [0,1]$ such that $p' \sim \alpha p + (1-\alpha)p''$
  - This tends to be a more useful definition than the one above. The idea is that those two sets cover the interval $[0,1]$ (by completeness). So if the sets are non-empty (which is guaranteed when $p \succsim p' \succsim p''$), then the sets must have a non-empty intersection which gives the indifference relation
- (Strict Independence) If $p \succ p'$, then for any $p'' \in \Delta X$ and $\alpha \in (0,1]$, it must be that $\alpha p + (1-\alpha)p'' \succ \alpha p' + (1-\alpha)p''$
- (Indifferent Independence) If $p \sim p'$, then for any $p'' \in \Delta X$ and $\alpha \in [0,1]$, it must be that $\alpha p + (1-\alpha)p'' \sim \alpha p' + (1-\alpha)p''$
- (Archimedian) If $p \succ p' \succ p''$, then there exists $\alpha, \beta \in (0,1)$ such that $\alpha p + (1-\alpha)p'' \succ p' \succ \beta p + (1-\beta)p''$
  - *Independence* + *Continuity* $\implies$ *Archimedian*, but also *Independence* + *Archimedian* $\implies$ *Continuity*

When we first studied preferences, we had to assume continuity to guarantee the existence of a (continuous) utility representation. However, if you recall, the definition for continuity there was slightly different, so make sure to use the right one depending on the context. It does still play a similar role in that continuity means that we shouldn't see preferences "jump" and it guarantees the existence of utility representation. Note that so far there is actually nothing overly special about lotteries. We could stop here and we would have similar results as in the first part of the course. Our goal though is to see if we can get something more special given that we are living in a world of uncertainty. For that we introduce the independence assumption.

The independence axiom essentially tells you to only compare lotteries at where they differ. If you're going to get lottery p'' with $1-\alpha$ probability anyway, then all that really matters should be the other lottery you get with α probability. The independence axiom is a lot easier to believe if you think about compound lotteries as two-stage lotteries (and see them expressed as such). However, if you think of compound lotteries as pure lotteries by themselves, then it can be quite easy to violate. Let me give one example in support of independence and one example against it.

Dutch Book

Suppose your preferences violated independence. Say that you strictly prefer lottery p over p', i.e.

$p \succ p'$. However, for some other lottery p'' and $\alpha \in [0,1]$, we have $\alpha p' + (1-\alpha)p'' \succ \alpha p + (1-\alpha)p''$, which clearly violates the axiom. Let's call the two compound lotteries L' and L (so we have $L' \succ L$). Imagine you have a ticket to play lottery L, but I come along and say that I can give you a ticket for L' if you give me your ticket for L plus some money. Since $L' \succ L$, you must be willing to pay some amount of money ε_1 to get L' over L. We do the trade and now you have lottery L' (and presumably you are better off). So now you're about to face getting either p' or p'' through the first-stage lottery. I come along again and say that if in the first stage you get p', I promise to replace the lottery with p if you give me some money now. Note that the money is made before the first-stage lottery is played. Since $p \succ p'$, you must be willing to pay some amount ε_2. So now you have paid a total of $\varepsilon_1 + \varepsilon_2$ to get a lottery of $\alpha p + (1-\alpha)p''$... but that's exactly lottery L, which is what you started with. You've made a trade that has made you strictly worse off than before. If you had satisfied independence then you couldn't have been "Dutch-booked".

Allais Paradox

Suppose you had a choice of two lotteries:

- p^1: Win \$1000 with 100% probability
- p^2: Win \$1500 with 80% probability and \$0 with 20% probability

Make your choice here. Say you pick p^1, so we conclude $p^1 \succsim p^2$. Now you face another choice of lotteries:

- p^3: Win \$1000 with 25% probability and \$0 with 75% probability
- p^4: Win \$1500 with 20% probability and \$0 with 80% probability

Again make a choice here - let's say you pick $p^4 \succsim p^3$. But doing this would violate independence! Let's express p^3 and p^4 in a different way:

- p^3: Play lottery p^1 with 25% probability and \$0 with 75% probability
- p^4: Play lottery p^2 with 25% probability and \$0 with 75% probability

Now you can clearly see by independence that you should have $p^1 \succsim p^2 \iff p^3 \succsim p^4$. Note that $0.25 \times p^1$ essentially means getting \$1000 with 25% probability; $0.25 \times p^2$ means getting \$1500 with $0.25 \times 0.8 = 0.2$ probability and \$0 with $0.25 \times 0.2 = 0.05$ probability. Adding these to a (degenerate) lottery where you get \$0 with $0.75 \times 1 = 0.75$ probability gives us exactly the first expressions for p^3 and p^4. Doing these extra computations can make it quite a lot easier to violate the independence axiom.

Let's give some graphical interpretation to our preferences. Imagine drawing indifference curves on the simplex (or look at MWG Figure 6.B.5). We know that indifference curves should connect two points that the consumer is indifferent between, say p and p'. Plot those two points on the simplex. The independence axiom tells us that we should have $\alpha p + (1-\alpha)p'' \sim \alpha p' + (1-\alpha)p''$, for any $p'' \in \Delta X$. Letting $p'' = p$ or p' and varying α shows us that this convex combination just means that there should a straight line connecting p and p'. If this is hard to see, you can pick a vertex V to start with and think of p and p' as vectors coming out of V. Then you know that a convex combination of vectors will get you to some point in the middle of them. So we know the indifference curves have to be straight, but we also know they have to be parallel. Remember that we interpreted the weight on a prize as being the perpendicular distance to the opposite edge. Therefore, changing the weighting of one prize (holding the relative weighting of the other two prizes) is just a parallel shift of the indifference curve. Finally, in our usual 2D no uncertainty case, it was very clear that higher indifference curves meant higher utility. You need to be a bit more careful with the simplex since it is actually a 3D representation. All you need to remember is that there are three prizes, which means that one of them has to be the 'best' one. So your most desirable lottery is the vertex corresponding to the best prize (we'll talk more about this line of argument later). Therefore indifference curves moving towards the best prize indicates higher utility.

3 Expected Utility

Just like last time, we want to move from preferences to utilities. Recall what we mean by utility representation: it means that we can find some function $U(\cdot) : \Delta X \to \mathbb{R}$ where:

$$p \succsim p' \iff U(p) \geq U(p')$$

This is already guaranteed by continuity. We're going to try to go a step further and see under what conditions we can have an *expected utility* representation.

Expected Utility Form
A utility function $U(\cdot) : \Delta X \to \mathbb{R}$ has an expected utility form if there is a function $u : X \to \mathbb{R}$ such that

$$U(p) = \sum_{x \in X} u(x)p(x) = \sum_{k=1}^{K} u(x_k)p_k$$

We call the little $u(\cdot)$ the Bernoulli utility function and it represents utility on the final outcomes (prizes). The big $U(\cdot)$ is called the von Neumann-Morgenstern (vNM) utility function, and it is utilities over lotteries. The idea here is that to evaluate lotteries we take all the possible outcomes, see what utility you would get from it, then weight them by the probability of the outcome occurring. So this is taking the expectations of the utilities. Note that this is not the same as taking the utility of the expectation! (we'll talk about that when we get to risk)

Let's see two useful properties of vNM utility:

1. **Utility has expected utility form if and only if it is a linear function**
   $U\left(\sum_{n=1}^{N} \alpha^n p^n\right) = \sum_{n=1}^{N} \alpha^n U(p^n)$ for N lotteries $p^1, ..., p^N$ and corresponding weights $\alpha^1, ..., \alpha^N$ such that $\sum_n \alpha^n = 1$

2. **Expected utility representation is defined up to an increasing affine transformation**[2]
   If $U(\cdot)$ is an expected utility representation for $\succsim$ then $V(\cdot)$ is another expected utility representation for $\succsim$ if and only if there exists $a > 0$ and $b \in \mathbb{R}$ such that $V(p) = aU(p) + b$ (or equivalently, $V(p) = \sum_{x \in X} v(x)p(x)$, where $v(x) = au(x) + b$)

Let's do the proof for the first one. This doesn't take anything tricky, but it does require you to be careful with notation. To make things easier we will use the k-indexed notation. Also note that $\sum \alpha^n p^n$ is just a convex combination of lotteries. So we can define this composite lottery, p^*, as $p_k^* = \sum_{n=1}^{N} \alpha^n p_k^n$.

Proof of (1)

[2] An affine transformation in $\mathbb{R}^N$ is a mapping $f : \mathbb{R}^N \to \mathbb{R}^N$ such that $f(x) = Ax + b, \forall x \in \mathbb{R}^N$, where A is an invertible $N \times N$ matrix and $b \in \mathbb{R}^N$

- *Expected Utility ⇒ Linearity:*

$$
\begin{aligned}
U\left(\sum_{n=1}^{N}\alpha^n p^n\right) = U(p^*) &= \sum_{k=1}^{K} u(x_k)p_k^* && \text{Definition of expected utility}\\
&= \sum_{k=1}^{K} u(x_k)\left[\sum_{n=1}^{N}\alpha^n p_k^n\right] && \text{Definition of } p^*\\
&= u(x_1)\alpha^1 p_1^1 + \dots u(x_1)\alpha^N p_1^N + \dots\\
&\quad \dots + u(x_K)\alpha^1 p_K^1 + \dots u(x_K)\alpha^N p_K^N && \text{Expand the double sum}\\
&= u(x_1)\alpha^1 p_1^1 + \dots u(x_K)\alpha^1 p_K^1 + \dots\\
&\quad \dots + u(x_1)\alpha^N p_1^N + \dots u(x_K)\alpha^N p_K^N && \text{Re-arrange}\\
&= \sum_{n=1}^{N}\alpha^n\left[\sum_{k=1}^{K} u(x_k)p_k^n\right] && \text{Simplify into new double sum}\\
&= \sum_{n=1}^{N}\alpha^n U(p^n) && \text{Definition of expected utility}
\end{aligned}
$$

- *Linearity ⇒ Expected Utility:*
  - Note that any lottery p^* can be written as the convex combination of K degenerate lotteries, where the weight for lottery δ_k is the corresponding probability in p^*. In the definition above, let $N = K$, $\alpha^n = p_k^*$, $p^n = \delta_k$, i.e. $p_k^* = \sum_k p_k\delta_k$.
  - Also note that for any degenerate lottery, $U(\delta_k) = \sum_{j\neq k}[0 \times u(x_j)] + 1 \times u(x_k) = u(x_k)$
  - Since $U(\cdot)$ is linear, then $U(\sum_k p_k\delta_k) = \sum_k p_k U(\delta_k) = \sum_k p_k u(x_k)$. Therefore it has expected utility form.

I will leave the proof for (2) until later when we establish some more ideas about the utility representation. Once we do, this proof will be very simple.

Finally, another important property of expected utility is that indifference curves on the simplex will be linear and parallel. Suppose we have two points on the simplex p and p', such that $U(p) = U(p') = U$ for some utility level U. Then clearly they must be connected by an indifference curve. By linearity, $U(\alpha p + (1-\alpha)p') = \alpha U(p) + (1-\alpha)U(p') = U, \forall\alpha \in [0,1]$. Therefore, any convex combination also achieves utility U and must be on the same indifference curve. So this shows that indifference curves are linear, but what about parallel? Take some other p''. If they were parallel, we should have that $U(\alpha p + (1-\alpha)p'') = U(\alpha p' + (1-\alpha)p'')$ (the same parallel shift is applied to both). But using linearity we see that this is true if $U(p) = U(p')$, which was our starting assumption.

Notice that the independence axiom and the expected utility form both gave us linear parallel indifference curves (and the logic is quite similar in the arguments too). This gives rise to the key theorem in this section:

> **Expected Utility Representation Theorem**
> $\succsim$ is independent and continuous over $\Delta X \iff$ there exists a function $u : X \to \mathbb{R}$ such that $U(p) = \sum_{x\in X} u(x)p(x)$ is a utility representation for $\succsim$

The proof for this is quite long, so we'll break it up into steps. Let me give a roadmap of what our plan will be:

1. Prove $\Longleftarrow$, i.e. vNM utility representation implies independent and continuous preferences
   1.1 Show independence
   1.2 Show continuity
2. Prove $\Longrightarrow$, i.e. independent and continuous preferences implies vNM utility representation
   2.1 Find a best and worst lottery
   2.2 Show that any lottery is equivalently a convex combination of the best and worst lotteries
   2.3 Define the unique value for $U(p)$ based on (2.2)
   2.4 Show that U is linear
   2.5 Show that $U(p)$ has expected utility form

Ok, let's get into this:

Step 1: Prove $\Longleftarrow$

1.1: Independence

Take any $p \succsim p'$ and an arbitrary p''. Since $U(\cdot)$ represents $\succsim$, then $U(p) \geq U(p')$. Since $U(\cdot)$ is of expected utility form it is also linear, and therefore: $U(\alpha q+(1-\alpha)p'') = \alpha U(q)+(1-\alpha)U(p''), \forall \alpha \in [0,1]$ and for $q \in \{p, p'\}$. Putting this together, we get that:

$$\begin{aligned}
\alpha U(p) + (1-\alpha)U(p'') &\geq \alpha U(p') + (1-\alpha)U(p'') && U(p) \geq U(p') \\
U(\alpha p + (1-\alpha)p'') &\geq U(\alpha p' + (1-\alpha)p'') && \text{Linearity} \\
\alpha p + (1-\alpha)p'' &\succsim \alpha p' + (1-\alpha)p'' && \text{Utility representation}
\end{aligned}$$

1.2: Continuity

Take any $p \succsim p' \succsim p''$. By utility representation, $U(p) \geq U(p') \geq U(p'')$. We can define the interval $[U(p''), U(p)]$ where $U(p') \in [U(p''), U(p)]$. Since any subset of $\mathbb{R}$ is convex, there exists $\alpha \in [0,1]$ such that $U(p') = \alpha U(p) + (1-\alpha)U(p'')$. Putting this together, we get that:

$$\begin{aligned}
U(p') &= \alpha U(p) + (1-\alpha)U(p'') && \text{Convexity} \\
&= U(\alpha p + (1-\alpha)p'') && \text{Linearity} \\
\therefore p' &\sim \alpha p + (1-\alpha)p'' && \text{Utility representation}
\end{aligned}$$

Step 2: Prove $\Longrightarrow$

2.1: Find best and worst lottery

Note that the 'best' option means it is weakly preferred to all other alternative and the 'worst' option means that all other alternatives are weakly preferred to it. X is a finite set so there are a finite number of prizes $x_1, ..., x_K$. We can always find a best and worst lottery. Compare x_1 and x_2. By completeness, one must be preferred and call it x^* and other one x_*. Then take x_3 and compare it to x^* and x_*. If $x_3 \succsim x^*$, then let $x^* = x_3$. Likewise, if $x_* \succsim x_3$, then let $x_* = x_3$. If neither, then keep x^* and x_* the same. By completeness, this covers all possible cases. Continue with this procedure until you go through all K prizes. By transitivity, we will have that $x^* \succsim x_k \succsim x_*, \forall k \in 1, ..., K$. x^* is the best prize and x_* is the worst prize.

Next we will find p^* and p_*, the best and worst *lotteries*. We claim that $p^* = \delta_{x^*}$ (the degenerate lottery for x^*) and similarly $p_* = \delta_{x_*}$ (the degenerate lottery for x_*). Note that we can write any arbitrary lottery p as $p = \sum_k p_k \delta_k$ (we showed this before). Since we had $x^* \succsim x_k \succsim x_*, \forall k$, and $x_k \sim \delta_{x_k}$, then we must also have $\delta_{x^*} \succsim \delta_{x_k} \succsim \delta_{x_*}, \forall k$. Therefore, take some arbitrary $p \in \Delta X$:

$$
\begin{aligned}
\delta^* &\sim \sum_{k=1}^{K} p_k \delta^* && \text{Since } \textstyle\sum_k p_k = 1 \\
&\succsim p_1\delta_1 + \sum_{k=2}^{K} p_k \delta^* && \text{Independence axiom; } \delta^* \succsim \delta_1 \\
&\succsim \sum_{j=1}^{2} p_j \delta_j + \sum_{k=3}^{K} p_k \delta^* && \text{Independence axiom; } \delta^* \succsim \delta_2 \\
&\ \vdots \\
&\succsim \sum_{j=1}^{K} p_j \delta_j && \text{Keep iterating} \\
&\sim p && \text{Any } p \text{ can be expressed as above}
\end{aligned}
$$

A similar argument holds for $p \succsim \delta_*$. Therefore, we showed for any arbitrary lottery p, $\delta^* \succsim p \succsim \delta_*$. So, our best and worst lotteries are indeed $p^* = \delta_{x^*}$ and $p_* = \delta_{x_*}$

2.2: Any lottery is a unique convex combination of the best and worst

Since $p^* \succsim p \succsim p_*$, then by continuity, there exists $\alpha \in [0, 1]$ such that $p \sim \alpha p^* + (1 - \alpha)p_*$. We want to show that this α is also unique. If $p \sim p^*$, then $\alpha = 1$ and if $p \sim p_*$, then $\alpha = 0$. So we only need to worry about the case where $p^* \succ p \succ p_*$.

Suppose not, and there are two scalars $\alpha, \beta \in [0, 1]$ such that $p \sim \alpha p^* + (1-\alpha)p_*$ and $p \sim \beta p^* + (1-\beta)p_*$. Without loss of generality, let's also suppose that $\beta > \alpha$. This means that $\frac{\beta-\alpha}{1-\alpha} \in [0, 1]$. Therefore:

$$
\begin{aligned}
p &\sim \beta p^* + (1 - \beta)p_* && \text{by assumption} \\
&\sim \left(\frac{\alpha}{\beta} + \left(1 - \frac{\alpha}{\beta}\right)\right) \beta p^* + (1 - \beta)p_* && \text{essentially multiplying by 1} \\
&\sim \alpha p^* + (\beta - \alpha)p^* + (1 - \beta)p_* \\
&\sim \alpha p^* + (1 - \alpha) \left[\frac{\beta - \alpha}{1 - \alpha} p_* + \frac{1 - \beta}{1 - \alpha} p_*\right] \\
&\sim \alpha p^* + (1 - \alpha) \left[\frac{\beta - \alpha}{1 - \alpha} p^* + \left(1 - \frac{\beta - \alpha}{1 - \alpha}\right) p_*\right] \\
&\succ \alpha p^* + (1 - \alpha) \left[\frac{\beta - \alpha}{1 - \alpha} p_* + \left(1 - \frac{\beta - \alpha}{1 - \alpha}\right) p_*\right] && \text{independence since } p^* \succ p_* \\
&\sim \alpha p^* + (1 - \alpha)p_* && \text{simplify} \\
&\sim p && \text{by assumption}
\end{aligned}
$$

We have a contradiction: $p \succ p$! This proves that there can only be one unique $\alpha \in [0, 1]$. Make sure you understand that key step when we use independence to switch to a strict preference. The whole expression represents a compound lottery, where the lottery that occurs with $1-\alpha$ chance (the expression in square brackets) is itself a compound lottery. Let $\gamma = \frac{\beta-\alpha}{1-\alpha}$, so independence gives us $P = \gamma p^* + (1-\gamma)p_* \succ \gamma p_* + (1-\gamma)p_* = Q$. But $P \succ Q$ also means that: $\alpha p_* + (1-\alpha)P \succ \alpha p_* + (1-\alpha)Q$ (again, by independence). That's the logic of that key step.

2.3: Define the unique value for $U(p)$

Define $U = \alpha$, i.e. $U(p)$ such that $p \sim U(p)p^* + (1-U(p))p_*$. The uniqueness proof we did in (2.2) effectively shows that $\beta p^* + (1-\beta)p_* \succsim \alpha p^* + (1-\alpha)p_* \iff \beta \geq \alpha$. This shows that $U(p)$ represents the preference: $p \succsim p' \iff U(p) \geq U(p')$.

2.4: Show that $U(p)$ is linear

First, note two claims:

1. $\alpha p + (1-\alpha)p' \sim [U(\alpha p + (1-\alpha)p')]\, p^* + [1 - U(\alpha p + (1-\alpha)p')]\, p_*$
2. $\alpha p + (1-\alpha)p' \sim \alpha\, [U(p)p^* + (1-U(p))p_*] + (1-\alpha)\, [U(p')p^* + (1-U(p'))p_*]$

In the first one, we are just treating $\alpha p + (1-\alpha)p'$ as a lottery and applying the definition of $U(\cdot)$. In the second one, we are plugging in the definition for $U(p)$ and $U(p')$ to replace p and p', respectively. By re-arranging (2) and transitivity:

$$\begin{aligned}
&\alpha\, [U(p)p^* + (1-U(p))p_*] + (1-\alpha)\, [U(p')p^* + (1-U(p'))p_*] \\
&= [\alpha U(p) + (1-\alpha)U(p')]\, p^* + [\alpha(1-U(p))p^* + (1-\alpha)(1-U(p'))]\, p_* \\
&= [\alpha U(p) + (1-\alpha)U(p')]\, p^* + \left[1 - \left(\alpha U(p) + (1-\alpha)U(p')\right)\right] p_* \\
&\sim [U(\alpha p + (1-\alpha)p')]\, p^* + [1 - U(\alpha p + (1-\alpha)p')]\, p_*
\end{aligned}$$

Since we proved that the α weight on the convex combination has to be unique, then it has to be the case that $\alpha U(p) + (1-\alpha)U(p') = U(\alpha p + (1-\alpha)p')$. By induction, we can generalize this from 2 lotteries to N lotteries: $U(\sum \alpha^n p^n) = \sum \alpha^n U(p^i)$. Therefore, $U(\cdot)$ is a linear function.

2.5: Show that $U(p)$ has expected utility form

Since $U(\cdot)$ is linear it also takes expected utility form. Another way to see this is that we can define $u(x_k) = U(\delta_k)$. And so for any $p = \sum_k p_k \delta_k$, we have $U(p) = U\left(\sum_k p_k \delta_k\right) = \sum_k p_k U(\delta_k) = \sum_k p_k u(x_k)$, where the second equality uses the linearity property of $U(\cdot)$.

Finally, as promised, let's show that expected utility representation is unique up to an affine transformation. By the definition of U, we have that $p \sim U(p)p^* + (1-U(p))p_*$. And since $V(\cdot)$ is another expected utility form, it must be linear: $V(\alpha p + (1-\alpha)p') = \alpha V(p) + (1-\alpha)V(p')$. Now we simply let $\alpha = U(p)$, $p = p^*$, $p' = p_*$:

$$\begin{aligned}
V(p) &= V\left(U(p)p^* + (1-U(p))p_*\right) \\
&= U(p)V(p^*) + (1-U(p))V(p_*) \\
&= \underbrace{\left(V(p^*) - V(p_*)\right)}_{=a} U(p) + \underbrace{V(p_*)}_{=b}
\end{aligned}$$

You can see that under the transformed utility function the possible range of utilities is $[V(p_*), V(p^*)]$, while the old interval $[U(p_*), U(p^*)] = [0, 1]$. So b normalizes the lowest value and a scales $U(p)$ to fit the range of the new interval.