Microeconomics 1 Lecture Notes

Enrico Mattia Salonia

November 8, 2025

Work in Progress!

Contents

Preamble

These notes accompany the first part of the PhD microeconomics sequence. They cover **choice under uncertainty** and **general equilibrium theory**. The write-up is still a work in progress, and I will continue to update it. If you spot any mistakes or typos, please let me know.

I have aimed for a conversational style rather than the more formal tone of, say, Mas-Colell et al. (1995). I assume no prior knowledge of the topics, though—as usual—some mathematical maturity helps (and I hope you will develop it along the way!). Each lecture summarizes what we cover in class, followed by exercises and suggestions for further reading. Whenever a result is proved, I have tried to give the simplest proof available. This often makes explanations and proofs a bit longer than strictly necessary, but, I hope, also more accessible.

Before diving in, you might enjoy some non-technical background that helps frame the topics we will study: Kreps (1988, ch. 1), Debreu (1959, pp. ix–xi), Myerson (1997, pp. 1–7), and Gilboa (2009, chs. 1–2).

You will occasionally see smaller text like this. These remarks are not essential for following the main exposition, but they add context or point to related ideas. Feel free to skip them on a first pass.

These notes draw on several sources. The main reference is Mas-Colell et al. (1995), but both here and in the text you will find pointers to alternative or complementary readings. A short reading list follows. If you would like more references or wish to discuss any of the material, just send me an email—I am always happy to talk.

Have fun!

Choice under uncertainty.

- **Mas-Colell et al. (1995), ch. 6.**
- Kreps (1988), chs. 4-6.
- Fishburn (1970), ch. 8.
- Kreps (2013), chs. 5-6.
- Gilboa (2009).

General equilibrium theory.

- **Mas-Colell et al. (1995) chs. 15–16.**
- Thomson (2011), sec 4.3 (no-envy).
- Kreps (2013), chs. 14–15.
- Debreu (1959).
- McKenzie (2005).

References

Debreu, G. (1959). *Theory of value: An axiomatic analysis of economic equilibrium* (Vol. 17). Yale University Press. 1

Fishburn, P. C. (1970). *Utility theory for decision making.* New York: Wiley. 1

Gilboa, I. (2009). *Theory of decision under uncertainty* (Vol. 45). Cambridge university press. 1

Kreps, D. M. (1988). *Notes on the theory of choice.* Westview Press. 1

Kreps, D. M. (2013). *Microeconomic foundations* (Vol. 1). Princeton university press. 1

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory.* Oxford university press New York. 1

McKenzie, L. W. (2005). Classical general equilibrium theory. *MIT Press Books*, *1*. 1

Myerson, R. B. (1997). *Game theory: Analysis of conflict.* Cambridge, Massachusetts, USA: Harvard University Press. 1

Thomson, W. (2011). Fair allocation rules. In *Handbook of social choice and welfare* (Vol. 2, pp. 393–506). Elsevier. 1

Lecture 1

Introduction to uncertainty

1.1 How to model uncertainty

The outcomes of our decisions are often uncertain, so we need a choice theory that takes uncertainty into account. Let us begin by thinking about how to represent uncertainty. Suppose you make a bet with a friend: if a fair coin toss results in heads, you receive 10 euros; otherwise, you pay 10 euros to your friend. There are two possible outcomes, 10 and -10, and since the coin is fair, each occurs with probability $1/2$. What are the main ingredients of this example?

First, we started from a set of possible outcomes—in this case, the monetary transfers 10 and -10. Second, we specified the probability of each outcome occurring, $1/2$ for both. We call such an object—a set of outcomes, each associated with a probability—a **lottery**. Denote the set of outcomes by X. Generic elements of X will be written x, y, z, or sometimes $x_1, x_2, \dots$. For simplicity, assume that X is finite. Outcomes alone are not enough to describe a lottery: we also need a probability distribution over outcomes, as in the $1/2$–$1/2$ distribution of the fair coin above. The set of all lotteries over X is denoted by $\Delta(X)$.[1] Each element of $\Delta(X)$ is a function $p\colon X \to [0,1]$ such that $\sum_{x\in X} p(x) = 1$; it maps each outcome x to a number $p(x) \in [0,1]$, representing the probability that x occurs.[2] We can equivalently represent a lottery as a vector, for example $p = (p(x), p(y), p(z))$ if $X = \{x, y, z\}$.

Example 1.1. In the example above, the set of outcomes is $\{10, -10\}$, and the lottery $p \in \Delta(\{10, -10\})$ induced by the fair coin toss satisfies $p(10) = p(-10) = 1/2$. ■

We can depict lotteries using a tree diagram, as in Figure 1.1.

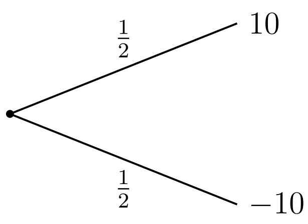


Figure 1.1: Lottery from Example 1.1.

Remark 1.1. *Notice that in this setup we are missing something: whether the coin lands on heads or tails is irrelevant; only the probabilities of the outcomes matter, not the events*

[1]Why the notation Δ? You will see soon.

[2]Why do we write a sum $\sum_{x\in X} p(x) = 1$ rather than an integral?

that generate them. This is a limitation of the model, which we will address later when we introduce a state-space representation of uncertainty.

The set of lotteries $\Delta(X)$ has *structure*: we can combine its elements in a meaningful way. For example, consider a lottery r that yields lottery p with probability α and lottery q with probability $1-\alpha$, where $\alpha \in [0,1]$. Such an object is called a **compound lottery**. It is still an element of $\Delta(X)$, and we write $r = \alpha p + (1-\alpha)q$.

For instance, if $p(10) = {}^1\!/_2$ and $q(10) = {}^1\!/_4$, the associated compound lottery is shown on the left of Figure 1.2. We can compute the probability that outcome 10 occurs in this compound lottery:

$$\alpha \times {}^1\!/_2 + (1-\alpha) \times {}^1\!/_4 = \tfrac{1+\alpha}{4}.$$

By calculating the probability of each outcome in a compound lottery, we can *reduce* it to an equivalent simple lottery, as shown on the right of Figure 1.2.

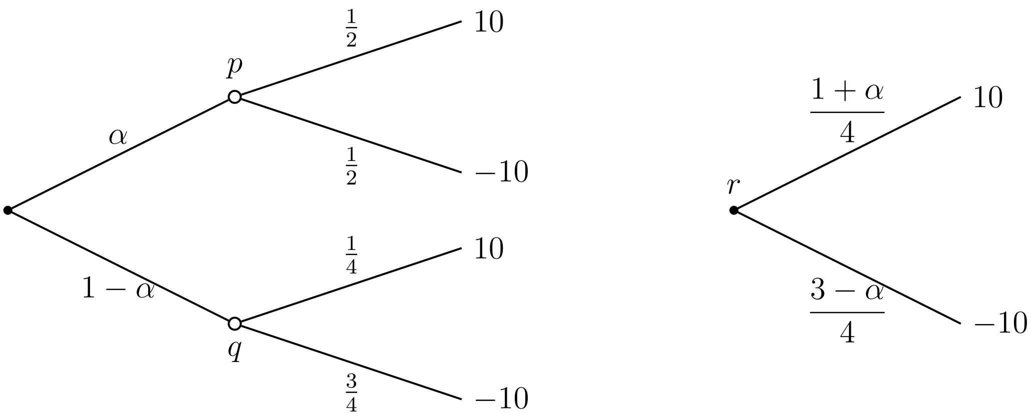


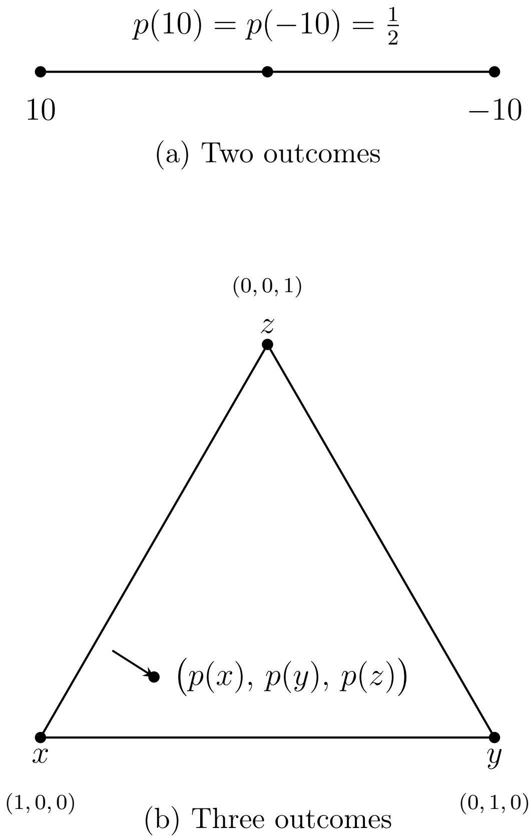
Figure 1.2: Compound lottery (left) and its reduced form (right).

We assume *reduction of compound lotteries*: individuals are indifferent between any compound lottery and its reduced form—that is, any two lotteries that induce the same probabilities over outcomes are treated as equivalent.

Can you think of reasons why someone might *not* be indifferent between a compound lottery and its reduced form? Violations of reduction generate interesting phenomena studied in behavioural economics. See, for example, Segal (1990) and Dillenberger & Raymond (2020).

This lottery *mixing* operation would not make sense with an unstructured set of outcomes. As an illustration, suppose the set of outcomes consists of fruits. We can have an apple or a banana, but there is no fruit that is a mixture of an apple and a banana. Imposing structure on the set of elements to be ranked is one of the key moves in microeconomic theory. In fact, we will later assume that the set of outcomes is $\mathbb{R}$, the set of real numbers representing monetary outcomes, which allows us to say more than we could with a generic set of outcomes.

There is another useful way to represent lotteries graphically. Consider again the coin toss that yields 10 euros with probability $1/2$ and -10 euros with probability $1/2$. We can represent this lottery as the midpoint of the line segment whose endpoints correspond to the degenerate lotteries that yield 10 and -10 with probability 1; see panel (a) of Figure 1.3. More generally, with n possible outcomes we can represent a lottery as a point in an $(n-1)$-dimensional simplex. For example, with three outcomes we can represent lotteries as points in an equilateral triangle, as in panel (b).[3] The vertices of the triangle correspond to degenerate lotteries that yield one outcome with probability 1, while any other point in the triangle represents a lottery that yields each of the three outcomes with some probability. Roughly speaking, the farther a point is from a vertex, the lower the probability of the corresponding outcome. For example, the lottery p in panel (b) yields outcome x with relatively high probability and outcomes y and z with relatively low probabilities.



(a) Two outcomes

(b) Three outcomes

Figure 1.3: Lotteries as points in simplexes: (a) a two–outcome lottery lies on a line segment; (b) with three outcomes, lotteries lie in an equilateral triangle.

For a finite outcome set X, the probability simplex over X is

$$\Delta(X) = \left\{ p\colon X \to [0,1] \;\middle|\; \sum_{x \in X} p(x) = 1 \right\},$$

[3]That's why the Δ notation!

or equivalently,

$$\left\{ \big(p(x_1), \ldots, p(x_n)\big) \in \mathbb{R}^n \;\middle|\; p(x_i) \geq 0, \ \sum_i p(x_i) = 1 \right\}.$$

This set is an $(n-1)$-dimensional simplex whose vertices correspond to the degenerate lotteries (unit vectors), e.g. $(1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)$.

1.2 Preferences over lotteries

Our goal is to understand how individuals choose between lotteries, whether they like or dislike risk, and how we can compare different individuals' attitudes toward risk. To do so, we need a way to express statements such as "an individual weakly prefers lottery p to lottery q". Introduce a binary relation $\succsim$ over $\Delta(X)$, where $p \succsim q$ reads "the individual weakly prefers lottery p to lottery q".[4] Contrary to choice under certainty, we are now comparing lotteries—that is, probability distributions over outcomes—rather than outcomes themselves.

Technically, $\succsim$ is a subset of $\Delta(X) \times \Delta(X)$: a set of ordered pairs of lotteries. For example, if $p, q \in \Delta(X)$, the statement "p is (weakly) preferred to q" is equivalent to $(p, q) \in \succsim$.

Recall that we can define strict preference and indifference in terms of weak preference. We write $p \succ q$, which reads "p is strictly preferred to q", if and only if $p \succsim q$ but not $q \succsim p$; and we write $p \sim q$, which reads "p is indifferent to q", if and only if both $p \succsim q$ and $q \succsim p$.

In principle, we could describe the preference relation $\succsim$ by listing, pair by pair, which lotteries are weakly preferred to which others. However, that would be rather inconvenient as a way to work with preferences. It is more practical to have a function that assigns a number to each lottery, so that we can compare lotteries by comparing their associated numbers. Such a function should "agree" with the preference relation $\succsim$ in the sense that, if p is weakly preferred to q, then the number assigned to p should be at least as large as the number assigned to q. This leads us to the notion of a **utility function** representing preferences.

Definition 1.1. *A utility function $U\colon \Delta(X) \to \mathbb{R}$ **represents** the preference relation $\succsim$ over $\Delta(X)$ if, for all lotteries p, q,*

$$p \succsim q \iff U(p) \geq U(q).$$

[4]Are you curious why we use the symbol $\succsim$ for preferences instead of $\geq$? The historian of economic theory Ivan Boldyrev told me that it originates from Herstein & Milnor (1953), who used it in their classic paper providing an axiomatic characterization of expected utility—which we will encounter soon.

What would be a reasonable utility function representing preferences over lotteries? A natural candidate is the **expected utility** function, defined as follows.

Definition 1.2. *Preferences* $\succsim$ *are represented by an* ***expected utility function*** *if there exists a function* $u\colon X \to \mathbb{R}$ *such that, for all lotteries* p,

$$U(p) = \sum_{x \in X} p(x)\, u(x). \tag{1.1}$$

In other words, an expected utility function assigns to each lottery p the *expected value* of the function u over the outcomes, where the expectation is taken with respect to the probability distribution p. The function u is sometimes called the **Bernoulli utility function**. An expected utility function is *linear in probabilities*; that is, for any lotteries p, q and any $\alpha \in [0, 1]$,

$$U\big(\alpha p + (1-\alpha)q\big) = \alpha U(p) + (1-\alpha)U(q),$$

meaning that the expected utility of a mixture of lotteries is the weighted average of their expected utilities. The other direction of the statement is also true: if U is linear, then it is an expected utility function.[5] Linearity is an extremely convenient property in applications, which partially explains the success of expected utility theory.

Suppose you observe an individual's choices and want to test whether their preferences can be represented by an expected utility function. How might you do that? One idea would be to design a choice task, predict the individual's choices using Equation (1.1), and see whether the predictions are accurate. However, this approach is hard to apply, because to make predictions you would need to assume a specific function u. This is sometimes done—certain functional forms work particularly well—but there is another approach.

We can instead look for *behavioural predictions* that are independent of any specific u; that is, properties of choices that any expected utility maximiser must satisfy. If we can identify such properties, we can design a choice task aimed at testing whether the individual's choices satisfy them. Linearity of expected utility is one such property: it holds regardless of the specific u, and it is the main behavioural prediction of expected utility theory.

In the next lecture, we will examine properties that *fully characterise* preferences representable by an expected utility function. In other words, violating these properties implies that preferences cannot be represented by an expected utility function, while satisfying them implies that they can be represented *only* by an expected utility function. Such characterisations are remarkably powerful. We will also discuss a second point of view on the role of these properties in defining a theory of choice.

[5]You are asked to prove this in Exercise 1.6.

Things to read. See Kreps (1988, pp. 31–33) for a brief, intuitive introduction to the lottery model in this chapter. For a similar treatment in a standard textbook, see Mas-Colell et al. (1995, pp. 168–170).

1.3 Exercises

Exercise 1.1. Can we still represent the set of lotteries and compound lotteries on the simplex if individuals are *not* indifferent between a compound lottery and its reduced form? Why or why not?

Exercise 1.2. Assume there are three outcomes x, y, z. Draw, in the simplex, the set of lotteries that yield each outcome with the same probability and the lottery that yields x with certainty. Now draw the set of all mixtures of these two lotteries. Assume that the individual is indifferent between the lottery yielding each outcome with the same probability, the lottery yielding x with certainty, and any mixture of the two. Which part of the simplex does this indifference "curve" correspond to? Is it really a curve?

Exercise 1.3. Assuming three outcomes x, y, z, draw in the simplex the set of lotteries that yield outcome x with probability at least $1/2$.

Exercise 1.4. In the main text, we assumed that individuals are indifferent between a compound lottery and its reduced form. State this indifference formally as a condition on the preference relation $\succsim$, using the notation introduced above.

Exercise 1.5. Show that preferences represented by an expected utility function satisfy reduction of compound lotteries.

Exercise 1.6. Show that a function U is an expected utility function if and only if it is linear in probabilities.

Exercise 1.7. An individual faces two choice problems. In the first problem, they choose between receiving 50 euros for sure and a lottery that yields 250 euros with probability 0.10, 50 euros with probability 0.89, and 0 euros with probability 0.01. In the second problem, they choose between two lotteries: the first yields 50 euros with probability 0.11 and 0 euros with probability 0.89; the second yields 250 euros with probability 0.10 and 0 euros with probability 0.90. Suppose that in the first problem the individual prefers the sure amount to the lottery, while in the second they prefer the second lottery (yielding 250 euros with probability 0.10) to the first (yielding 50 euros with probability 0.11). Assume the individual prefers having more money. Can these preferences be represented by an expected utility function? Why or why not?

References

Dillenberger, D., & Raymond, C. (2020). *Additive-belief-based preferences* (PIER Working Paper No. 20-020). Philadelphia, PA, USA: Penn Institute for Economic Research (PIER). 4

Herstein, I. N., & Milnor, J. (1953). An axiomatic approach to measurable utility. *Econometrica, Journal of the Econometric Society*, 291–297. 6

Kreps, D. M. (1988). *Notes on the theory of choice.* Westview Press. 8

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory.* Oxford university press New York. 8

Segal, U. (1990). Two-stage lotteries without the reduction axiom. *Econometrica: Journal of the Econometric Society*, 349–377. 4

Lecture 2

Expected utility theory

2.1 Assumptions on preferences

We now impose properties on preferences over lotteries and study their behavioural implications. But first, a brief methodological aside on what we are doing. Before discussing properties of $\succsim$, we should make explicit what the interpretation of $\succsim$ is. Different methodological stances are possible. Is $\succsim$ tracking what an individual has in mind? What he would say if asked? How he chose in the past?

Under *revealed preference theory*, we interpret $\succsim$ as a description of how an individual **chooses**. Therefore, there is no psychological content to $\succsim$. Revealed preference theory has been the standard methodological stance in economics for a long time. But why? Wouldn't it be better to develop a theory that exploits psychological insights?

Revealed preference theory is a methodological stance, not a psychological (or, for that matter, a moral) one. The assumption is not that choices are unrelated to psychological motives, but that we abstract from these motives and look for patterns in choices directly. There is a strong advantage in doing so: psychological motives are hard to observe, while choices can be observed easily. The implication is that a choice theory based on revealed preferences is empirically testable: if we observe choices that violate the theory's assumptions, we can reject it. Therefore, revealed preference theory is **not** a claim about how individuals make choices or about what drives them. On the contrary, it is deliberately silent about these issues.[1] This is often misunderstood: there is a plethora of critics claiming that economics views individuals as cold robots.[2]

Such critics mostly come from behavioural economics, a field that aims to incorporate psychological insights into economic models.[3] Is it therefore impossible to do behavioural economics within the revealed-preference framework? Not at all. Good behavioural theories do what the name suggests: they characterise the *behavioural* content of a theory, so that we, as economists, can understand how individuals behave. Two behavioural theories with different psychological content but that are observationally equivalent—i.e., they make the same predictions about choices—have the same economic implications.[4]

---

[1] If you are interested, see Thoma (2021) for a discussion of the current status of revealed preference theory and Moscati (2025) for the role of psychological narratives in choice theory.

[2] By the way, if you read Asimov's books you know that robots are not cold at all!

[3] See Spiegler (2024) for an account of the motivations of the founding fathers of behavioural economics.

[4] There is a huge debate on this topic. Among many, I suggest reading Gul & Pesendorfer (2008) and the reply by Camerer (2008). A more recent discussion is Spiegler (2019).

An interesting case study is Masatlioglu & Raymond (2016), where the authors show that the famous model by Köszegi & Rabin (2007) is behaviourally equivalent to the intersection of rank-dependent utility and quadratic utility—two older models—despite having a different psychological interpretation. Another example that is quite relevant today is in Eliaz & Spiegler (2006).

In what follows, you can have in mind the interpretation of $\succsim$ that you prefer, but keep in mind that assumptions may have different flavours under different interpretations. Recall that we want to find properties that single out expected utility preferences in Equation (1.1). Therefore, it might be worthwhile to first understand some implications of having expected utility preferences.

Having expected utility preferences over lotteries implies that indifference curves on the simplex are straight lines. That is, say that if $p \sim q$, then, for any $\alpha \in (0, 1)$, it holds that $\alpha p + (1 - \alpha)q \sim p$, as illustrated in Figure 2.1.

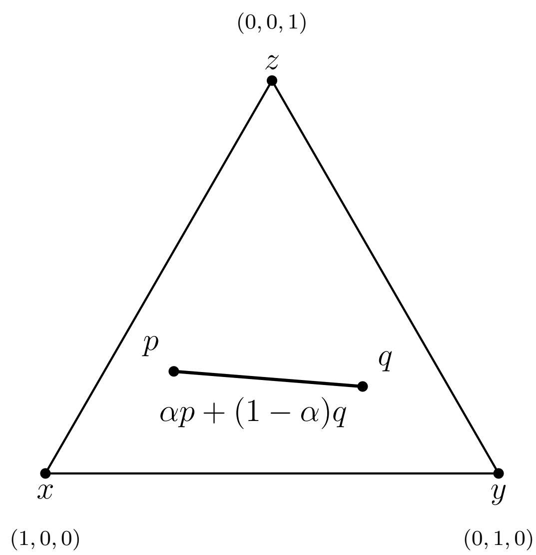


Figure 2.1: If $p \sim q$, then any mixture of p and q is also indifferent to p and q.

Let's show this formally. Assume that $p \sim q$. Then, by the definition of expected utility, we have

$$\sum_{x \in X} p(x)u(x) = \sum_{x \in X} q(x)u(x).$$

Applying expected utility again, for any $\alpha \in (0, 1)$, the utility of the lottery $\alpha p + (1 - \alpha)q$ is

$$
\begin{aligned}
\sum_{x\in X}\big(\alpha p(x)+(1-\alpha)q(x)\big)u(x) &= \sum_{x\in X}\alpha p(x)u(x)+\sum_{x\in X}(1-\alpha)q(x)u(x)\\
&= \alpha\sum_{x\in X}p(x)u(x)+(1-\alpha)\sum_{x\in X}q(x)u(x)\\
&= \alpha\sum_{x\in X}q(x)u(x)+(1-\alpha)\sum_{x\in X}q(x)u(x)\\
&= \sum_{x\in X}q(x)u(x).
\end{aligned}
$$

Indifference curves are also parallel; you are asked to show this in Exercise 2.1. Of course, the fact that indifference curves are straight lines is related to the linearity of expected utility, which in turn follows from a specific axiom, as we will see shortly.

Let's now turn to the properties of $\succsim$ we will consider. First, we assume that preferences form a **weak order**.

Axiom 2.1. *(**Weak order**) Preferences $\succsim$ are complete and transitive.*

Recall that preferences are **complete** if, for any two lotteries p, q, either $p \succsim q$ or $q \succsim p$, or both. They are **transitive** if, for any three lotteries p, q, r, whenever $p \succsim q$ and $q \succsim r$, then $p \succsim r$.

Sometimes Weak order is referred to as **rationality** of preferences (see e.g. Mas-Colell et al. (1995, p. 6)). However, I think this is an unfortunate name. It suggests that it is "irrational" to violate Weak order, but there are reasons why people might have intransitive or incomplete preferences (can you think of any?).

Weak order is a necessary condition for having any utility representation (see Mas-Colell et al. (1995, p. 9)). It is not the core assumption of expected utility theory, but rather one shared by most theories of choice.

Axiom 2.2. *(**Continuity**) For any three lotteries p, q, r, if $p \succ q \succ r$, then there exist $\alpha, \beta \in (0,1)$ such that $\alpha p + (1-\alpha)r \succ q \succ \beta p + (1-\beta)r$.*

Continuity says that there is no lottery p so good that, for $q \succ r$, a small probability β of p and a large probability $1-\beta$ of r is always better than q. Similarly, there is no gamble r so bad that, for $p \succ q$, a large probability α of p and a small probability $1-\alpha$ of r is always worse than q. In essence, this axiom implies that preferences do not have "jumps" when probabilities change slightly—i.e., that preferences are *continuous* in probabilities. Continuity allows us to obtain a continuous utility representation of preferences (see Mas-Colell et al. (1995, p. 47)), but again, it is not the core assumption of expected utility theory—the next one is.

Axiom 2.3. *(**Independence**) For any three lotteries p, q, r and for any $\alpha \in (0, 1)$, we have $p \succsim q$ if and only if $\alpha p + (1-\alpha) r \succsim \alpha q + (1-\alpha) r$.*

In words, Independence says that if p is preferred to q, then mixing both lotteries with any third lottery r, using the same probability $1-\alpha$, does not change their ranking. One way to justify Independence is as follows. Suppose $p \succsim q$. Now consider two compound lotteries obtained by tossing a coin: the first yields p if the coin shows heads and r otherwise; the second yields q if the coin shows heads and r otherwise. Ex ante, one might reason that what happens if the coin shows tails is the same in both compound lotteries, so that part should not matter—while if the coin shows heads, p is preferred to q. Therefore, the first compound lottery should be preferred to the second.

This argument relies on the meaning of the mixing operation within the set of lotteries. By contrast, consider the case of mixing foods. One might prefer pasta to cake, yet mixing both with whipped cream could make the cake better than the pasta.[5] You are asked in Exercise 2.2 to elaborate on the relation between Independence and the linearity of indifference curves.

Before stating Theorem 2.1 in the next section, we need to prove a preliminary result. Lemma 2.1 establishes that, under the assumptions introduced so far, there exist two lotteries that are the best and the worst possible ones.

Lemma 2.1. *Let $\succsim$ satisfy Weak order and Independence. Then there exist two lotteries $\overline{p}$ and $\underline{p}$ such that*

$$\overline{p} \succsim p \succsim \underline{p} \quad \text{for all } p.$$

Proof. The proof proceeds in two steps.

Step 1. By Weak order, the restriction of $\succsim$ to the set of degenerate lotteries $\{\delta_x \in \Delta(X) : \delta_x(x) = 1\}$ is a complete and transitive order on a finite set. Hence there exist outcomes x^*, x_* such that

$$\delta_{x^*} \succsim \delta_x \succsim \delta_{x_*} \qquad \text{for all } x.$$

Fix $\overline{p} := \delta_{x^*}$ and $\underline{p} := \delta_{x_*}$.

Step 2. For any lottery p, let $\text{supp}(p) = \{x \in X : p(x) > 0\}$ and denote its size by $|\text{supp}(p)|$. We prove by induction on $k := |\text{supp}(p)|$ that

$$\overline{p} \succsim p \succsim \underline{p}.$$

Base case. If $k = 1$, then $p = \delta_x$ for some x, and the claim follows from **Step 1**.

Inductive step. Assume the statement holds for all lotteries whose support size is at most $k-1$. Let p have support size $k \geq 2$. Pick any $x \in \text{supp}(p)$ and write

$$p = \alpha\, \delta_x + (1-\alpha)\, q, \qquad \alpha := p(x) \in (0, 1),$$

[5] Feel no shame if you are unconvinced by this example.

where q is the renormalized remainder, defined by

$$q(y) = \begin{cases} \dfrac{p(y)}{1-\alpha} & \text{if } y \neq x, \\ 0 & \text{if } y = x. \end{cases}$$

Then $q \in \Delta(X)$ and $|\text{supp}(q)| \leq k - 1$.

By the inductive hypothesis, $\overline{p} \succsim q$; and by **Step 1**, $\overline{p} \succsim \delta_x$. We apply Independence twice. From $\overline{p} \succsim q$, mix with $\overline{p}$:

$$\overline{p} = (1-\alpha)\,\overline{p} + \alpha\,\overline{p} \;\succsim\; (1-\alpha)\,q + \alpha\,\overline{p} = \alpha\,\overline{p} + (1-\alpha)\,q.$$

From $\overline{p} \succsim \delta_x$, mix with q:

$$\alpha\,\overline{p} + (1-\alpha)\,q \;\succsim\; \alpha\,\delta_x + (1-\alpha)\,q = p.$$

By transitivity,

$$\overline{p} \succsim p.$$

A symmetric argument yields $p \succsim \underline{p}$. Indeed, by the inductive hypothesis $q \succsim \underline{p}$ and by **Step 1** $\delta_x \succsim \underline{p}$. Using Independence with the same reasoning gives

$$p = \alpha\,\delta_x + (1-\alpha)\,q \;\succsim\; \alpha\,\underline{p} + (1-\alpha)\,q \;\succsim\; \underline{p}.$$

Therefore, $\overline{p} \succsim p \succsim \underline{p}$ for all lotteries with support size k, completing the induction. The fixed degenerate lotteries $\overline{p} = \delta_{x^*}$ and $\underline{p} = \delta_{x_*}$ bound every p, as claimed. □

2.2 Expected utility representation

We are ready to state and prove the theorem relating the properties of preferences over lotteries to the expected utility functional form.

Theorem 2.1. *Preferences over lotteries $\succsim$ satisfy Weak order, Continuity, and Independence if and only if there exists a utility function $u\colon X \to \mathbb{R}$ such that*

$$p \succsim q \text{ if and only if } \sum_{x\in X} p(x)\,u(x) \geq \sum_{x \in X} q(x)\,u(x) \quad \text{for all } p, q. \tag{2.1}$$

The proof essentially follows Mas-Colell et al. (1995, pp. 176–178), complemented by intuition and figures.

Proof. We proceed by steps.

Step 1. If $p \succsim q$, then $p \succsim \alpha p + (1-\alpha)q \succsim q$ for any $\alpha \in (0,1)$.

The intuition behind this step is simple: if p is better than q, then any mixture of the two is worse than p and better than q. Figure 2.2 illustrates the idea.

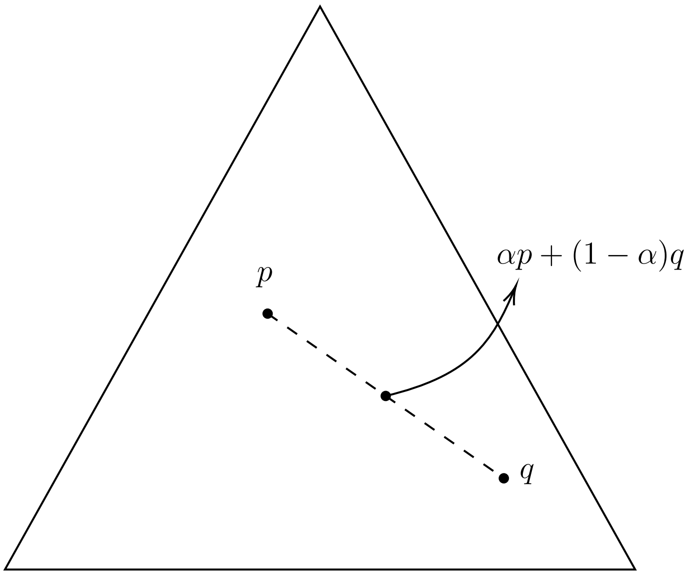


Figure 2.2: Step 1.

This follows from Independence.

$$p \succsim q \implies (1-\alpha)\,p + \alpha\,p \succsim (1-\alpha)\,q + \alpha\,p \implies p \succsim \alpha p + (1-\alpha)q. \tag{2.2}$$

$$p \succsim q \implies \alpha p + (1-\alpha)q \succsim \alpha q + (1-\alpha)q \implies \alpha p + (1-\alpha)q \succsim q. \tag{2.3}$$

The conclusion follows from Equations (2.2) and (2.3).

Step 2. $\beta > \alpha$ if and only if $\beta\overline{p} + (1-\beta)\,\underline{p} \succ \alpha\overline{p} + (1-\alpha)\,\underline{p}$, where $\overline{p}$ and $\underline{p}$ are the best and worst lotteries identified in Lemma 2.1.

The idea of this step is as follows. From **Step 1**, we know that a mixture of p and q, where $p \succsim q$, is worse than p and better than q. Now, since $\overline{p}$ is the best lottery available, we have $\overline{p} \succ \alpha\overline{p} + (1-\alpha)\,\underline{p}$. We want to show that $\beta\overline{p} + (1-\beta)\,\underline{p}$ can be written as a mixture of $\overline{p}$ and $\alpha\overline{p} + (1-\alpha)\,\underline{p}$; therefore, by **Step 1**, it must be preferred to $\alpha\overline{p} + (1-\alpha)\,\underline{p}$. The idea is illustrated in Figure 2.3.

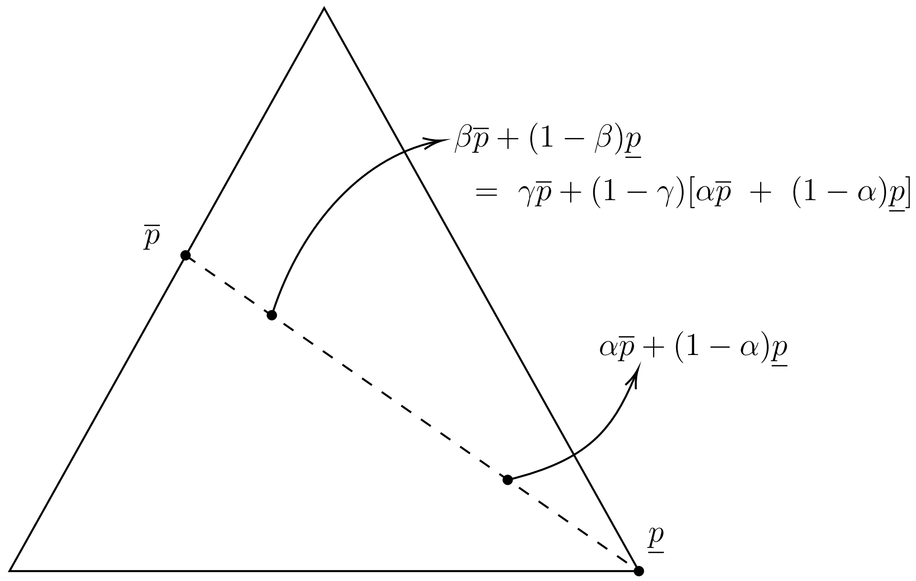


Figure 2.3: Step 2.

We want to express $\beta\overline{p} + (1-\beta)\,\underline{p}$ as a mixture of $\overline{p}$ and $\alpha\overline{p} + (1-\alpha)\,\underline{p}$. That is, we look for some $\gamma \in (0,1)$ such that

$$\beta\overline{p} + (1-\beta)\,\underline{p} = \gamma\overline{p} + (1-\gamma)\big[\alpha\overline{p} + (1-\alpha)\,\underline{p}\big].$$

A short calculation shows that $\gamma = \dfrac{\beta-\alpha}{1-\alpha}$. By **Step 1** we know that $\overline{p} \succ \alpha\overline{p} + (1-\alpha)\,\underline{p}$; therefore,

$$\gamma\overline{p} + (1-\gamma)\big[\alpha\overline{p} + (1-\alpha)\,\underline{p}\big] \succ \alpha\overline{p} + (1-\alpha)\,\underline{p}.$$

Since $\beta\overline{p} + (1-\beta)\,\underline{p} = \gamma\overline{p} + (1-\gamma)\big[\alpha\overline{p} + (1-\alpha)\,\underline{p}\big]$, the conclusion follows.

Up to this point we have proved that if $\beta > \alpha$, then $\beta\overline{p} + (1-\beta)\,\underline{p} \succ \alpha\overline{p} + (1-\alpha)\,\underline{p}$. But the statement says "if and only if", so we must also show the converse: if $\alpha \geq \beta$, then it cannot be that $\beta\overline{p} + (1-\beta)\,\underline{p} \succ \alpha\overline{p} + (1-\alpha)\,\underline{p}$. When $\beta = \alpha$, the two lotteries coincide and are therefore indifferent. The relevant case is $\alpha > \beta$. By the argument above, $\alpha\overline{p} + (1-\alpha)\,\underline{p} \succ \beta\overline{p} + (1-\beta)\,\underline{p}$, and that completes the proof of this step.

Step 3. [6] For any p, there exists a unique $\alpha_p \in [0,1]$ such that $p \sim \alpha_p\overline{p} + (1-\alpha_p)\,\underline{p}$.

We can derive this step as a consequence of the previous ones together with Continuity. This step involves some algebra, but you can get intuition from Figure 2.4.

[6] In this step we use a proof by contradiction. Before diving in, make sure you are familiar with the logic of such proofs.

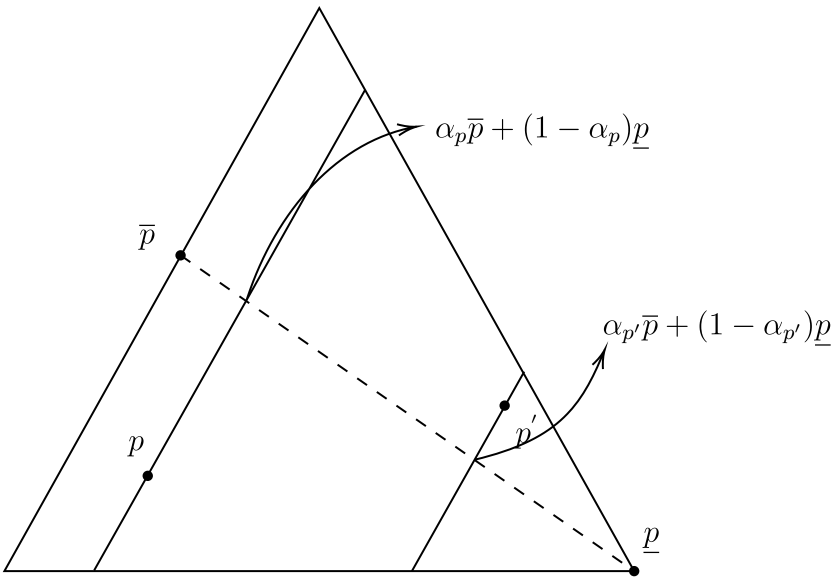


Figure 2.4: Step 3.

First, notice that if α_p exists, it must be unique. Suppose there are two such numbers, α_p and α_p', with $\alpha_p > \alpha_p'$. Then, by **Step 2**, $\alpha_p\overline{p}+(1-\alpha_p)\,\underline{p} \succ \alpha_p'\overline{p}+(1-\alpha_p')\,\underline{p}$, contradicting indifference to p.

Now we need to show that such an α_p exists. If $\overline{p} \sim p$, then $\alpha_p = 1$ works; if $\underline{p} \sim p$, then $\alpha_p = 0$ works. The interesting case is when $\overline{p} \succ p \succ \underline{p}$.

Define

$$\alpha_p \;=\; \sup\big\{\, \alpha \in [0,1] : p \succsim \alpha\overline{p} + (1-\alpha)\,\underline{p} \,\big\}. \tag{2.4}$$

Since $\alpha = 0$ belongs to this set, the supremum is well defined and the set is non-empty.

We now establish two auxiliary claims. The first is

$$\text{If } 1 \geq \alpha > \alpha_p, \text{ then } \alpha\overline{p} + (1-\alpha)\,\underline{p} \succ p. \tag{2.5}$$

Indeed, if $p \succsim \alpha\overline{p}+(1-\alpha)\,\underline{p}$ held for such α, then α_p would not satisfy Equation (2.4). Moreover,

$$\text{If } 0 \leq \alpha < \alpha_p, \text{ then } p \succ \alpha\overline{p} + (1-\alpha)\,\underline{p}. \tag{2.6}$$

The reasoning is as follows. By the definition of α_p, there exists some α' such that $\alpha < \alpha' \leq \alpha_p$ and $p \succsim \alpha'\overline{p} + (1-\alpha')\,\underline{p}$. Since $\alpha < \alpha'$, **Step 2** implies that

$$p \succ \alpha'\overline{p} + (1-\alpha')\,\underline{p} \succ \alpha\overline{p} + (1-\alpha)\,\underline{p}.$$

Now, there are three possibilities to consider: $\alpha_p\overline{p}+(1-\alpha_p)\,\underline{p} \succ p$, $p \succ \alpha_p\overline{p}+(1-\alpha_p)\,\underline{p}$, or indifference between them.

If $\alpha_p\overline{p} + (1-\alpha_p)\,\underline{p} \succ p$, then by Continuity there exists $\beta \in (0,1)$ such that

$$\beta\big(\alpha_p\overline{p} + (1-\alpha_p)\,\underline{p}\big) + (1-\beta)\,\underline{p} \succ p.$$

Notice that

$$\begin{aligned}\beta\big(\alpha_p\overline{p} + (1-\alpha_p)\,\underline{p}\big) + (1-\beta)\,\underline{p} &= \beta\alpha_p\,\overline{p} + \beta(1-\alpha_p)\,\underline{p} + (1-\beta)\,\underline{p}\\ &= \beta\alpha_p\,\overline{p} + \big[\beta(1-\alpha_p) + (1-\beta)\big]\,\underline{p}\\ &= \beta\alpha_p\,\overline{p} + \big(1-\beta\alpha_p\big)\,\underline{p} \;\succ\; p.\end{aligned}$$

Since $\beta\alpha_p < \alpha_p$, by Equation (2.6) we must have $p \succ \beta\alpha_p\,\overline{p} + (1-\beta\alpha_p)\,\underline{p}$, which is a contradiction.

If instead $p \succ \alpha_p\overline{p} + (1-\alpha_p)\,\underline{p}$, then by Continuity there exists $\beta \in (0,1)$ such that

$$\begin{aligned}p &\succ \beta\big(\alpha_p\overline{p} + (1-\alpha_p)\,\underline{p}\big) + (1-\beta)\,\overline{p}\\ &= \big[\beta\alpha_p + (1-\beta)\big]\,\overline{p} + \beta(1-\alpha_p)\,\underline{p}\\ &= \big(1-\beta(1-\alpha_p)\big)\,\overline{p} + \beta(1-\alpha_p)\,\underline{p}.\end{aligned}$$

Since $1-\beta(1-\alpha_p) > \alpha_p$, by Equation (2.5) we must have

$$\big(1-\beta(1-\alpha_p)\big)\,\overline{p} + \beta(1-\alpha_p)\,\underline{p} \succ p,$$

which is again a contradiction.

Step 4. Define a utility function $U\colon \Delta(X) \to \mathbb{R}$ that assigns to each lottery a number representing its utility, defined by $U(p) = \alpha_p$. This function represents preferences $\succsim$.

Take two lotteries p and p'. By **Step 3**, there exist unique α_p and $\alpha_{p'}$ such that

$$p \sim \alpha_p\overline{p} + (1-\alpha_p)\,\underline{p}, \quad p' \sim \alpha_{p'}\overline{p} + (1-\alpha_{p'})\,\underline{p}.$$

Therefore,

$$p \succsim p' \text{ if and only if } \alpha_p\overline{p} + (1-\alpha_p)\,\underline{p} \succsim \alpha_{p'}\overline{p} + (1-\alpha_{p'})\,\underline{p}.$$

By **Step 2**,

$$\alpha_p\overline{p} + (1-\alpha_p)\,\underline{p} \succsim \alpha_{p'}\overline{p} + (1-\alpha_{p'})\,\underline{p} \text{ if and only if } \alpha_p \geq \alpha_{p'}.$$

The last condition holds if and only if $U(p) \geq U(p')$, which proves the claim.

Step 5. The function U is linear and therefore, by Exercise 1.6, has the expected utility form.

From the previous steps we know that, for any lottery p, there is a unique number $U(p) \in [0,1]$ such that

$$p \sim U(p)\,\overline{p} + \big(1-U(p)\big)\,\underline{p}, \qquad p' \sim U(p')\,\overline{p} + \big(1-U(p')\big)\,\underline{p}.$$

Applying Independence, we get

$$\begin{aligned}\beta p + (1-\beta)p' &\sim \beta\big[U(p)\,\overline{p} + \big(1-U(p)\big)\,\underline{p}\big] + (1-\beta)p'\\ &\sim \beta\big[U(p)\,\overline{p} + \big(1-U(p)\big)\,\underline{p}\big] + (1-\beta)\big[U(p')\,\overline{p} + \big(1-U(p')\big)\,\underline{p}\big]\\ &= \big[\beta U(p) + (1-\beta)U(p')\big]\,\overline{p} + \Big(1-\big[\beta U(p) + (1-\beta)U(p')\big]\Big)\,\underline{p}.\end{aligned}$$

Let $\gamma := \beta U(p) + (1-\beta)U(p')$. By **Step 4**, for the lottery $\beta p + (1-\beta)p'$ there is a *unique* number γ such that $\beta p + (1-\beta)p' \sim \gamma\,\overline{p} + (1-\gamma)\,\underline{p}$. Therefore,

$$U(\beta p + (1-\beta)p') = \beta U(p) + (1-\beta)U(p').$$

$\square$

Recall that a functional representation of preferences need not be unique: multiple functions can represent the same preferences. However, for expected utility representations we have a very specific characterization of all possible representations, as stated in the following corollary.

Corollary 2.1. *Suppose U is an expected utility representation of $\succsim$. Then $\widetilde{U}\colon \Delta(X) \to \mathbb{R}$ is another expected utility representation of $\succsim$ if and only if there exist $\beta > 0$ and $\gamma \in \mathbb{R}$ such that*

$$\widetilde{U}(p) \;=\; \beta\, U(p) + \gamma \qquad \textit{for all } p. \tag{2.7}$$

Proof. First, suppose that Equation (2.7) holds. Then $\widetilde{U}$ is an expected utility representation of $\succsim$. Assume $\widetilde{U} = \beta U + \gamma$ with $\beta > 0$. Then

$$\widetilde{U}(p) \;=\; \beta \sum_x p(x)u(x) + \gamma \;=\; \sum_x p(x)\big[\beta u(x) + \gamma\big].$$

Hence, $\widetilde{U}$ has the expected utility form with $\widetilde{u}(x) := \beta u(x) + \gamma$. Since $\beta > 0$, it follows that $p \succsim q \iff U(p) \geq U(q) \iff \widetilde{U}(p) \geq \widetilde{U}(q)$.

Second, suppose that U and $\widetilde{U}$ are both expected utility representations of the same $\succsim$. Then they must be related by an affine transformation as in Equation (2.7). By Lemma 2.1, there exist $\overline{p}, \underline{p}$ such that $\overline{p} \succ \underline{p}$ and $\overline{p} \succsim p \succsim \underline{p}$ for all p. For any lottery p, define $\alpha_p \in [0,1]$ by

$$U(p) \;=\; \alpha_p U(\overline{p}) + (1-\alpha_p)U(\underline{p}), \quad \text{so that} \quad \alpha_p \;=\; \frac{U(p) - U(\underline{p})}{U(\overline{p}) - U(\underline{p})}.$$

Applying the same construction to $\widetilde{U}$ and using the *same* α_p (since both functions represent the *same* preferences), we obtain

$$\widetilde{U}(p) \;=\; \alpha_p \widetilde{U}(\overline{p}) + (1-\alpha_p)\,\widetilde{U}(\underline{p}) \;=\; \widetilde{U}(\underline{p}) \;+\; \frac{\widetilde{U}(\overline{p}) - \widetilde{U}(\underline{p})}{U(\overline{p}) - U(\underline{p})}\big[U(p) - U(\underline{p})\big].$$

Rearranging yields the affine relation

$$\widetilde{U}(p) \;=\; \beta\, U(p) + \gamma \quad \text{for all } p,$$

where

$$\beta \;:=\; \frac{\widetilde{U}(\overline{p}) - \widetilde{U}(\underline{p})}{U(\overline{p}) - U(\underline{p})} \;>\; 0, \qquad \gamma \;:=\; \widetilde{U}(\underline{p}) - \beta\, U(\underline{p}).$$

Positivity of β follows because $\overline{p} \succ \underline{p}$ implies $U(\overline{p}) > U(\underline{p})$ and $\widetilde{U}(\overline{p}) > \widetilde{U}(\underline{p})$. $\square$

In Exercise 2.7, you are asked to elaborate on the significance of Corollary 2.1 for interpreting utility numbers as *cardinal* representations of preferences.[7]

Things to read. For a textbook treatment of the material covered in this lecture, see Mas-Colell et al. (1995, pp. 170–178). Theorem 2.1 is known as the *von Neumann–Morgenstern representation theorem*, after von Neumann & Morgenstern (2007), and it has enormous historical importance. An excellent discussion of the historical context and significance of expected utility theory can be found in Moscati (2018).

2.3 Exercises

Exercise 2.1. Show that if preferences are represented by an expected utility function, then indifference curves in the triangle are parallel lines.

Exercise 2.2. Explain why the fact that indifference curves are straight and parallel lines follows from Independence.

Exercise 2.3. Prove the direction of Theorem 2.1 that we did not prove in class. Show that if U represents $\succsim$, then $\succsim$ satisfies Weak order, Continuity, and Independence. (It is not difficult, I promise!)

Exercise 2.4. Revisit the choice problem in Exercise 1.7. Show that the preferences exhibited there do not satisfy Independence.

Exercise 2.5. Consider the **Betweenness Axiom** introduced by Dekel (1986): for all lotteries p, q and $\alpha \in [0,1]$, if $p \sim q$, then $\alpha p + (1-\alpha)q \sim p$. Show that Independence implies Betweenness, but Betweenness does not imply Independence. (Hint: for the second part, construct a preference relation that satisfies Betweenness but not Independence.) Are indifference curves still linear under Betweenness? Are they parallel?

Exercise 2.6. Show that the choices of the individual in Exercise 1.7 are compatible with Betweenness. (Hint: drawing a picture might help.)

Exercise 2.7. Explain why Corollary 2.1 allows us to make statements such as "the utility of winning the lottery is twice the utility of not winning."

References

Camerer, C. (2008). The case for mindful economics. In *The Foundations of Positive and Normative Economics: A Hand Book* (pp. 43–69). Oxford University Press. 10

[7]If you are interested in this aspect, you may read von Neumann & Morgenstern (2007, ch. 3). An illuminating discussion of measurement in the social sciences is in Krantz et al. (1971, ch. 1).

Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic theory*, *40*(2), 304–318. 20

Eliaz, K., & Spiegler, R. (2006). Can anticipatory feelings explain anomalous choices of information sources? *Games and Economic Behavior*, *56*(1), 87–104. 11

Gul, F., & Pesendorfer, W. (2008). The case for mindless economics. In *The Foundations of Positive and Normative Economics: A Hand Book* (pp. 3–42). Oxford University Press. 10

Köszegi, B., & Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, *97*(4), 1047–1073. 11

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement, vol. I: Additive and polynomial representations.* New York: Academic Press. 20

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory.* Oxford university press New York. 12, 14, 20

Masatlioglu, Y., & Raymond, C. (2016). A behavioral analysis of stochastic reference dependence. *American Economic Review*, *106*(9), 2760–2782. 11

Moscati, I. (2018). *Measuring utility: From the marginal revolution to behavioral economics.* Oxford Studies in History of Economics. 20

Moscati, I. (2025). Psychological narratives in decision theory: What they are and what they are good for. *Journal of Economic Methodology*, 1–22. 10

Spiegler, R. (2019). Behavioral economics and the atheoretical style. *American Economic Journal: Microeconomics*, *11*(2), 173–194. 10

Spiegler, R. (2024). *The Curious Culture of Economic Theory.* The MIT Press. 10

Thoma, J. (2021). In defence of revealed preference theory. *Economics & Philosophy*, *37*(2), 163–187. 10

von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior* (60th Anniversary Commemorative ed.). Princeton, NJ: Princeton University Press. 20

Lecture 3

Money lotteries

3.1 Structuring the set of outcomes

In the previous section, we studied preferences with the expected utility form over lotteries on a *finite* outcome set X. We now study a setting where the outcome set is the set of real numbers $\mathbb{R}$, representing monetary outcomes. This setting is particularly important in economics and finance, as it allows us to model decisions such as investments, insurance, and consumption.

You may wonder whether a form of Theorem 2.1 extends to this setting. The answer is yes, see Kreps (1988, pp. 59–78) or Fishburn (1970, ch. 10).

Since the outcome set is now infinite, we should be careful about how we define lotteries. We will use cumulative distribution functions (CDFs) to represent lotteries over monetary outcomes. A CDF $F\colon \mathbb{R} \to [0,1]$ maps each monetary outcome x to the probability that the outcome is less than or equal to x. It satisfies:

- F is nondecreasing, i.e. if $x \leq y$, then $F(x) \leq F(y)$.
- $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to+\infty} F(x) = 1$.
- F is right–continuous, i.e. for every $x \in \mathbb{R}$, $\lim_{y\downarrow x} F(y) = F(x)$.[1]

Example 3.1. Consider a lottery that pays 1 dollar with probability $\frac{1}{4}$, 4 dollars with probability $\frac{1}{2}$, and 6 dollars with probability $\frac{1}{4}$. The corresponding CDF F is

$$F(x) = \begin{cases} 0 & \text{if } x < 1, \\ \frac{1}{4} & \text{if } 1 \leq x < 4, \\ \frac{3}{4} & \text{if } 4 \leq x < 6, \\ 1 & \text{if } x \geq 6, \end{cases}$$

and it is represented in Figure 3.1.

[1]The notation $y \downarrow x$ means that y approaches x from above.

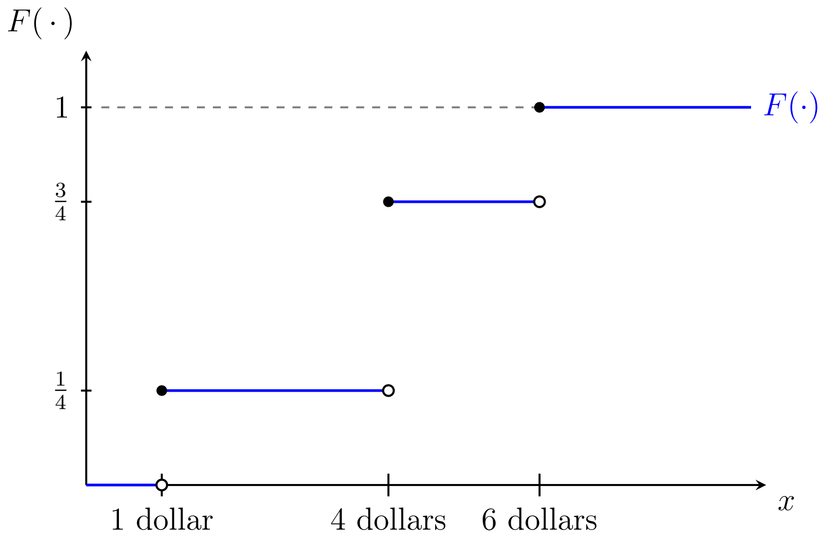


Figure 3.1: Cumulative distribution function (CDF) representing a lottery over monetary outcomes.

■

A second approach to studying lotteries over $\mathbb{R}$ is via *simple* probability distributions, i.e. probability distributions that assign positive probability to only a finite number of outcomes.

Notice that mixtures of CDFs are also CDFs, so we can employ the same mixture operation defined in Section 1.1. In particular, given two CDFs F and G, and $\alpha \in [0,1]$, the mixture $H = \alpha F + (1-\alpha)G$ is also a CDF, where $H(x) = \alpha F(x) + (1-\alpha)G(x)$ for all x.

We now define preferences $\succsim$ over the set of CDFs on $\mathbb{R}$ that have the expected utility form. The idea is the same as before: we weight the utility of each monetary outcome by its probability and sum these weighted utilities to obtain the expected utility of the lottery. A preference relation $\succsim$ over the set of CDFs has the expected utility form if there exists a utility function $u\colon \mathbb{R} \to \mathbb{R}$ such that for any two CDFs F and G,

$$F \succsim G \quad \Longleftrightarrow \quad \int u(x)\,\mathrm{d}F(x) \;\geq\; \int u(x)\,\mathrm{d}G(x).$$

Earlier, the Bernoulli utility was defined on a finite X as $u\colon X \to \mathbb{R}$; now the outcome set is $\mathbb{R}$, hence the domain differs. This lets us impose properties of u that are specific to monetary outcomes. From now on we assume the following two.

Definition 3.1. *The utility function u is* ***increasing*** *if for any x, y with $x > y$, we have $u(x) > u(y)$.*

Definition 3.1 captures the idea that more money is preferred to less. When the outcome set was a generic X, the inequality $x > y$ had no meaning.[2]

[2] For instance, if x is an apple and y is a banana, what would $x > y$ even mean?

Definition 3.2. *The utility function* $u\colon \mathbb{R} \to \mathbb{R}$ *is* ***continuous*** *if for any* x *and any* $\varepsilon > 0$*, there exists* $\delta > 0$ *such that for all* y *with* $|x - y| < \delta$*, we have* $|u(x) - u(y)| < \varepsilon$*.*

Definition 3.2 ensures that small changes in money lead to small changes in the Bernoulli utility u. This could not be stated with a generic outcome set, where expressions like $x - y$ are undefined.

Definition 3.2 is continuity in *money*. What about continuity in *probabilities*?

3.2 Risk aversion

We now have the tools to define and discuss *risk aversion*. Defining this concept allows us to answer questions such as: does an individual dislike risk? how much? As we will see in the next Lecture, the answer has important implications for economic behaviour.

The definition of risk aversion is quite intuitive. Consider the following choice: receive 5 euros for sure, or take a lottery that pays 0 euros with probability 0.5 and 10 euros with probability 0.5. Both options have the same expected monetary value, 5 euros. If the individual prefers the certain 5 to the lottery, he *dislikes* risk—he prefers getting the mean outcome for sure rather than facing uncertainty. If instead he prefers the lottery, he *likes* risk—he is willing to face uncertainty for the chance of a higher payoff.

Did you notice what we just did? We needed to develop a definition of an intuitive, but ex-ante vague concept, risk aversion. We did it by developing a simple thought experiment that "keeps everything fixed" except for the presence of risk. These thought experiments are very useful to develop effective definitions.

For a general lottery, we say an individual is risk averse if he prefers the certain amount equal to the lottery's expected value to the lottery itself. For a CDF F, the expected value is

$$\int x \, \mathrm{d}F(x). \tag{3.1}$$

Evaluating money through u, the certain amount equal to that expected value yields utility

$$u\left(\int x \, \mathrm{d}F(x)\right), \tag{3.2}$$

whereas the lottery yields expected utility

$$\int u(x) \, \mathrm{d}F(x). \tag{3.3}$$

Definition 3.3. *An individual with expected utility preferences and Bernoulli utility u is* ***risk averse*** *if for each CDF F,*

$$u\Big(\int x \,\mathrm{d}F(x)\Big) \;\geq\; \int u(x)\,\mathrm{d}F(x). \tag{3.4}$$

If u is not increasing (Definition 3.1), Definition 3.3 loses its intended meaning.

Inequality (3.4) is precisely Jensen's inequality and is equivalent to the *concavity* of u. Thus, the intuitive notion of risk aversion is equivalent to concavity of u. If u is twice differentiable, concavity means $u''(x) \leq 0$ for all x; see Figure 3.2.

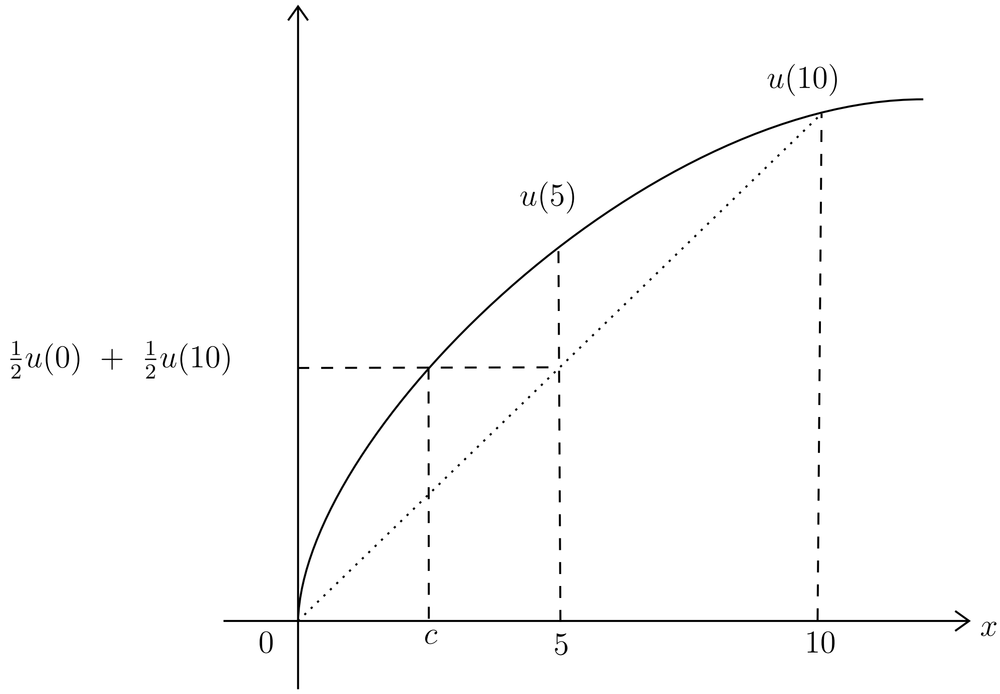


Figure 3.2: Example of a concave u.

Analogously, an individual is *risk loving* if (3.4) is reversed, and *risk neutral* if it holds with equality.

There are other, equivalent ways to define risk aversion. This is good news: it means the definition is robust. One convenient route is via the **certainty equivalent**—the sure amount of money that makes the individual indifferent to the lottery.

Definition 3.4. *The* ***certainty equivalent*** *of a lottery with CDF F for an individual with utility u is the number $c(F, u)$ solving*

$$u\big(c(F,u)\big) = \int u(x)\,\mathrm{d}F(x). \tag{3.5}$$

As an illustration, the certainty equivalent of the lottery paying 0 with probability 1/2 and 10 with probability 1/2 in Figure 3.2 is c.

Intuitively, if an individual is risk averse, his certainty equivalent must be less than the expected value of the lottery, as he prefers receiving the expected value for sure rather than facing the lottery. To capture this intuition we can define the **risk premium** of a lottery as the difference between the expected value of the lottery and its certainty equivalent.

Definition 3.5. *The* ***risk premium*** *of a lottery with CDF F for utility u is*

$$\pi(F, u) = \int x \, \mathrm{d}F(x) - c(F, u). \tag{3.6}$$

You will show in Exercise 3.2 that an individual is risk averse if and only if the risk premium is nonnegative for every lottery.

We now have a notion of risk aversion, but not a way to compare the risk attitudes of two individuals. Again, we start from intuition, how could we compare two risk aversion individuals? The risk premium might be a starting point, the higher the risk premium, the more risk averse the individual, as he requires a lower certainty equivalent to face the lottery. Consider two individuals with utility function u and v. If for each lottery F, the risk premium of the first individual is higher than that of the second, i.e. $\pi(F, u) \geq \pi(F, v)$, we can say that the first individual is more risk averse than the second. However, such condition boils down to comparing certainty equivalents:

$$\pi(F, u) \geq \pi(F, v) \quad \Longleftrightarrow \quad c(F, u) \leq c(F, v),$$

you should show this. We therefore have the following definition.

Definition 3.6. *An individual with utility u is* ***more risk averse*** *than one with utility v if, for every lottery F,*

$$c(F, u) \leq c(F, v).$$

We now want to develop a measure of risk aversion that is related to the rate at which the certainty equivalent changes as we change the lottery. Consider a lottery over monetary outcomes that pays $x + \varepsilon$ with probability $1/2$ and $x - \varepsilon$ with probability $1/2$, call it F_ε. By Definition 3.4

$$u\big(c(F_\varepsilon, u)\big) = \tfrac{1}{2}u(x + \varepsilon) + \tfrac{1}{2}u(x - \varepsilon). \tag{3.7}$$

Since both sides of Equation (3.7) are twice differentiable in ε and $u'(c(F_\varepsilon, u)) > 0$ since u is increasing, the implicit function theorem implies that $c(F_\varepsilon, u)$ is twice differentiable in a neighborhood of 0. Differentiating (3.7) with respect to ε gives

$$u'\big(c(F_\varepsilon, u)\big)\, c'(\varepsilon) = \tfrac{1}{2}u'(x + \varepsilon) - \tfrac{1}{2}u'(x - \varepsilon).$$

Evaluating at $\varepsilon = 0$,

$$u'(x)\, c'(0) = 0 \quad \Longrightarrow \quad c'(0) = 0, \qquad c(0) = x.$$

Differentiating again with respect to ε,

$$u''\big(c(F_\varepsilon, u)\big)\big(c'(\varepsilon)\big)^2 + u'\big(c(F_\varepsilon, u)\big)c''(\varepsilon) = \tfrac{1}{2}u''(x+\varepsilon) + \tfrac{1}{2}u''(x-\varepsilon).$$

Evaluating at $\varepsilon = 0$ and using $c'(0) = 0$ and $c(0) = x$, we obtain

$$u'(x)c''(0) = u''(x) \quad \Longrightarrow \quad c''(0) = \frac{u''(x)}{u'(x)}.$$

The ratio between the second and first derivative of the utility function is the **Arrow-Pratt coefficient of absolute risk aversion**. It is not by chance that it appears here. As we noticed already, risk aversion is related to the concavity of the utility function, which is captured by its second derivative. In principle we could use the second derivative alone as a measure of risk aversion, but this would not be satisfactory, as multiplying the utility function by a positive constant would change the second derivative but not risk aversion. Dividing the second derivative by the first derivative solves this problem, as multiplying the utility function by a positive constant multiplies both derivatives by the same constant, leaving their ratio unchanged.

Definition 3.7. *The **Arrow-Pratt coefficient of absolute risk aversion** for an individual with utility function u at outcome x is*

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

Hence, we just showed that the limit of the second derivative of the certainty equivalent as $\varepsilon \to 0$ is exactly $-r(x, u)$. You should notice that, if u is increasing and concave, then $r(x, u)$ is positive. In the exercises, you are asked to show the following equivalence between certainty equivalents and the Arrow-Pratt coefficient.

Proposition 3.1. *An individual with utility u is **more risk averse** than an individual with utility v if and only if, for each x,*

$$r(x, u) \;\geq\; r(x, v).$$

You should notice that, "more risk averse then" is a partial order on the set of utility functions, it is not complete. There might be two utility functions u and v such that neither u is more risk averse than v, nor v is more risk averse than u. This happens when the Arrow-Pratt coefficients cross, i.e. there exist x and y such that $r(x, u) > r(x, v)$ and $r(y, u) < r(y, v)$.

Things to read. This section mostly draws from Mas-Colell et al. (1995, ch. 6.C). Alternatives treatments can be found in Kreps (1988, ch. 6) and Kreps (2013, ch. 6).

3.3 Exercises

Exercise 3.1. Check that the CDF in Figure 3.1 satisfies the three properties of a CDF.

Exercise 3.2. Show that, if an individual with expected utility preferences and utility u is risk averse, his risk premium is nonnegative for each lottery.

Exercise 3.3. We noted that risk aversion is linked to concavity. Can you define "more risk averse than" directly via concavity? Define when a function is "more concave than" another, and show that your notion is equivalent to Definition 3.6. (Hint: it is easiest to go via the Arrow–Pratt coefficient.)

Exercise 3.4. Prove Proposition 3.1. (If you are stuck, check exercises 6.C.6 and 6.C.7 in Mas-Colell et al. (1995).)

Exercise 3.5. A specific utility function that is often used in economics and finance is the **exponential utility function**, defined as $u(x) = 1 - e^{-\alpha x}$, where $\alpha > 0$ is a parameter. Such function has an interesting property related to how it handles risk. Can you find it? Can you elaborate on what this property implies for risk taking at different wealth levels?

References

Fishburn, P. C. (1970). *Utility theory for decision making.* New York: Wiley. 22

Kreps, D. M. (1988). *Notes on the theory of choice.* Westview Press. 22, 27

Kreps, D. M. (2013). *Microeconomic foundations* (Vol. 1). Princeton university press. 27

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory.* Oxford university press New York. 27, 28

Lecture 4

Stochastic dominance and applications

4.1 Stochastic dominance

We just developed an analysis of properties of expected utility preferences. However, we have not yet discussed relevant properties of lotteries, which is what we do now. As an example, one might want to have a language to say that a "lottery yields higher returns than another one".

A simple way of capturing this idea is the following: for each given return, the probability of getting at least that return is higher in one lottery than in the other one. Formally, for each return x, if $F(x) \leq G(x)$, then the probability of getting at least x is higher in lottery F than in lottery G, bacause the area "on the left" of F is smaller. In other words, the CDF of a "better" lottery lies everywhere below that of a "worse" one — meaning that, for any return x, more probability mass lies above that return. This idea is captured by the concept of first-order stochastic dominance.

Definition 4.1. *The lottery F* ***first-order stochastically dominates*** *G if*

$$F(x) \leq G(x) \quad \textit{for all } x.$$

As illustrated in Figure 4.1, lottery F first-order stochastically dominates lottery G if its graph is always below the graph of lottery G.

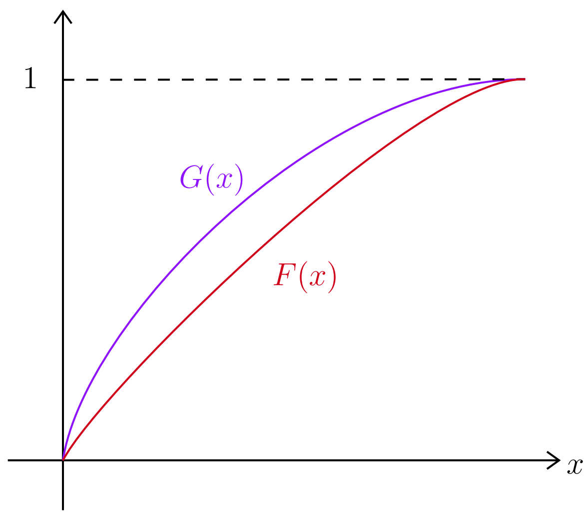


Figure 4.1: Lottery F first-order stochastically dominates lottery G.

There is a second way of capturing the idea: an individual with expected utility preferences should prefer lottery F to lottery G for any possibile utility function u. Formally,

$$\int u(x)dF(x) \geq \int u(x)dG(x) \quad \text{for every nondecreasing } u. \tag{4.1}$$

The two criteria (4.1) and (4.1) are equivalent, as stated by the following result.

Proposition 4.1. *Lottery F first-order stochastically dominates lottery G if and only if Equation* (4.1) *holds.*

Notice that first-order stochastic dominance is an *incomplete* ordering over lotteries: there are pairs of lotteries (F, G) such that neither F first-order stochastically dominates G nor G first-order stochastically dominates F.

First order stochastic dominance is a comparison of returns. We now develop a notion allowing us to compare riskiness. Again, we start from an intuitive idea: say that two lotteries have the same expected return, but a risk averter prefers one lottery to the other. Since the individual is risk averse, she must be preferring the less risky lottery. In this case, we say that the first lottery second-order stochastically dominates the second one.

Definition 4.2. *The lottery F* ***second-order stochastically dominates*** *G with the same mean if*

$$\int u(x)dF(x) \geq \int u(x)dG(x) \quad \textit{for every nondecreasing concave } u.$$

Recall that if an expected utility maximiser has a concave utility function, he is risk averse, which explains the qualifier in Definition 4.2.

4.2 Applications

come va.

Things to read. leggi Mas-Colell et al. (1995).

4.3 Exercises

Exercise 4.1. Prove one direction of Proposition 4.1: show that if Equation (4.1) holds, then F first-order stochastically dominates G in the sense of Definition 4.1. (check Mas-Colell et al. (1995, p. 195) if you are stuck)

References

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory.* Oxford university press New York. 30