MICROECONOMICS I: CHOICE UNDER UNCERTAINTY

MARCIN PĘSKI

Please let me know about any typos, mistakes, unclear or ambiguous statements that you find.

1. Introduction

1.1. **Suggested readings.**
MWG chapter 6.A. Kreps "Notes on the Theory of Choice", chapters 4 and 7 (the first part only).

1.2. **Describing the uncertainty.**

Example 1. A company develops a product of an unknown quality. The product can be either good or bad. Company manager must decide whether to advertise the product or not. The payoffs (say, in monetary units) for the manager are

	good	bad
Advertise	10	-1
Don't advertise	0	0

The manager's choice will reveal his preferences over the two actions.

In this example, the agent chooses between a pair of decisions "Advertise" and "Don't advertise". Each decision is described by a pair of numbers c_b and c_g that denote the monetary consequences of the decision given one of two states of the world, b and g.

We are going to assume that the pair of numbers (c_b, c_g) contains all the information necessary for the manager to make the decision. The assumption can be divided into two logically independent (sub) assumptions:

- there are only two states of the world that are relevant for the problem at hand,
- the monetary payoffs in each state describe all the relevant consequences of the decision.

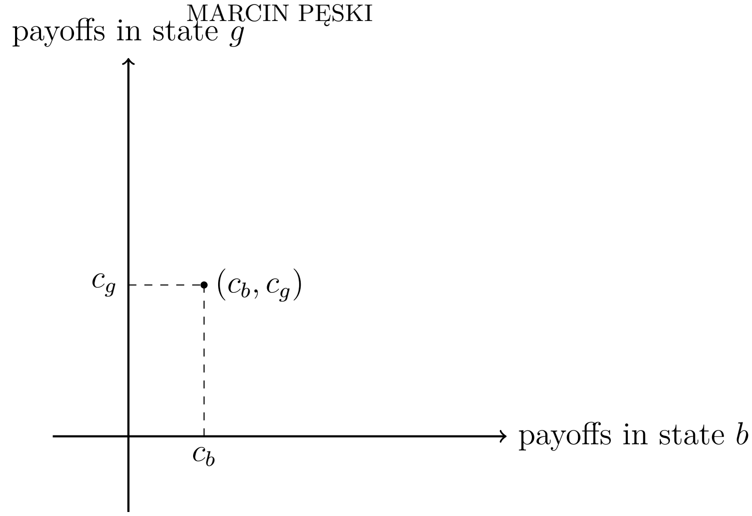

FIGURE 1.1. Space of acts, $S = \{c_b, c_g\}, Z = R$.

Mathematically, the pair of numbers can be represented as a mapping from the space of the states of the world $S = \{b, g\}$ into the space of "prizes" (i.e., monetary consequences) $Z = R$. Each mapping of this form $c : S \to R$ can be described as a pair of numbers $c = (c_b, c_g)$. We refer to such objects as *acts* and let $A = Z^S$ be the space of the acts.

With only two states, we can represent the space of acts on a two-dimensional diagram.

1.3. **Preferences.** The decision theory under uncertainty is a continuation of the decision theory that you learned in the beginning of the course. In particular, as in the standard decision theory, we typically make two assumptions:

- The decision maker has well-defined preferences $\preceq$ over acts. That implies (and is almost implied by) that the choices over the acts satisfy WARP.

We can use our diagram of the space of the acts to describe preferences o the manager, or at least draw her indifference curves. For example, if the manager prefers more money, then the indifference curves will be downward sloping (see Fig. Indifference curves).

- The preferences are continuous. Then, the utility theory (Proposition 3.C.1 from MWG) says that there exists a utility representation of the preferences $\preceq$:

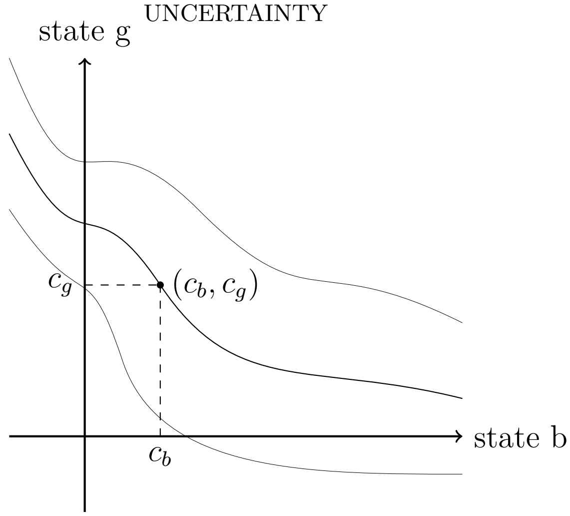

FIGURE 1.2. Indifference curves

there is a function $V : A \to R$ such that for any two acts $(c_b, c_g), \left(c'_b, c'_g\right) \in A$,

$$(c_b, c_g) \preceq \left(c'_b, c'_g\right) \text{ if and only if } V(c_b, c_g) \leq V\left(c'_b, c'_g\right).$$

As in the standard utility theory, the utility representation is a convenient, compact, and typically illuminating way of describing preferences.

1.4. **Preferences and probability.** In what follows, we will use the acts and the revealed preferences over acts to develop a theory of decision making under uncertainty.

Some of you might be surprised that so far, we haven't mentioned anything about the probabilities of the states. That is by design. Our theory does not require any other information about the probabilities of the states. Remember - that our goal is to describe the behavior of the manager. Given that we already know the manager's preferences - we have all the information we need and we do not need anything more.

To see how it works, suppose that the manager's preferences are given by function

$$V(c_b, c_g) = \frac{3}{4}c_b + \frac{1}{4}c_g. \tag{1.1}$$

So, the manager compares any two acts by taking a (weighted) average of monetary payoffs in each of the states of the world. The manager assigns three times as high weight to the "bad" state than the the "good" state.

There are many different reasons why the manager could have had such a preferences. Consider the following two examples:

- The manager cares only about the expected payoff from her decision. The manager is very experienced and she knows that 75% of the time the state of the world is bad. She applies her experience to the problem at hand to estimate that the probability that the state is "bad" is equal to $\frac{3}{4}$.
- The manager cares only about the expected payoff from her decision. She is a new manager on her first managerial position. She doe snot have any experience to estimate the probability of the "bad" state. She does not want to mess up, and she believes that she can reduce the chances of it is safer to assign higher weights to "bad" states of the world. The exact value of the weight, $\frac{3}{4}$, is taken from looking at the ceiling.
- The manager is a computer program designed by an engineer in a different country who was told by his boss to come up with some reasonable formula fast, no matter what as long as it is reasonable. The engineer picked up some formula from Wikipedia.
- If the state is "good", her utility is equal to c_g. If the state is "bad", her utility is equal to $\frac{1}{2}c_b$ (because, possibly, her company will =go bankrupt with probability $\frac{1}{2}$, in which case, she gets no profits at all). She assigns probability $\frac{6}{7}$ to the "bad" state. Her utility from profits in a bad state is equal to $\frac{1}{2}c_b$. The manager cares about the expected utility:

  $$\frac{6}{7}\left(\frac{1}{2}c_b\right)+\frac{1}{7}c_g=\frac{3}{7}c_b+\frac{1}{7}c_g=\frac{4}{7}\left(\frac{3}{4}c_b+\frac{1}{4}c_g\right)=\frac{4}{7}V\left(c_b,c_g\right).$$

  Notice that the above formula describes the same preferences as (1.1) - if we multiply utility function by a positive constant, the preferences are unchanged.

Each of the above cases describes different reasons why the manager has preferences represented by (1.1). Only in the first case, we can say that parameter $\frac{3}{4}$ can be *interpreted* as a probability of the state. Nevertheless, each of the cases above (including the first one) leads to the same preferences, and hence, the same decision theory.

We emphasize that the "probabilities" are not necessary for any decision theory. If we can interpret the preferences *as if* they arose from some probabilistic model of the

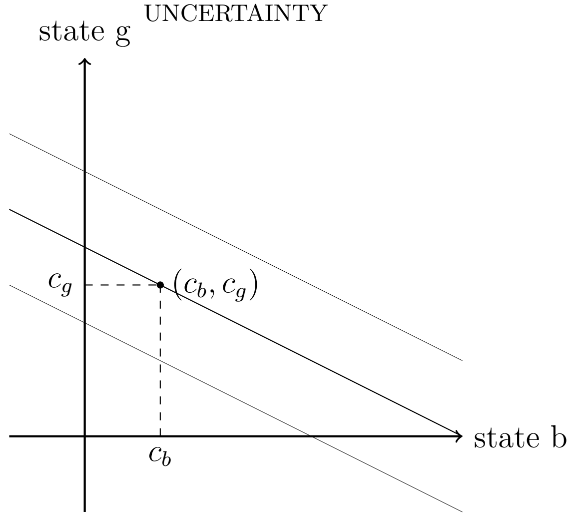

FIGURE 1.3. Indifference curves for expected payoff representation

world, that's great for us and it is helpful. But, the probabilities are nothing more (and nothing less) than that - interpretation.

More generally, all decision theory proceeds as follows: Find conditions (i.e., known as "axioms") on the behavior (i.e., preferences) that guarantee that the behavior can be interpreted *as if* the decision maker acts according to some model (for example, expected utility). The model is an *interpretation, "as if"*, and it does not really have to correspond to anything that "really" happens in the world (say, whether the decision maker in her brain computes the probabilities).

We need to remember this remark when we study the expected utility models.

1.5. **Examples of utility representations.** We list and compare few types of utility representations.

Expected payoff. Suppose that

$$V(c_b, c_g) = pc_b + (1-p)\,c_g$$

for some $p \in [0,1]$. The indifference curves are linear (Fig. 1.3), and the slope depends on p.

We can interpret p as a "probability" (we will talk about what exactly it means later) of the state "good". Then, V is the expected payoff from an act.

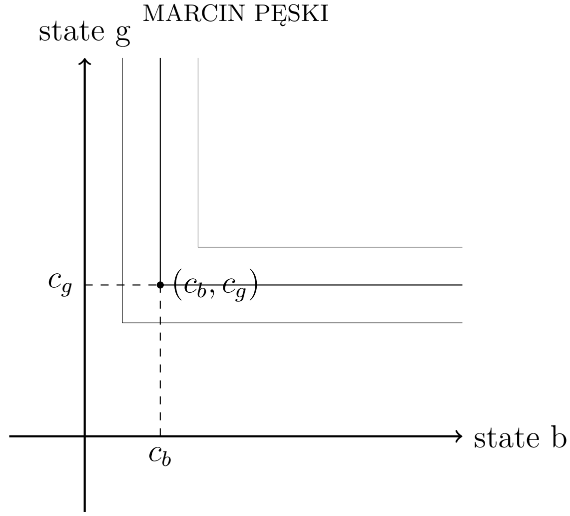

FIGURE 1.4. Indifference curves for the Maxmin representation

Expected utility. Let $u : R \to R$ be a standard (increasing, continuous) utility from monetary payoffs. Suppose that

$$V(c_b, c_g) = pu(c_b) + (1-p)u(c_g). \tag{1.2}$$

We say that V is the expected utility representation with (state-independent) utility function u and probabilities p for the "good" state.

Be careful to distinguish utility over acts $V(.)$ and "small" utility $u(.)$ from consumption given a state!

Exercise 1. Is it possible that the indifference curves of model (1.2) look like those drawn on Fig. 1.2? Draw the indifference curves if you know that $u()$ is concave.

Maxmin utility. Suppose that

$$V(c_b, c_g) = \min\{c_b, c_g\}.$$

Here, $V(c_b, c_g)$ is equal to the payoff in the worst possible state. The manager chooses to maximize the worst possible payoff. There is no probabilities in the representation - they are not relevant. The only relevant thing is the worst possible payoff. The indifference curves look like the indifference curves in the Leontieff utility - they have a corner on 45 degree line (see Fig. 1.4).

State-dependent utility.

Example 2. Alan likes ice cream and he wants to eat some ice cream tomorrow. Beth is able to provide the ice cream for Alan, but her offer is contingent on tomorrow's weather. There are two possible states "sun" and "rain". Beth offers two "insurance plans":

	sun	rain
plan 1	2	0
plan 2	1	1

Alan must choose today, before tomorrow's weather is known (and he is committed to his choice).

Alan's choice, intuitively will depend on

(1) how likely is each of the states of the world,
(2) how much he likes more ice cream vs. less (i.e., his marginal utility of ice cream), and
(3) how much he likes ice cream given the weather.

One example of Alan's preferences is *state-dependent expected utility:*

$$V(c_1, c_2) = pu_b(c_b) + (1-p)\,u_g(c_g) \tag{1.3}$$

Notice that the utility function depends on the state.

The representation (1.3) is somehow misleading in that it suggests a well-defined probability distribution over states. However, notice that if $p \in (0,1)$ we write down an equivalent representation for any other $p' \in (0,1)$. Indeed, define functions

$$u_b'(c_b) = \frac{p}{p'}u_b(c_b),$$
$$u_g'(c_g) = \frac{1-p}{1-p'}u_g(c_q).$$

Then,

$$\begin{aligned} V(c_b, c_g) &= p'\frac{p}{p'}u_b(c_b) + (1-p')\,\frac{1-p}{1-p'}u_g(c_g) \\ &= p'u_b'(c_b) + (1-p')\,u_g'(c_g). \end{aligned}$$

In other words, we can represent preferences V with two different sets of probabilities and state-dependent utility functions. Because both representations lead to exactly the same choices, there is no way that we can tell which representation is more correct.

For this reason, it is more appropriate to incorporate the probabilities in the definition of utility functions. Let

$$u''_b(c_b) = pu(c_b) \text{ and } u''_g(c_g) = (1-p)\,u(c_g).$$

Then,

$$V(c_b, c_g) = u''_b(c_b) + u''_g(c_g). \tag{1.4}$$

We refer to the additive form (1.4) as the state-dependent utility.

To summarize, the state-dependent utility representation does not determine the probabilities over states.

2. Model

2.1. Subjective vs. objective uncertainty.

The nice thing about our approach so far is that we do not need to make any assumptions about the true probabilities of the states. The only thing that matters are the preferences of the agent. In fact, we can try to identify the probabilities from the agent's preferences (see Section 1.4).

There is a fundamental difference about the concept of probability that we discuss here, and that is inferred from revealed choices of an individual, and the more familiar concept of the probability of the fact that the coin will fall Heads. The former is called "subjective" to emphasize the fact that it exists only in the mind of the individual (and only if we are willing to treat our interpretation as something more than "as if"). The other is "objective" - it can be measured, tested, repeated, etc.

From now on, we will distinguish between two types of probabilities: "subjective" and "objective". Our model of objective probabilities has four elements:

(1) The space of states S. The state space contains all possible realizations of the world. Or, more precisely, all possible and relevant to the choice at hand (the assumption that the agent is able to conceive all possible states of the world is often criticized, because there is too many of them.)

(2) The set of "prizes" Z. The prizes are things that directly affect our utility (consumption in monetary terms, amount of ice-cream, etc.)
(3) Acts: an *act* is a function $f : S \to Z$. The interpretation is that $f(s)$ is a "prize" that the agent receives if the true state of the world turns out to be s. This definition of an act was introduced by L. Savage, so we call them *Savage acts*, to distinguish from other type of acts (we will see the other ones soon). An example of an act, is a *constant act* δ_z that specifies the same prize z in each of the state of the world, $\delta_z(s) = z$ for each $s \in S$.
(4) Preference relation $\preceq$ over acts. Specifically, for any two acts $f$and g, $f \preceq g$ means that the agent (weakly) prefers act f to g.

For simplicity, we assume that S and Z are finite.

We won't make any assumption about the probabilities over states S, in fact we will try to infer it from the preferences.

However, we will assume that the individual agrees on the objective probabilities of some events. For example, we assume that the individual understands that the fair coin falls Heads with probability $\frac{1}{2}$ and that the dice shows 6 with probability $\frac{1}{6}$.

We are going to distinguish the objective probabilities by defining lotteries. A *lottery* is a probability distribution over prizes $q \in \Delta Z$. Hence, $q(z)$ is a objective probability of receiving prize z.

Example 3. Suppose that you have a raffle ticket that wins with probability $\frac{1}{100}$. The prize in the raffle is a free dinner in the best restaurant in town. We represent the ticket using a two-element set $Z = \{\text{dinner, nothing}\}$ and a lottery $q =$ "dinner"$^{\frac{1}{100}}$"nothing"$^{\frac{99}{100}}$ such that $q(\text{dinner}) = \frac{1}{100}$ and $q(\text{nothing}) = \frac{99}{100}$.The key here is that there is an objective probability with which the ticket wins. The objectivity here means that both the agent and the modeler and everybody else agree on the exact value and the nature of the probabilities.

2.2. **Anscombe-Aumann acts.** We are going to combine the subjective and objective probability in one model by extending our definition of an act: Instead of prizes, we assume that the act specifies lotteries over prizes given each state. Formally, an *Anscombe-Aumann act* (*AA act*, for short) is a mapping $f : S \to \Delta Z$ with the interpretation that $f(s)$ is an (objective) lottery which the agent receives if the state of the world turns out to be s, and $f(s)(z)$ is an objective probability with which the agent receives prize z if the world is s.

The distinction between subjective and objective uncertainty that lies at the heart of AA act should be clear at the theoretical, or conceptual levels. It is not so clear in the "real life", where it is often difficult to draw a border between two types of uncertain. One can construct some examples (For instance: Let

$$S = \{\text{cure for cancer within next 20 years, no cure for cancer within next 20 years}\}.$$

For each state s, the life length is a lottery with a distribution that can be easily computed from the mortality tables and an information about the reasons for each death. The probability of state $s \in S$ is not so easy to determine), but these examples are more often than not somehow artificial or forced.

On the other hand, our theory simply is going to assume that the decision maker is able to rank the AA acts Kreps says that the lotteries and AA acts are more often than not thought exercises than real objects. That's OK with us.

Do we need AA acts to develop a decision theory? No. Savage developed his whole expected utility theory without worrying about objective probabilities. But, his theory is difficult. AA acts make our life much easier. Why? Roughly, because they provide more data for our theory. To see it, notice first that there are more AA acts than Savage acts. (Each Savage act is an AA act with a degenerate lotteries in each state.) So, if we observe the preferences over all AA acts, we get much more information about the rankings than if we could only observe the preferences over Savage acts.

2.3. **Compound acts and convex combination of acts.** One of the reason why working with AA acts is easier is that the space of AA acts can be equipped with a natural operation of taking convex combinations. It turns out that (a) this operation is very useful mathematically and (b) it has a somehow natural interpretation. For any $\alpha \in [0,1]$, for any two acts f and g, define an act $\alpha f \oplus (1-\alpha) g$ so that for each s and z,

$$(\alpha f \oplus (1-\alpha) g)(s)(z) := \alpha f(s)(z) + (1-\alpha) g(s)(z). \tag{2.1}$$

This left-hand side is the (objective) probability that the agent receives prize z in state s given the act $\alpha f \oplus (1-\alpha)\, g$. The right hand side is a convex combination with weight α of the probability of prize z given state s and acts f and g. In other words, is a well-defined act.

We still need to check whether $\alpha f \oplus (1-\alpha)\, g$, the convex combination of two acts f and g with the weight α on the first act, is a well-defined act. That means, we need to check whether it assigns well-defined lotteries to each state. But,

$$\begin{aligned} &\sum_z (\alpha f \oplus (1-\alpha)\, g)\,(s)\,(z) \\ =&\sum_z \alpha f\,(s)\,(z) + \sum_z (1-\alpha)\, g\,(s)\,(z) \\ =&\alpha + (1-\alpha) = 1. \end{aligned}$$

So, we are good.

Mathematically, the above operation corresponds to taking convex combinations. There is also another, natural even if somewhat subtle interpretation. We will describe the interpretation very carefully to make sure that we understand what it says and what it does not say. First, notice that for any two acts f and g, and for any number $\alpha \in [0,1]$, we can consider a lottery over two acts $f^\alpha g^{1-\alpha}$. The lottery is objective (hence, the probability is known and well-understood) and with probability α chooses act f and with probability $1-\alpha$ chooses act g. We refer to such a lottery as a “compound act” and we can interpret it as just another (a slightly more general version of) act.

Second, the right-hand side of (2.1) is a probability of prize z in a equivalent reduced lottery with the timing of uncertainty reversed: first, Nature decides on the state s, then we draw with probabilities α and $1-\alpha$ the lotteries, respectively, $f\,(s)$ or $g\,(s)$, and then, finally, we use the drawn lottery to draw the prize. Thus, we can think about $(\alpha f \oplus (1-\alpha)\, g)\,(s)$ as a reduced form of the compound act $f^\alpha g^{1-\alpha}$.

To see how it works, consider the following example.

Example 4. You are an analyst in a macro consulting firm. You are asked to make a prediction about the future state of economy. if your prediction turns out to be correct, but only if so you will get a bonus. In short, you choose between two acts

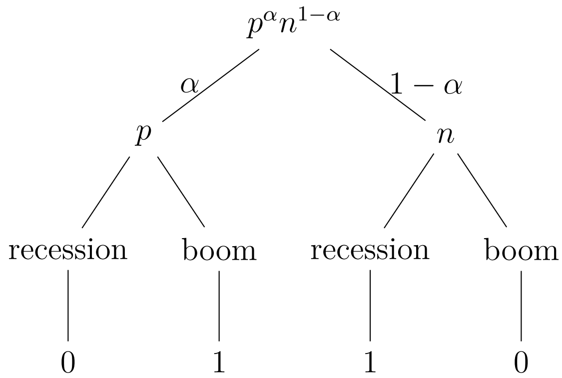

FIGURE 2.1. Compound act $p^\alpha n^{1-\alpha}$.

that can be described as follows:

	recession	boom
p(ositive)	0	1
n(egative)	1	0

Alternatively, you can also flip a coin (privately, of course) and offer a prediction based on the outcome of the coin. Given that the coin chooses p(ositive) prediction with probability α, such a behavior would correspond to a compound act

$$p^\alpha n^{1-\alpha}.$$

See Fig. 2.1. The reduced Anscombe-Aumann act corresponding to the above compound act can be described as follows:

	recession	boom
$\alpha p \oplus (1-\alpha)\, n$	$\alpha^0 (1-\alpha)^1$	$(1-\alpha)^0 \alpha^1$

If $\alpha = \frac{1}{2}$, then the reduced act yields the same lottery (50% chances for the bonus) regardless of the state of the world. See Fig. 2.2.

The underlying assumption for the definition (2.1) is that compound acts can be treated as if they are equivalent to their reduced forms. If we believe that the agents are good in reducing compound lotteries, then this seems a natural construction. Of course, it is an assumption - and, as with all other assumptions, people tend to make mistakes when they reduce compound lotteries. If we think that the problems

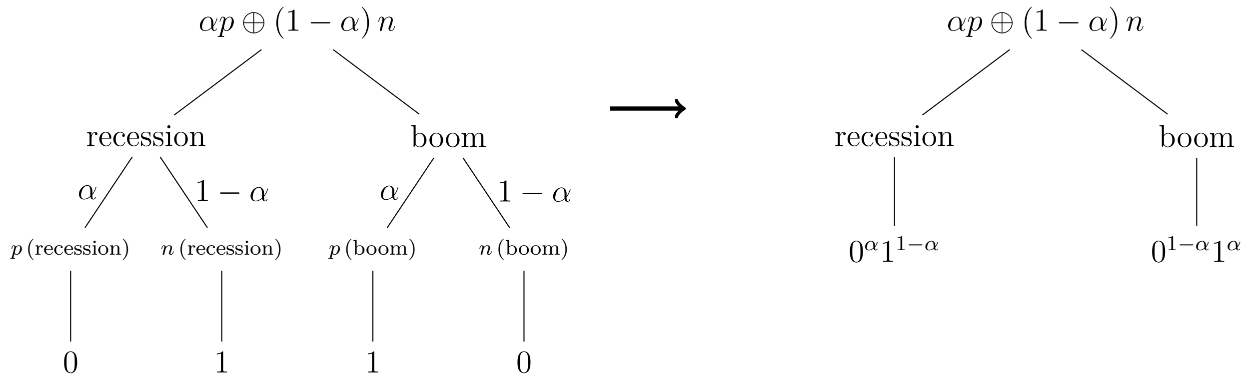

FIGURE 2.2. Reduced act $\alpha p \oplus (1-\alpha)\, n$ (right side).

with reducing compound lotteries are important, then the construction (2.1), or more precisely, the axioms based on this construction, will tend to be violated. We will go back to this issue soon.

So far, we used symbol $\oplus$ to describe the convex combination. We did this to emphasize the conceptual difference between compound acts and reduced acts. In the same time, as the above discussion suggests, from now on we will typically interpret both operations as the same. For this reason, it will be easier to use one symbol. From now on, we will write simply $\alpha f + (1-\alpha)\, g$ instead of $\alpha f \oplus (1-\alpha)\, g$.

2.4. **Two states, two prizes representation.** Suppose that the sets of prizes and states contain two elements each. Specifically, suppose that $Z = \{a, b\}$, and $S = \{s_1, s_2\}$.

An Anscombe-Aumann act assigns probability distribution over the prizes to each of the states. So, $f(a|s)$ is the probability of receiving prize a in state s. Because $f(b|s) = 1 - f(a|s)$, we can deduce the probability of prize b from the probability of a. In other words, the knowledge of probabilities $f(a|s)$ for each state s is enough to characterize an act.

Because $f(a|s) \in [0,1]$ for each state, we can interpret the Anscombe-Aumann act $f \in \mathcal{F}$ as a point in the "consumption space" $\mathcal{X} = [0,1]^2$ with the interpretation that

$(f_1, f_2) \in \mathcal{X}$ corresponds to the act f such that

$$f(a, s_1) = f_1, f(b|s_1) = 1 - f_1,$$
$$f(a, s_2) = f_2, f(b|s_2) = 1 - f_2,$$

For example, act $f = \left(\frac{1}{2}, 0\right) \in X$ means that the individual receives prize a with probability $\frac{1}{2}$ if the state is 1, and receives prize b with probability 1 if the state is 2.

Notice that for any two acts $f, g \in \mathcal{F}$, the convex combination of two acts, $\alpha f + (1-\alpha) g$, corresponds to the convex combination of the points in space X,

$$\alpha (f_1, f_2) + (1-\alpha)(g_1, g_2) = (\alpha f_1 + (1-\alpha) g_1, \alpha f_2 + (1-\alpha) g_2).$$

Now, the preferences over AA acts can be represented using the familiar diagrams with indifference curves.

2.5. **Purely objective theory (MWG).** The first development of the expected utility was purely objective. In fact, the discussion of the expected utility theory that you can read in MWG follows the purely objective path.

You will be happy to learn that our present model is strictly more general. To see why, notice that you can eliminate subjective uncertainty by assuming that the state space S contains a single element, say $S = \{*\}$. In such a case, an AA act $f : S \to \Delta Z$ is equivalent to an objective lottery $f(*) = q \in \Delta Z$. The decision maker's preferences over acts are equivalent to preferences over (objective) lotteries.

You may worry that given that our theory is more general, it is also going to be more complicated. Worry not! It turns out that we have already learned the main difference between theories - the idea of AA act. Everything else, including the axioms and the proofs that you will see in the next section, follow almost exactly the same lines.

Graphical representation with three prizes $Z = \{a, b, c\}$!

3. State-Dependent Expected Utility

3.1. **Suggested readings.** MWG chapter 6.B., Kreps "Notes on the Theory of Choice", chapter 7.

3.2. **Model.** Model:

(1) Finite state space S.
(2) Finite set of prizes Z. The set of lotteries over prizes ΔZ (i.e., the set of probability distributions over Z.
(3) (Anscombe-Aumann acts): functions $f : S \to \Delta Z$. Let $\mathcal{F}$ be the space of acts.
(4) Binary relation $\preceq$ over acts. The theory assumes that we can observe the agent choices over pairs of AA acts. The choices reveal the relation $\preceq$.

3.3. **Representation.** We say that relation $\preceq$ has a *state dependent expected utility (SDEU) representation* if and only if there exist functions $u_s : Z \to R$ such that for any two acts f and g,

$$f \preceq g \text{ iff}$$
$$\sum_{s,z} u_s(z) f_s(z) \leq \sum_{s,z} u_s(z) g_s(z).$$

We say that relation$\preceq$ has a *state independent expected utility (SIEU) representation* if and only if there exist function $u : Z \to R$ and probability distribution $\pi \in \Delta S$ such that for any two acts f and g,

$$f \preceq g \text{ iff}$$
$$\sum_{s,z} \pi_s u(z) f_s(z) \leq \sum_{s,z} \pi_s u(z) g_s(z).$$

3.4. **Axioms. Axiom 1. (Preference)** Binary relation $\preceq$ is a rational preference relation (i.e., it is transitive and complete).

The first axiom is obvious. Because $\preceq$ is a preference relation, we can use it to define the relations of strict preference $\prec$ and indifference $\sim$.

Axiom 2. (Continuity) For each act $g \in \mathcal{F}$, the upper and the lower contour sets of g,

$$\{f : g \preceq f\} \text{ and } \{f : f \preceq g\}$$

are closed.

[1]

The Continuity Axiom is a version of the same axiom that we used in the utility theory to prove the existence of continuous utility representation.

Axiom 3. **(Independence)**. For any three acts f, g, h, if $f \prec g$, then for any $\alpha \in (0,1)$, $\alpha f + (1-\alpha) h \prec \alpha g + (1-\alpha) h$.

The axiom seems natural: taking convex combination with the same act should not change the preference relations. This is the key axiom of the theory.

To understand the axiom a little bit better, take three acts three acts f, g, h, and suppose that the decision maker prefers act f to act g, $f \prec g$. Next, consider compound acts $f^{\alpha}h^{1-\alpha}$ and $g^{\alpha}h^{1-\alpha}$. The two compound acts are obtained by a lottery with probability α acts, respectively, f or g, and with probability $1-\alpha$ act h. We make two claims:

(1) If the decision maker's preferences were defined over the compound acts (and not only over the AA acts), our decision maker would rank the compound acts in the same way as the original acts, i.e. she would prefer $f^{\alpha}h^{1-\alpha}$ to $g^{\alpha}h^{1-\alpha}$.
(2) The ranking over the compound acts should be the same as the ranking over their reduced versions, $\alpha f + (1-\alpha) h$ and $\alpha g + (1-\alpha) h$.

Claim 1 can be explained by the fact that the two compound acts differ only in the event drawn with probability α, in which case the difference is equivalent to the difference between acts f and g. Claim 2 follows from our discussion from section 2.3 about the equivalence of compound acts and their reduced versions. If you believe that the two claims are correct, you must accept the Independence Axiom.

Example 5. Recall the situation from Example 4. Suppose that the manager has no clue which of the recession or boom is more likely. In fact, he may treat the two events completely symmetrically. In such a case, he would be indifferent between two acts p and n. Also, as in Claim 1, the manager is indifferent between choosing p, n, or flipping a coin and choosing the prediction based on the outcome of the coin. Using our terminology, he is indifferent $p^{\frac{1}{2}}n^{\frac{1}{2}}$.

[1] The literature usually describes a slightly weaker axiom, called Archimedean: For any three acts $f \prec g \prec h$, there exist sufficiently small $a, b \in (0,1)$ such that $af + (1-a) h \prec g \prec bh + (1-b) f$. We use the stronger version to compare to the consumer utility theory and for simplicity.

Further, suppose that the manager is not only utterly unqualified, but also very confident and he expects that given otherwise equal odds, the Nature is going to be on his side (that is not supposed to be a precise statement, but rather a description of a psychological state of his mind. I am sure that you know people like this). It seems at least plausible that such a manager would prefer act $p^{\frac{1}{2}}n^{\frac{1}{2}}$ to its reduced version $\frac{1}{2}p + \frac{1}{2}n$. In the former case, he believes that the sympathetic Nature will help him to win; in the latter case, the Nature has no chances in manipulating the objective lottery. Such a manager will fail Claim 2 and the Independence Axiom.

In this example, the manager fails Claim 2 because he prefers the uncertainty about his decision to be resolved *before* Nature chooses the state of the world. Later, in section 5.2, we will see a slightly different (in some sense, opposite) type of violation of the Independence Axiom.

Of course, if you don't believe Claim 1 or Claim 2 (for example, because you don't think that people are able quickly reduce compound lotteries, or because their attitude to the timing of the resolution of uncertainty), you won't believe the Independence Axiom. As we will prove soon, it turns out that if you don't believe the Independence Axiom, you must think that the decision maker is not an expected utility maximizer.

3.5. **Main result.**

Theorem 1. *Relation $\preceq$ has a state dependent utility representation if and only if it satisfies Axioms 1,2,3.*

Thus, by testing Axioms 1,2, and 3, we can test (more precisely, reject) the state-dependent theory.[2] The second part of the Theorem is also called invariance to affine transformations.

[2]The axioms are stated in such a way that they can be tested, in a particular way. An axiom can be violated by pointing to a pair (or triple) of acts for which the agent chooses differently from what the axiom says. The violation of the axioms means that the expected utility theory must be rejected.

We cannot imagine a real world experiment that would allow us to accept the expected utility theory. The problem, of course, is that the set of acts is infinite and we cannot possibly ask the agent to list all possible choices over all pairs of acts.

3.6. **Proof.**

Exercise 2. Show that if relation $\preceq$ has a state dependent utility representation, then it satisfies Axioms 1,2, and 3.

We will show that Axioms 1, 2, and 3 imply the state-dependent utility representation. The proof has two basic steps. The first step is to find an affine utility representation of relation $\preceq$. We show that there exists a function $F : \mathcal{F} \rightarrow R$ such that (a) F represents $\preceq$: for any two acts f and g,

$$f \preceq g \text{ iff } F(f) \leq F(g),$$

and (b) F is affine: for any two acts f and g, and for any number $\alpha \in (0,1)$,

$$F(\alpha f + (1-\alpha) g) = \alpha F(f) + (1-\alpha) F(g). \tag{3.1}$$

In the second step, we show that there exist functions $u_s : Z \rightarrow R$ such that for any act f,

$$F(f) = \sum_{s,z} u_s(z) f_s(z). \tag{3.2}$$

Parts (a) and (b) together establish the existence of state-dependent representation.

3.6.1. *Affine utility representation.* We start with some preliminary results.

Lemma 1. *There exists acts f^* and g^* such that for any act h, $f^* \preceq h \preceq g^*$.*

The acts f^* and g^* are respectively the worst and the best act.

Proof. Due to Axiom 1 and 2, the (standard) utility theory shows that there exists a continuous utility function $U : \mathcal{F} \rightarrow R$ that represents preferences on acts $f \in \mathcal{F}$. Because the space of acts is compact, there exists

$$f^* \in \arg\min_{f \in \mathcal{F}} U(f) \text{ and } g^* \in \arg\max_{f \in \mathcal{F}} U(f).$$

Because function U represents the preferences, the two acts are respectively, the worst and the best act. □

Lemma 2. *For any two acts $f \prec g$, any $\alpha, \beta \in (0,1)$, $\alpha > \beta$ if and only if $\alpha f + (1-\alpha) g \prec \beta f + (1-\beta) g$.*

Proof. Suppose that $\alpha > \beta$. Notice first that Axiom 2 implies that

$$\alpha f + (1-\alpha) g \prec \alpha g + (1-\alpha) g = g.$$

Using Axiom 3 again, we get

$$\begin{aligned}\beta f + (1-\beta) g &= \frac{\beta}{\alpha}(\alpha f + (1-\alpha) g) + \left(1-\frac{\beta}{\alpha}\right) g \\ &\succ \frac{\beta}{\alpha}(\alpha f + (1-\alpha) g) + \left(1-\frac{\beta}{\alpha}\right)(\alpha f + (1-\alpha) g) \\ &= \alpha f + (1-\alpha) g.\end{aligned}$$

The other direction follows from a similar argument applied to $\alpha \leq \beta$. □

Lemma 3. *Suppose that $f^* \prec g^*$. For each act h, there exists unique $\alpha_h \in [0,1]$ such that $h \sim \alpha_h g^* + (1-\alpha_h) f^*$.*

The Lemma says that any act that lies between f and g (in the sense of preference relation) can be calibrated to some mixture of acts f and g.

Proof. The result holds trivially if $h \sim g^*$ or $h \sim f^*$. So, assume that $f^* \prec h \prec g^*$. Recall that the preferences have a continuous utility representation U. Let $u : [0,1] \to R$ be defined as

$$u(\alpha) = U(\alpha g^* + (1-\alpha) f^*).$$

Then, u is continuous and $u(0) = U(f^*) < U(h) < U(g^*) = u(1)$. By the intermediate value theorem, there exists $\alpha_h \in (0,1)$ such that

$$U(h) = u(\alpha_h) = U(\alpha_h g^* + (1-\alpha_h) f^*).$$

□

We use the Lemmas to prove our claim. Let f^* and g^* be as in Lemma 1. The claim holds trivially when $f^* \sim g^*$ (or, in other words, the decision maker is indifferent among all acts). So, now we assume that $f^* \prec g^*$.

- Define function F using Lemma 3: For each act h, let $F(h) = \alpha_h$.

- Property (a) follows from the equivalence of the sequence of claims for any two acts h and h':

$$\begin{aligned} & h \preceq h', \\ \Longleftrightarrow_{\text{by definition of } \alpha} \quad & \alpha_h g^* + (1-\alpha_h) f^* \preceq \alpha_{h'} g^* + (1-\alpha_{h'}) f^* \\ \Longleftrightarrow_{\text{by Lemma } 3} \quad & \alpha_h \leq \alpha_{h'} \\ \Longleftrightarrow_{\text{by definition of } F(.)} \quad & F(h) \leq F(h'). \end{aligned}$$

- Property (b) follows from the following observations. Take $f^* \preceq h \preceq h' \preceq g^*$. Then, for any $x \in [0,1]$, by Lemma 3 and the definition of function F (fill all the remaining steps),

$$\begin{aligned} & xh + (1-x) h' \\ & \sim x((1-F(h)) f^* + F(h) g^*) + (1-x)((1-F(h')) f^* + F(h') g^*) \\ & = [x(1-F(h)) + (1-x)(1-F(h'))] f^* + [xF(h) + (1-x) F(h')] g^*. \end{aligned}$$

But this implies that

$$F(xh + (1-x) h') = xF(h) + (1-x) F(h').$$

3.6.2. *State-dependent expected utility.* We start with few definitions. First, we fix one price $z^* \in Z$ and let

$$\mathbf{z}^* = \delta_{z^*} = (z^*, ..., z^*)$$

be a constant act that always yields prize z^*. Next, for each price $z \in Z$ and each state $s \in S$, define deterministic act

$$f_{z,s} = (z^*, ..., z^*, z_{\text{state } s}, z^*, ..., z^*).$$

In other words, act $f_{z,s}$ yields prize z in state s and prize z^* in any other state $s' \neq s$. Let $F(f_{z,s})$ be the value of act $f_{z,s}$ according to function F. And, finally, define

$$u_s(z) = F(f_{z,s}) - \frac{n-1}{n} F(\mathbf{z}^*).$$

We will show that (3.2) is satisfied. We will do it in two short steps. First, we will show the claim for deterministic act f, i.e., $f(s) = z_s \in Z$ for each state s. Let $n = |S|$ be the number of states. To shorten the notation, we assume that we can enumerate all states $s \in S$ from 1 to n. Notice that the mixture act $\frac{1}{n}f + \frac{n-1}{n}\mathbf{z}^*$ is equal to the mixture (with equal weights) of acts $f_{z_1,1}, ..., f_{z_n,n}$.

$$\begin{aligned}\frac{1}{n}f + \frac{n-1}{n}\mathbf{z}^* = &\left(\frac{1}{n}z_1 + \frac{n-1}{n}z^*, ..., \frac{1}{n}z_n + \frac{n-1}{n}z^*\right)\\ =&\frac{1}{n}\left(z_1, z^*, ..., z^*\right)\\ &+ \frac{1}{n}\left(z^*, z_2, ..., z^*\right)\\ &+ ...\\ &+ \frac{1}{n}\left(z^*, ..., z^*, z_n\right)\\ =&\frac{1}{n}f_{z_1,1} + \frac{1}{n}f_{z_2,2} + ... + \frac{1}{n}f_{z_n,n}\end{aligned}$$

By repeated application of (3.1),

$$\begin{aligned}\frac{1}{n}F\left(f\right) + \frac{n-1}{n}F\left(\mathbf{z}^*\right) =&F\left(\frac{1}{n}f + \frac{n-1}{n}\mathbf{z}^*\right)\\ =&\frac{1}{n}F\left(f_{z_1,1}\right) + ... + \frac{1}{n}F\left(f_{z_n,n}\right)\\ =&\frac{1}{n}\left(u_1\left(z_1\right) + ... + u_n\left(z_n\right)\right) + \frac{n-1}{n}F\left(\mathbf{z}^*\right).\end{aligned}$$

But this implies that

$$F\left(f\right) = u_1\left(z_1\right) + ... + u_n\left(z_n\right).$$

To conclude, we need to show (3.2) holds for all acts, not necessarily deterministic. We leave this part as an exercise.

3.7. **Invariance to affine transformations.**

Theorem 2. *Collections of functions (u_s) and (v_s) are two SDEU representations of the same preference relation if and only if there exists constant $a > 0$ and $b_s \in R$ for each state s such that for each s, $v_s = au_s + b_s$.*

Proof. In the exercise, you are asked to show that a positive affine transformation of utility functions does not change the preferences. The converse also holds. Any two expected utility representations must be affine transformations of each other. Here is the argument.

- Suppose that two different sets of utility functions (u_s) and (v_s) are expected utility representations of the same relation $\preceq$. For each AA act f, let

$$U(f) = \sum_s u_s(z) f_s(z),$$
$$V(f) = \sum_s v_s(z) f_s(z)$$

- Let f^* and g^* be the single worst and single best acts for the preference relation $\preceq$. (Note that such acts always exist by Lemma 1.) Then, it must be that

$$f^* \in \arg\min_f U(f) \text{ and } f^* \in \arg\min_f V(f)$$

  Similarly,

$$g^* \in \arg\max_f U(f) \text{ and } g^* \in \arg\max_f V(f)$$

  Let

$$a = \frac{V(g^*) - V(f^*)}{U(g^*) - U(f^*)}$$

  and for each state s,

$$b_s = v_s(f_s^*) - a u_s(f_s^*)$$

- Fix state s and choose any prize z. Consider act $f^* z_s = \left(f_1^*, ..., f_{s-1}^*, z, f_{s+1}^*, ..., f_S^*\right)$ that is obtained from f^* by replacing the price in state s by z. Find $\alpha_{zs} \in [0,1]$ such that $f^* z_s$ is indifferent to the mixture $\alpha_{zs} f^* + (1 - \alpha_{zs}) g^*$. Then,

$$U(f^* z_s) = \alpha_{zs} U(f^*) + (1 - \alpha_{zs}) U(g^*), \text{ and}$$
$$U(f^* z_s) - U(f^*) = (1 - \alpha_{zs}) [U(g^*) - U(f^*)].$$

  Because the left hand side of the second equality is equal to $u_s(z) - u_s(f_s^*)$, we get

$$u_s(z) = u_s(f_s^*) + (1 - \alpha_{zs}) [U(g^*) - U(f^*)].$$

  A similar relation holds for the second representation:

$$v_s(z) = v_s(f_s^*) + (1 - \alpha_{zs}) [V(g^*) - V(f^*)].$$

- Hence,

$$\begin{aligned}v_s(z) - au_s(z) - b_s =& v_s(f_s^*) + (1-\alpha_{zs})[V(g^*) - V(f^*)] \\ & - au_s(f_s^*) - (1-\alpha_{zs})a[U(g^*) - U(f^*)] \\ & - (v_s(f_s^*) - au_s(f_s^*)) \\ =& (1-\alpha_{zs})[(V(g^*) - V(f^*)) - a(U(g^*) - U(f^*))] \\ =& 0.\end{aligned}$$

Because it is true for all states and prizes, we get the result.

□

Notice that this is a much stronger result than invariance to monotone transformations that you already know from the classic utility theory without uncertainty. With uncertainty, we can pin down the class of utilities much more precisely. The reason is that, here, we observe preferences over lotteries, which gives us much more data.

4. State-independent expected utility

So far we did not manage to fulfill our promise and determine subjective probabilities. The problem is that with state-dependent utility it is impossible. Here we show that if preferences are state-independent, we uniquely determine probabilities.

We need one more axiom. We start with a notation. Take any act f. For each state s and a lottery $p \in \Delta Z$, define an act $f_s p$ so that

$$f_s p = (f_1, ..., f_{s-1}, p, f_{s+1}, ..., f_S).$$

In other words, act $f_s p$ is equal to act f in all states but s and it yields lottery p (instead of lottery f_s) in state s.

Axiom 4. (State-independence). Suppose that for some act f, two lotteries $p, q \in \Delta Z$, and state s, we have

$$f_s p \preceq f_s q.$$

Then, for any other state s',

$$f_{s'} p \preceq f_{s'} q.$$

The axiom says that the preference over the lotteries p and q do not depend on the state. This is a very strong axiom. It is certainly inappropriate for some situations in which clearly your utility should be affected by state (like utility of carrying an umbrella should depend on the weather), but it is good for some other (your utility of carrying umbrella is presumably independent on whether Hillary Clinton wins US presidential elections).

Theorem 3. *Suppose that rational preferences $\preceq$ over acts satisfy Independence, Continuity, and State-independence. Then, $\preceq$ has state-independent expected utility representation. If SIEU representation exists and all states have strictly positive probability ($\pi_s > 0$ for all s), then the preferences satisfy State-Independence (and other axioms as well). Moreover, the utility function u is identified up to affine transformations.*

Proof. Steps:

(1) The Independence and Continuity imply (Theorem 1) that there exist functions $u_s : Z \to R$ such that for any two acts f and g,

$$f \preceq g \text{ iff}$$

$$\sum_{s,z} u_s(z) f_s(z) \leq \sum_{s,z} u_s(z) g_s(z).$$

(2) Fix state $s_0 \in S$. By the State-Independence Axiom, for any lottery p and q and for any state s,

$$f_{s_0} p \preceq f_{s_0} q \Longleftrightarrow f_s p \preceq f_s q. \tag{4.1}$$

Using the SDEU representation from step 1 and the fact that acts $f_s p$ and $f_s q$ agree everywhere but in state s, we notice that the left-hand of (4.1) is equivalent to

$$\sum_z u_{s_0}(z) p(z) \leq \sum_z u_{s_0}(z) q(z), \tag{4.2}$$

and the right-hand side is equivalent to

$$\sum_z u_s(z) p(z) \leq \sum_z u_s(z) q(z). \tag{4.3}$$

Thus, for each state s, any two lotteries p, q, we have (4.2) if and only if (4.3).

(3) Following similar arguments to those used in the proof of the invariance to affine transformations, we can show that the equivalence between (4.2) and (4.3) implies that there exists $a_s > 0$ and b_s such that

$$u_s(z) = a_s u_{s_0}(z) + b_s.$$

(4) Define $\pi_s := \frac{a_s}{\sum_{s'} a_{s'}}$ and $u(z) := u_{s_0}(z)$. Then, for any two acts $\Longleftrightarrow$

$$\begin{aligned}
& f \preceq g \\
\Longleftrightarrow & \sum_{s,z} u_s(z) f_s(z) \leq \sum_{s,z} u_s(z) g_s(z) \\
\Longleftrightarrow & \sum_{s,z} (a_s u_{s_0}(z) + b_s) f_s(z) \leq \sum_{s,z} (a_s u_{s_0}(z) + b_s) g_s(z) \\
\Longleftrightarrow & \left(\sum_{s'} a_{s'}\right) \sum_{s,z} \pi_s u(z) f_s(z) + \sum_s b_s \sum_z f_s(z) \leq \left(\sum_{s'} a_{s'}\right) \sum_{s,z} \pi_s u(z) g_s(z) + \sum_s b_s \sum_z g_s(z).
\end{aligned}$$

Because for each state s, $\sum_z f_s(z) = \sum_z g_s(z) = 1$, the last terms on both sides of the above inequality cancel out. By further dividing the inequality by $\sum_{s'} a_{s'}$, we obtain that

$$f \preceq g \Longleftrightarrow \sum_{s,z} \pi_s u(z) f_s(z) \leq \sum_{s,z} \pi_s u(z) g_s(z).$$

This demonstrates that the preferences $\preceq$have SIEU representation (π, u).

□

We will also show an analogue of the Invariance to Affine Transformations for the state-independent expected utility.

Theorem 4. *Suppose that preferences $\preceq$ over acts are represented by (π, u) and (ρ, v) for some $\pi, \rho \in \Delta S$ and $u, v : Z \to R$. Then, $\pi = \rho$ and there exists $a > 0$ and $b \in R$ such that for each z, $v(z) = au(z) + b$.*

Proof. Take any two prizes $x, y \in Z$ such that $u(x) < u(y)$ (if such prizes do not exist, then the result is trivial). Consider the following acts

$$\mathbf{x} = (x, ..., x),$$

$$\mathbf{y} = (y, ..., y).$$

For each state s, let $\mathbf{x}_s y$ be an act obtained from $\mathbf{x}$ by replacing the prize in state s by y.

Notice that

$$\mathbf{x} \preceq \mathbf{x}_s y \preceq \mathbf{y}.$$

This follows from the fact that the preferences have SIEU representation (you can use any of the representations) and $u(x) < u(y)$. Moreover, because $u(x) < u(y)$ (and by the same argument as in the proof of Lemma 3), there exists a unique $\alpha_s \in [0,1]$ such that

$$\mathbf{x}_s y \sim \alpha_s \mathbf{y} + (1-\alpha_s)\,\mathbf{x}.$$

Let $U(f)$ and $V(f)$ denote the expected utility of act f using, respectively, the first and the second representations. Then,

$$U(\mathbf{x}_s y) = \pi_s u(y) + (1-\pi_s)\,u(x),$$
$$U(\alpha_s \mathbf{y} + (1-\alpha_s)\,\mathbf{x}) = \alpha_s u(y) + (1-\alpha_s)\,u(x),$$

and $U(\mathbf{x}_s y) = U(\alpha_s \mathbf{y} + (1-\alpha_s)\,\mathbf{x})$ implies that $\alpha_s = \pi_s$. Because the latter holds for each state s, and because a similar argument shows that $\alpha_s = \rho_s$ for each state s, we obtain $\pi = \rho$.

To see the second part, notice that each SIEU is also a SDEU. Specifically, preferences $\preceq$ have two SDEU representations (u_s) and (v_s), where, for each state s and prize z ,

$$u_s(z) := \pi_s u(z) \text{ and } v_s(z) := \rho_s v(z).$$

Because $\pi = \rho$, and by Theorem 2, there exists $a > 0$ and b_s such that for each state s and prize z

$$\pi_s v(z) = v_s(z) = a u_s(z) + b_s = a\pi_s u(z) + b_s$$

By dividing both sides by π_s, we obtain, for each state s and prize z

$$v(z) = a u(z) + \frac{b_s}{\pi_s}.$$

Because the above is true for each state s, it must be that $\frac{b_s}{\pi_s}$ does not depend on state s. Let $b = \frac{b_s}{\pi_s}$. □

5. Famous violations of expected utility

Next, we discuss some violations of the expected utility theory.

5.1. **Allais paradox.**

There are three prizes $\{10, 1, 0\}$ (in million dollars). The agent is asked about two choices:

	10	1	0
q_1	0	1	0
q_2	.10	.89	.01

and

	10	1	0
q_1'	0	0.11	0.89
q_2'	0.10	0	0.90

Most people choose $q_1$over q_2, and q_1' over q_2'. Nevertheless such choices violate the Independence axiom (and hence, violate the expected utility model). Indeed, define lotteries

$$g_1 = \frac{1}{2}q_1 + \frac{1}{2}q_1' = \frac{1}{2}q_1' + \frac{1}{2}q_1,$$
$$g_2 = \frac{1}{2}q_2 + \frac{1}{2}q_1' = \frac{1}{2}q_2' + \frac{1}{2}q_1.$$

By independence, the ranking between g_1 vs g_2 should be the same as $q_1$vs q_2, and q_1' vs. q_2'. But, of course, this means that the two latter rankings must be equal. (PICTURE!)

The story is that people have a trouble of correctly analyzing small probabilities. Here, they exaggerate the probability of the prize 0 in lottery q_2, which leads to inverse ranking among lotteries q_1 and q_2.

A simple example of a preference models that exhibits Allais paradox type of behavior would be a *probability weighing*: Let $h : R \to R$ be a strictly increasing function such that $h(0) = 0$. For each act f, compute

$$U_h(f) = \sum_{s,z} h(p_s) f(z|s) u(z).$$

We use the "probability-weighed" expected utility $U_h(.)$ to compare acts. If $h(p) = p$, then we have standard (state-independent) expected utility. If $h(p)$ is concave, then U_h overweighs small probabilities relative to large ones, which could rationalize the above choices.

5.2. **Ellsberg paradox.** Imagine the following thought experiment. There is an urn that contains 90 balls. There are 29 balls of color red; the remaining 61 balls are either blue or green. (It is possible that all the remaining balls have the same color). No further information is provided.

A ball is drawn randomly from the urn and its color is not revealed. The subject is asked to compare four bets:

Bet A: You get \$100 if the ball is Red and ,

Bet B : You get \$100 with probability 1/2 if the ball is Blue or Green,

Bet C: You get \$100 if the ball is Blue,

Bet D: You get \$100 if the ball is Green.

Most people choose

$$C \sim D \prec A \prec B. \tag{5.1}$$

However, these strict preferences are inconsistent with the expected utility theory. To see it, suppose that $q_R, q_B, q_G > 0$ are the probabilities of colors red, blue or green, $q_G = 1 - q_B - q_R$. Then, assuming for simplicity that $u(0) = 0$ and that $u(100) > 0$,

$$\begin{aligned} C \sim D &\Rightarrow q_B = q_G, \\ D \prec A &\Rightarrow q_R > q_G, \\ A \prec B &\Rightarrow \frac{1}{2}(q_B + q_G) > q_R. \end{aligned}$$

But these claims are inconsistent with the Independence Axiom.

The idea is that people see bets A and B as relatively safe, the probabilities of outcomes are well-understood. Bet B is better because it leads to strictly higher probability of winning the prize. On the other hand, bets C and D are not “safe”, the subject does not know how many Blue or Green balls are in the urn. So, given a priori symmetry between Blue and Green, they are indifferent between C and D and they prefer A to both of them.

Define the state as the color of the ball, $S = \{R, B, G\}$, and let $Z = \{0, 100\}$. The bets correspond to the following acts:

$$A = (1, 0, 0)\,,$$
$$B = \left(0, \frac{1}{2}, \frac{1}{2}\right),$$
$$C = (0, 1, 0)\,,$$
$$D = (0, 0, 1)\,.$$

We claimed that choices (5.1) are inconsistent with expected utility. So, some axioms must be violated. Indeed, the independence axiom is violated. Notice that

$$B = \frac{1}{2}C + \frac{1}{2}D.$$

If the decision maker is indifferent between C and D, she should be also indifferent between any of these two and B. But, $C, D \prec B$.

The Ellsberg paradox can be interpreted in many different ways:

- Most of the recent literature prefers the interpretation that there is a fundamental difference between two types of uncertainty (a) risk - a situation you don't know what outcome is going to occur but you understand very well the objective probabilities of this outcome (it is like choosing between A and B) and (b) ambiguity - a situation, where you don't even know or you are not willing to trust the probabilities. In this case, the Ellsberg paradox is an example of "ambiguity aversion".

  An example of preferences that rationalizes choices (5.1) is maxmin utility of Gilboa and Schmeidler: For each act $f = (f_R, f_B, f_G)$, let

  $$u(f) = \frac{29}{90} f_R + \frac{61}{90} \min(f_R, f_G)\,.$$

  Here, two types of uncertainty are handled differently: The risk corresponds to the expected utility part with well-defined probabilities of Red ball and Blue-or-Green ball. The ambiguity corresponds to uncertainty between Blue and Green. Observe that

  $$u(C) = u(D) < u(A) < u(B)\,.$$

- There are two other interpretations that attribute the Ellsberg paradox to bounded rationality of the decision maker. First, it seems that the decision makers perceive two types of uncertainty: one is about drawing the ball from the urn, and the other about the composition of the urn. They do not combine theses uncertainties in a proper way. This interpretation suggests that the subjects make a mistake in evaluating the compound lotteries.
- Second, suppose that, in real life (as contrasted with the laboratory experiment) whenever the decision maker faces a situation in which she is arbitrarily refused information that is available to somebody else, she may suspect that the other party is trying to trick her. There are multiple examples (lemon problem, buying insurance policy with lots of unreadable fine print, etc.) It is easy to imagine that the decision maker confuses the laboratory experiment with her real life experiences (or, in other words, does not put enough mental effort to distinguish between those two situations).

6. Updating

Which axiom allows us to do updating? TBA

7. Decision making under uncertainty

The goal of this lecture is to provide introduction to the decision making under uncertainty. In the first two parts, we explain the idea of stochastic dominance. In the last two parts, we use the dominance to derive two comparative statics results about the decisions in an uncertain environment.

For technical reasons, we assume that all lotteries have a bounded support contained in the interval $[a, b]$ for some $a < b$. For each cdf F on interval $[a, b]$, each function $u : [a, b] \to R$, we denote the expected value of function u with respect to cdf F as

$$F[u] = \int_a^b u(x)\, dF(x).$$

7.1. **Stochastic dominance.**

Definition 1. Say that F *first-order stochastically dominates* G (we write $F \rhd^1 G$) if for each x, $F(x) \leq G(x)$. Say that F *second-order stochastically dominates* G (we write $F \rhd^2 G$) if for each $t \leq b$, $\int_a^t F(x)\,dx \leq \int_a^t G(x)$.

The first result states that the first-order stochastic dominance is a stronger condition of the two.

Lemma 4. *For any two lotteries F and G, $F \rhd^1 G$ implies $F \rhd^2 G$.*

Proof. Obvious. □

The next two results provide a partial characterization of the second order stochastic dominance.

Lemma 5. *For any two lotteries F and G, $F \rhd^2 G$ implies $E_F \geq E_G$.*

Proof. By integration by parts,

$$E_F = \int_a^b x dF(x) = bF(b) - aF(a) - \int_a^b F(x)\,ds = b - \int_a^b F(x)\,dx.$$

The last equality comes from the fact that $F(a) = 0$ and $F(b) = 1$. Because similar equalities hold for lottery G, we obtain,

$$E_F - E_G = \int_a^b (G(x) - F(x))\,dx.$$

The result follows. □

Lemma 6. *Suppose that there exists $x_0 \in [a, b]$ such that for each $x \leq x_0$, $G(x) \geq F(x)$ and for each $x \geq x_0$, $G(x) \leq F(x)$. Then, if $E_F \geq E_G$, then $F \rhd^2 G$.*

Proof. Define function

$$h(t) = \int_a^t (G(x) - F(x))\,dx.$$

Notice that $h(a) = 0$ and because of the proof of the previous result,

$$h(b) = \int_a^b (G(x) - F(x))\,dx = b - \int_a^b F(x)\,dx - \left(b - \int_a^b G(x)\,dx\right) = E_F - E_G \geq 0.$$

Moreover, the assumptions imply that function $h(t)$ is increasing for $t \leq x_0$ and decreasing for $t \geq x_0$. It follows that $h(t) \geq 0$ for each $t \in [a,b]$, which implies that $F \rhd^2 G$. □

7.2. **Comparison of lotteries.**

Theorem 5. *Cdf. F first-order dominates G if and only if $F[u] \geq G[u]$ for each increasing utility function $u : [a,b] \to R$.*

Proof. Suppose that $u(.)$ is differentiable. Then, by integration by parts,

$$F[u] = \int_a^b u(x)\, dF(x) = u(b)\, F(b) - u(a)\, F(a) - \int_a^b F(x)\, u'(x)\, d(x),$$

and, because $F(a) = G(a) = 0$, $F(b) = G(b) = 1$,

$$F[u] - G[u] = \int_a^b (G(x) - F(x))\, u'(x)\, d(x)$$

Because $u(.)$ is increasing, $u'(x) \geq 0$. Then, $F[u] \geq G[u]$ for any differentiable and increasing u if and only if $G(x) \geq F(x)$ for each $x \in [a,b]$. The result follows from the fact that any increasing utility function can be approximated by a differentiable and increasing utility function. □

Cdf. F second-order dominates G if and only if $F[u] \geq G[u]$ for each increasing and concave utility function $u : [a,b] \to R$.

Proof. For each $t \in [a,b]$, denote

$$\bar{F}(t) = \int_a^t F(x)\, dx,$$
$$\bar{G}(t) = \int_a^t G(x)\, dx.$$

Suppose that $u(.)$ is twice differentiable. Then, by two applications of integration by parts,

$$F[u] = \int_a^b u(x)\, dF(x) = u(b)\, F(b) - u(a)\, F(a) - \int_a^b F(x)\, u'(x)\, d(x)$$
$$= u(b)\, F(b) - u(a)\, F(a) - u'(b)\, \bar{F}(b) + u'(a)\, \bar{F}(a) + \int_a^b \bar{F}(x)\, u''(x)\, d(x),$$

and, because $\bar{F}(a) = \bar{G}(a) = 0$,

$$F[u] - G[u] = u'(b)\left(\bar{G}(b) - \bar{F}(b)\right) - \int_a^b \left(\bar{G}(x) - \bar{F}(x)\right) u''(x)\, d(x).$$

Because $u' \geq 0$, and $u'' \leq 0$, $F[u] \geq G[u]$ for each twice differentiable, increasing and concave u if and only if $\bar{G}(x) \geq \bar{F}(x)$ for each $x \in [a, b]$

The result follows from the fact that any increasing concave utility function can be approximated by a twice differentiable, increasing, and concave utility function. □

7.3. **Comparative statics under uncertainty.** In this section, we consider the following problem. A decision maker has utility $u(y, x)$ that depends on the value of the decision taken x and the uncertain state of the world y. The state of the world is drawn from cdf F_t that is indexed by a parameter t. The goal of the decision maker is to choose x that maximizes the expected utility:

$$\max_x \int u(y, x)\, dF_t(y). \tag{7.1}$$

Let $x^*(t)$ denote the set of optimal choices in the above problem.

Examples of applications include the optimal choice of insurance given prices, the optimal investment allocations, etc.

Say that family F_t is ordered by the first-order stochastic dominance (f.o.s.d.), if for each $t < t'$, distribution $F_{t'}$ first-order stochastically dominates F_t.

Theorem 6. *Let $x^*(t)$ be the maximizer set of problem (7.1). If family F_t is ordered by f.o.s.d., and u has increasing differences, then $x^*(t)$ is (weakly) increasing.*

Proof. Let

$$V(x, t) = \int u(y, x)\, dF_t(y).$$

By the Theorem from the Section Increasing differences in Compareative Statics notes, it is enough to show that $V(x, t)$ has increasing differences. For any $t < t'$, and $x < x'$, we have

$$\begin{aligned}&[V(x', t') - V(x, t')] - [V(x', t) - V(x, t)]\\ &= \int [u(y, x') - u(y, x)]\, dF_{t'}(y) - \int [u(y, x') - u(y, x)]\, dF_t(y).\end{aligned}$$

By the assumption, function $f(y) = u(y, x') - u(y, x)$ is increasing. But then, the stochastic dominance and Theorem 5 implies that

$$\int f(y)\, dF_{t'}(y) - \int f(y)\, dF_t(y) \geq 0.$$

□

7.4. **Monotone Likelihood Ratio Property.** In this section, we consider a special case of problem (7.1), in which the parameter t is interpreted as a signal about the unknown state of the world.

Consider the following (interim) decision problem. A state of the world y is drawn from distribution $\pi \in \Delta Y$. For simplicity, we assume that the space of states of the world is a finite subset of real numbers, $Y \subseteq R$ and $|Y| < \infty$. Then, $\pi(y)$ is the prior probability of state y. The "finiteness" assumption is for simplicity of the exposition only and all our results can be easily generalized.

The decision maker does not observe y, but observes a signal $t \in R$ drawn from distribution $H(.|y) \in \Delta R$. After learning t, the decision maker chooses an action $x \in R$. Finally, she receives a payoff $u(y, x)$. We assume that the distribution of signals has Lebesgue density $h(t|y)$. In such a case, the decision problem can be written as follows:

$$\max_x \int u(y, x)\, f_\pi(y|t)\, dy, \tag{7.2}$$

where

$$f^\pi(y|t) = \frac{h(t|y)\, \pi(y)}{\sum_{y'} h(t|y')\, \pi(y')}$$

is the conditional (posterior) probability of the state of the world given signal t. Let $F_\pi(y|t) = \sum_{y' \leq y} f_\pi(y'|t)$ be the cdf of the posterior distribution. We want to know how the action changes with the signal.

Are there any natural conditions on the fundamentals (i.e., distribution of signals conditionally on the state of the world, $H(.|y)$) that guarantee that family $F_\pi(.|t)$ is ordered by stochastic dominance? Milgrom ("Good news, bad news") provided such conditions.

Definition 2. For two signals $t, t' \in T$, say that signal t' is *more favorable than* t if for each prior $\pi \in \Delta Y$, $F_\pi(.|t') \rhd^1 F_\pi(.|t)$.

We emphasize that the above definition must work for *all* priors π.

The reason why the definition is useful because it ties with the above comparative statics results. Let

$$x^*_\pi(t) = \arg\max_x \sum_y u(y,x) f_\pi(y|t)$$

be the set of optimal solutions to the optimization problem (7.2) with $F_t = F_\pi(.|t)$. Theorem 6 implies that $x^*(t)$ is increasing if utility function u has increasing differences, and the conditional distributions $F_\pi(.|t)$ are ordered by stochastic dominance. Thus, we have an immediate corollary to the Theorem 6.

Corollary 1. *Suppose that $u(y,x)$ has increasing differences. If t' is more favorable than t, then for each prior π,*

$$x^*_\pi(t) \leq_S x^*_\pi(t').$$

The next result presents a characterization of more favorable signals in terms of the conditional density h:

Theorem 7. *For any two signals, signal t' is more favorable than signal t if and only if for each $y' < y$,*

$$\frac{h(t'|y')}{h(t|y')} \leq \frac{h(t'|y)}{h(t|y)}. \tag{7.3}$$

We refer to inequality (7.3) as the marginal likelihood ratio property. The inequality says that the ratio of likelihoods of signals t' to t increases if the state shifts from y' to y. Intuitively, given higher state y, signal t' is relatively more likely than given the lower state y'.

Observe that the MLR property has a flavor of increasing differences. In fact, it is equivalent to

$$\log h(t'|y) - \log h(t'|y') \geq \log h(t|y) - \log h(t|y'),$$

which means that the log-likelihood function has increasing differences.

Proof. We first show that if t' is more favorable than signal t then (7.3) must hold. Indeed, take a prior π so that $\pi(y) = 1 - \pi(y') = \frac{1}{2}$. Because t' is more favorable than t, distribution $F_\pi(.|t)$ is stochastically dominated by $F_\pi(.|t')$. Because $y' < y$, and because the conditional distributions F_π assign positive probability to only two states y and y', it must be that $f_\pi(y'|t') < f_\pi(y'|t)$. Because $f_\pi(y|s) = 1 - f_\pi(y'|s)$ for any s, we get $f_\pi(y|t) < f_\pi(y|t')$, which implies that

$$\frac{f_\pi(y|t)}{f_\pi(y'|t)} < \frac{f_\pi(y|t')}{f_\pi(y'|t')}.$$

Because the prior probabilities of the two states are equal, we have

$$\frac{f_\pi(y|t')}{f_\pi(y'|t')} = \frac{h(t'|y)}{h(t'|y')}, \text{ and } \frac{f_\pi(y|t)}{f_\pi(y'|t)} = \frac{h(t|y)}{h(t|y')},$$

which implies the MLR property.

Next, suppose that inequality (7.3) holds. We will show that t' is more favorable than t. In the proof, we assume for simplicity that the set of signals Y is finite. The result extends to the infinite case.

Take any prior distribution π. Recall that the conditional cdf is defined so that for each y, and $s = t', t$,

$$\begin{aligned} F_\pi(y|s) = \sum\nolimits_{y' \leq y} f_\pi(y'|s) &= \frac{\sum_{y' \leq y} h(s|y')\pi(y)}{\sum_{y'} h(s|y')\pi(y)} \\ &= \frac{1}{1 + \frac{\sum_{y'>y} h(s|y')\pi(y)}{\sum_{y' \leq y} h(s|y')\pi(y)}}. \end{aligned} \tag{7.4}$$

We need to show that the posterior cdf. $F_\pi(.|t')$ first-order dominates $F_\pi(.|t)$. By (7.3), for any $y' \leq \bar{y} \leq y$,

$$h(t'|y)\, h(t|y')\, \pi(y') \geq h(t|y)\, h(t'|y')\, \pi(y').$$

Thus, for any $\bar{y} \leq y$

$$h(t'|y) \sum\nolimits_{y' \leq \bar{y}} h(t|y')\pi(y') \geq h(t|y) \sum\nolimits_{y' \leq \bar{y}} h(t'|y')\pi(y')$$

which implies that

$$\frac{h(t'|y)}{\sum_{y' \leq \bar{y}} h(t'|y')\pi(y')} \geq \frac{h(t|y)}{\sum_{y' \leq \bar{y}} h(t|y')\pi(y')}.$$

Summing over ys such that $y > \bar{y}$ yields,

$$\frac{\sum_{y'>\bar{y}} h\left(t'|y'\right)\pi\left(y'\right)}{\sum_{y'\leq\bar{y}} h\left(t'|y'\right)\pi\left(y'\right)} \geq \frac{\sum_{y'>\bar{y}} h\left(t|y'\right)\pi\left(y'\right)}{\sum_{y'\leq\bar{y}} h\left(t|y'\right)\pi\left(y'\right)}.$$

Together with (7.4), this implies that

$$F_{\pi}\left(y|t'\right) \leq F_{\pi}\left(y|t\right).$$

□