Contents

1 Math Notation

Notation	*Meaning*
$\Longrightarrow$	implies.
$\Leftrightarrow$	implies and is implied by, also denoted iff (if and only if)
$\rightarrow$	denotes a mapping from the set on the left hand side (LHS) to the set on the right hand side (RHS).
$\Re$	the set of all real numbers.
$\Re^n$	n-dimensional Euclidean vector space.
$\exists$	there exists.
$\forall$	for all.
$\succ$	is strictly preferred to.
I_e	the indicator function for event e, so $I_e = 1$ in event e and otherwise $I_e = 0$.
$s \in S$	state s is an element of the set of possible states S.
	Example: $S = \{\text{rain, sun}\}$, $s = \text{rain}$.
π	probability distribution, $\pi : S \rightarrow [0, 1]$.
	Example: $\{\pi(\text{rain}) = 0.3, \pi(\text{sun}) = 0.7\}$
E_π	expected value over the probability distribution given by π
$a \in A$	action a is an element of the set of possible actions A.
	Example: $A = \{\text{carry umbrella, don't carry umbrella}\}, a = \text{carry umbrella}$.
$x \in X$	outcome x is an element of the set of possible outcomes $X \subset \Re^n$.
	oucome depends on state and action ($a : S \rightarrow X$ or $x(a, s)$).
	Example: $x(a, s) = $ dry but burdened by umbrella.
U	utility function, $U : X \rightarrow \Re$.
	For two possible outcomes $y, z \in X$, $y \succ z \Leftrightarrow U(y) > U(z)$.
	Preferences map outcomes to the set of real numbers via a utility function.
U_s	state dependent utility function. $U_s(x)$ is utility of outcome x in state s.

2 Static Decision Theory

Decision theory answers the basic question: "Which action $a \in A$ to choose?"

Begin with the simplest setting:

- a single once-and-for-all decision,
- among a finite set of alternative actions $A = \{a_1, ..., a_M\}$,
- given a finite number of alternative states $S = \{s_1, ..., s_N\}$
- with known probabilities $(p_1, ..., p_N)$.

The actions $a \in A$ represent alternative choices with typically uncertain outcomes, such as which pension plan to buy, if any, or where to live while looking for a job. The states $S \subset \Re^m$ represent uncertain events that the decider cares about, for example political outcomes; other examples include health events, or tomorrow's weather, or possible job offers. Actions affect the probabilities of each outcome, and can thus affect the set of possible outcomes (those with positive probability). For example, the action [move to Detroit] and the state [high demand for aeroponics] may produce, with positive probability, the outcome [hired as chief economist for Detroit aeroponics startup firm].

The outcome space can be 1-dimensional. Here the outcomes $X \subset \Re$ can be thought of as possible wealth increments or, a bit more generally, the monetary values that the decider assigns to possible outcomes. The goal for the next several subsections is to formulate the choice problem in the wealth increments case (the results generalize, but this case is the most useful), and to show how expected utility theory answers the "basic question."

2.1 Preferences over lotteries

Suppose that there is a given finite set of possible outcomes $x_1 > x_2 > ... > x_N \in \Re$, and that each action pins down a particular vector of associated probabilities $p_i = \pi(x_i), i = 1, ..., N$ over that given set of outcomes. Of course, the probabilities must be non-negative and sum to 1.0. The set of all possible probability vectors is called the *N-simplex* and denoted $\mathbf{\Sigma}^N = \{(q_1, ..., q_N) : q_i \geq 0, q_1 + ... + q_N = 1\}$.

Define a *lottery* $L = ([x_1, x_2, ..., x_N], [p_1, p_2, ..., p_N])$ as a finite list of monetary outcomes $[x_1, x_2, ..., x_N] \in \Re^N$ with corresponding probabilities $[p_1, p_2, ..., p_N] \in \mathbf{\Sigma}^N$. The *space of all lotteries* with a fixed list $[x_1, x_2, ..., x_N]$ of outcomes is denoted $\mathcal{L}[x_1, x_2, ..., x_N]$, or, if the list of outcomes is understood, just $\mathcal{L} = \mathbf{\Sigma}^N$ for short.

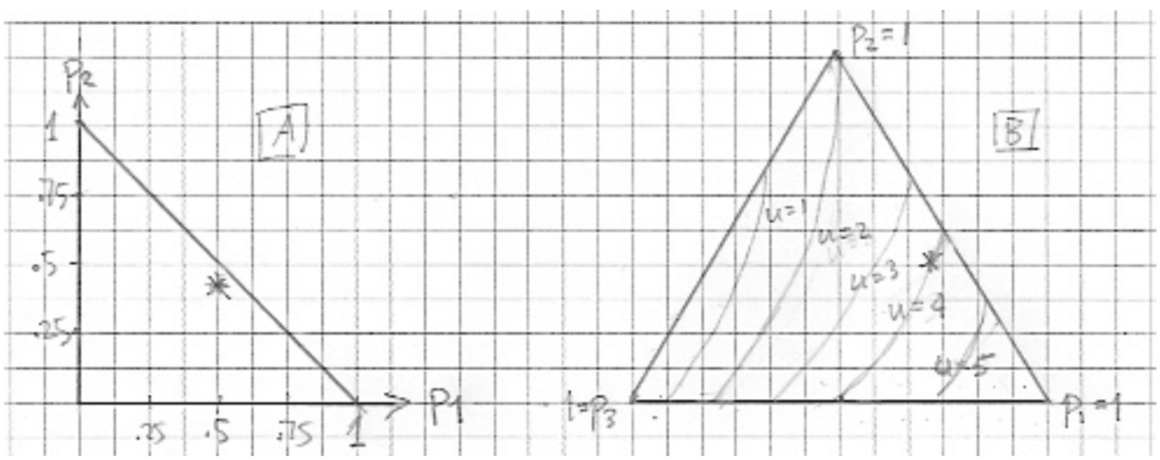

Figure 1: The N-Simplex for $N = 3$. These represent all lotteries over given states $x_1 > x_2 > x_3 \in \Re$. Panel A shows the first two probabilities using the usual coordinate system, with $p_3 = 1 - p_1 - p_2$ not shown explicitly. Panel B uses barycentric coordinates (in which lines of constant p_i are parallel to the edge where $p_i = 0$). The point $p^* = (.5, .4, .1)$ is marked by an asterisk (*) in both panels. Indifference curves for some (non-Bernoulli) utility function U are shown in Panel B.

The *expected value* of lottery L is $E_L x = \sum_{i=1}^{N} p_i x_i$. For example, the lottery $([10, 0, -1], [.1, .5, .4])$ has expected value $.1(10) + .5(0) + .4(-1) = 0.6$.

A utility function over monetary outcomes (henceforth called a *Bernoulli function*) is a strictly increasing function $u : \Re \rightarrow \Re$. Given Bernoulli function u, the *expected utility* of lottery $L = ([x_1, x_2, ..., x_k], [p_1, p_2, ..., p_k])$ is $E_L u = \sum_{i=1}^{k} p_i u(x_i)$. For example, if $u(x) = 1 - \exp(-2x)$, then the expected utility of the lottery just mentioned is $1 - [.5e^0 + .4e^2 + .1e^{-20}] \approx -2.46$.

Preferences $\succeq$ over any set refer to a complete and transitive binary relation. A utility function U *represents* preferences $\succeq$ if $x \succeq y \iff U(x) \geq U(y) \quad \forall\, x, y$. In particular, a utility function $U : \mathcal{L} \rightarrow \Re$ represents preferences $\succeq$ over $\mathcal{L}$ if

$$L \succeq L' \iff U(L) \geq U(L') \quad \forall\, L, L' \in \mathcal{L}. \tag{1}$$

Preferences $\succeq$ over $\mathcal{L}$ have the *expected utility property* if they can be represented by a utility function U that is the expected value of some Bernoulli function u. That is, there is some Bernoulli function u, such that for all $L, L' \in \mathcal{L}$, where $L \equiv ([x_1, x_2, ..., x_N], [p_1, p_2, ..., p_N])$ and $L' \equiv ([x_1, x_2, ..., x_N], [p'_1, p'_2, ..., p'_N])$, we have $L \succeq L' \iff$

$$U(L) \equiv E_L u \equiv \sum_{i=1}^{N} p_i u(x_i) \geq \sum_{i=1}^{N} p'_i u(x_i) \equiv E_{L'} u \equiv U(L'). \tag{2}$$

2.2 Expected utility theorem

It might seem that preferences with the expected utility property are quite special, and indeed they are. For example, their indifference surfaces are parallel and flat. Thus preferences over for lotteries with $k = 3$ given monetary outcomes but varying probabilities (as in Figure 1) must have indifference curves on the probability simplex that are all straight lines with the same slope. Thus the preferences indicated by the curved indifference curves in Panel B of the Figure do *not* have the expected utility property.

The expected utility theorem (EUT) is therefore surprising. It states that preferences over lotteries that satisfy a seemingly mild set of conditions will automatically satisfy the expected utility property, and thus be representable via a Bernoulli function.

Over the decades since the original results of Von Neumann and Morgenstern, many different sets of conditions have been shown to be sufficient. Mas-Colell et al. [2010] work with the following four conditions that an individual's preferences $\succeq$ on $\mathcal{L}$ should satisfy:

1. Rationality: Preferences $\succeq$ are complete and transitive on $\mathcal{L}$.

2. Continuity: The precise mathematical expressions are rather indirect (they state that certain subsets of real numbers are closed sets), but they capture the intuitive idea that U doesn't take jumps on the space of lotteries. This axiom rules out lexicographic preferences.

3. Reduction of Compound Lotteries: Compound lotteries have outcomes that are themselves lotteries in $\mathcal{L}$. One obtains the *reduced* lottery, a simple lottery in $\mathcal{L}$ over all possible outcomes in the final lotteries, using the probabilities obtained in the obvious way (multiplication). The axiom states that the person is indifferent between any compound lottery and the corresponding reduced lottery.

4. Independence: Let $L, L', L'' \in \mathcal{L}$ and $\alpha \in (0, 1)$. Suppose that $L \succeq L'$. Then $\alpha L + (1 - \alpha)L'' \succeq \alpha L' + (1 - \alpha)L''$. "In other words, if we mix two lotteries with a third one, then the preference ordering of the resulting two mixtures does not depend upon (is independent of) the particular third lottery used." (Mas-Colell et al. p. 171).

Theorem 1 (EUT) *Let preferences $\succeq$ on $\mathcal{L}$ satisfy axioms 1-4 above. Then $\succeq$ has the expected utility property, i.e., there is a Bernoulli function u such that (2) holds.*

As Mas-Colell et al. point out, all four axioms seem innocuous. Someone who cares only about the ultimate monetary payoffs and whose calculations are not affected by indirect ways of stating the probabilities will satisfy the third axiom. For example, such a person would be indifferent between the compound lottery "get 0 with probability 0.5 and with probability 0.5 play the lottery that pays 10 with independent probability 0.5 and 0 otherwise," and the reduced lottery "get 0 with probability 0.75 and 10 with probability 0.25." The fourth axiom enforces a degree of consistency by requiring that preference rankings over lotteries are not changed by nesting each of those lotteries within a generic compound lottery. The first two axioms are even less controversial or problematic.[1]

For a proof of the EUT, see Mas-Colell et al. and the references cited therein. Here is a sketch of how the function u can be constructed for given preferences. Denote by x_+ and x_- the maximum and minimum monetary outcomes in the lottery. (Given our convention on ordering the outcomes, $x_+ = x_1$ and $x_- = x_N$.) Set $u(x_+) = 1$ and $u(x_-) = 0$. Consider any other monetary outcome x, and the set of lotteries $\{([x_+, x_-], [p, (1-p)]) : p \in [0, 1]\}$. For $p = 1$ the lottery is preferred to x and for $p = 0$ the outcome x is preferred to the lottery. Using the continuity axiom, one can show that for some intermediate p^* the person is indifferent between x and that lottery. Set $u(x) = p^*$. Then use the other axioms to verify that the Bernoulli function so constructed indeed represents the given preferences.

2.3 Expected utility hypothesis (EUH)

The expected utility *theorem* (EUT) is a mathematical result with a rigorous proof. There can be no doubt that it is true. But what are the empirical implications?

The expected utility *hypothesis* (EUH) states that an actual person will choose as if maximizing the expectation of her own personal Bernoulli function. To state this with some generality beyond lottery choice, let $x(a, s)$ denote the outcome resulting from action $a \in A$ when the state of Nature turns out to be $s \in S$, and let u be a utility function

[1] On the other hand, the EUT's conclusion is quite strong, and not consistent with some actual choice data. One response is to accommodate some of the anomalous data by weakening the axioms, usually the third or fourth. For an extensive and skeptical review of these matters, see *Risky Curves* (Routledge, 2014) by D. Friedman, M. Isaac, D. James and S. Sunder.

defined over outcomes $x \in X$. Then

EUH: Each person has a Bernoulli function $u : X \to \Re$, and always chooses an action $a \in A$ that maximizes her expected utility, i.e., chooses an action that solves

$$\max_{a \in A} E_\pi u = \max_{a \in A} \sum_{s \in S} \pi(s) u \left[x(a, s) \right]. \tag{3}$$

Of course, the sum in equation (3) is replaced by an appropriate integral when the state space S is continuous.

In particular, if the action is to chose a lottery from $A = \mathcal{L}$, then the EUH would follow from the EUT if the four axioms hold, which seems reasonable. But how can we know the Bernoulli function for a particular person? The EUT proof sketch offers a suggestion: for each outcome x_i, vary the probabilities on best and worst outcomes to try to find a person's point of indifference. There are many variations on this theme; one of the currently more popular is the Multiple Price List scheme introduced by Holt and Laury (*American Economic Review*, 2002).

Of course, in many (if not most) choice problems in our uncertain world, the state probabilities π are not given to us. How can you then calculate (3) and apply the EUH?

The fallback position is called subjective expected utility (SEU). Leonard Savage (1954) first developed a list of sufficient conditions for the expected utility property, and others since have tweaked them. For example, Kreps' (1990) list includes asymmetry, negative transitivity, substitution and Archimedian — so far, all very plausible — plus a fifth condition noted below. That is, if choices over A satisfy Kreps' five conditions, then there is some some Bernoulli function $u : \Re \to \Re$ and some (subjective) probability distribution π such that (3) holds. This result is called the Strong SEU Theorem.

- "strong" refers to state-independent utility, a strong assumption that is not always valid.
- Example: utility for ice cream on a hot, sunny day is different than utility for ice cream in the winter.

Weak SEU allows for state dependent utility, i.e., the Bernoulli function u_s can vary with the state of the world s. Weak SEU then gives the expected utility property as

$$a \succ a' \Leftrightarrow E_\pi u_s\left[x(a,s)\right] > E_\pi u_s\left[x(a',s)\right] \tag{4}$$

The SEU Theorem says that Kreps' first four (very plausible) conditions are sufficient to ensure (4); here we drop the fifth condition, which basically just says that preferences are state-independent.

Before turning to other matters, we now mention something that will be helpful later in the course. The so-called *Harsanyi doctrine* connects the beliefs π across two or more decision makers:

- The π used in SEU comes via Bayes' Rule from a common prior belief and person-specific information.
- This means that differences in beliefs arise solely from differences in information to which different deciders have been exposed.
- As we will see later in this chapter, Bayes' rule is the only logically consistent way to combine new information with prior beliefs.
- Bayes' Rule (and, implicitly, the Harsanyi doctrine) is the foundation for "rational expectations" in macroeconomics.

2.4 Do people really behave this way?

Three questions:

Q1. Is EUH (or SEU) a good normative theory?

A1: These theories offer prescriptions for how best to achieve your goals. A rational person should indeed act as if maximizing the expected value of a Bernoulli function when the decision is stand-alone. Here's an important caveat: the prescription doesn't apply to *partial* decisions. In particular,

(a) portfolio effects. For example, if you look at the choice of which financial asset to choose in isolation, it would be irrational to pick gold because its returns are first-order stochastically dominated. (We'll see soon what that means.) But gold probably has a negative covariance with the rest of your portfolio (returns are not independent), so if you look at the overall decision of which entire portfolio to choose, you might rationally include some gold, depending on your Bernoulli function.

(b) path dependence. Your current action may affect the outcomes of future decisions. For example, the return to holding cash is first order stochastically dominated by returns to holding bonds, but its liquidity better enables you to cope with opportunities and risks that might turn up later. We'll deal with such matters in the next chapter.

Q2. Do actual people really behave according to EUH (or SEU)?

A2: Well, a few come pretty close, especially if they get expert advice, or have good MBA training. But most people tend to depart in some systematic ways, at least in unfamiliar situations. These departures are called choice anomalies, and we will discuss some later, when we return to the questionnaire distributed earlier.

Q3. As an economist, how should I react to choice anomalies?

A3: This is an active and controversial area of research.

(a) Orthodox behavioral economists seek generalized models of EUH (or SEU); a leading example is cumulative prospect theory.

(b) Hardline neoclassical economists believe that anomalies are ignorable because the markets reflect "smart money."

(c) The authors of these notes believe that under favorable conditions, deciders learn to behave more consistently with EUH (or Bayesian SEU), and so it is a mistake to work with static generalized (weaker) models, but it is also a mistake to ignore anomalies. They think that the way forward is to supplement neoclassical models with dynamic models of adaptation and learning.

2.5 A questionnaire

1. An urn contains 90 balls. 30 are red and 60 are some unknown mixture of blue and green. One ball from the urn is chosen at random.

Choice A: You receive $100 if the ball is red.

Choice B: You receive $100 if the ball is blue.

Your choice, A or B? ____

Choice C: You receive $100 the ball is red or green.

Choice D: You receive $100 if the ball is blue or green.

Your choice, C or D? ____

2. Choose between:

Choice E. You win $100,000 with probability 1.

Choice F. You win $500,000 with probability 0.10; $100,000 with probability 0.89; or $0 with probability 0.01.

Your choice, E or F? ____

Choice G. You win $100,000 with probability 0.11; or $0 with probability 0.89.

Choice H. You win $500,000 with probability 0.10; or $0 with probability 0.90.

Your choice, G or H? ____

3. Six hundred people in a village have a fatal disease.

Choice I: Treatment I will cure $\frac{1}{3}$ of the people for sure and $\frac{2}{3}$ will die.

Choice J: Treatment J potentially cures everyone with a probability of success of $\frac{1}{3}$.

Your choice, I or J? ____

2.6 Questionnaire "solutions"

- The first pair of choices constitute the Ellsburg Paradox (1961) *Quarterly Journal of Economics* 75: 643-69.
  - The most common choice pair is: $[A, D]$, that is $A \succ B$ and $D \succ C$.
  - This violates SEU, which implies (i) if $A \succ B$ then $C \succ D$ or (ii) if $B \succ A$ then $D \succ C$.
  - Here is why. Let r, b, and g be the individual's subjective probabilities for the balls in the urn. Note that we are not requiring that $r = \frac{1}{3}$, or even $r+b+g = 1$, or that we know anything about the Bernoulli function. Choosing A over B (if not indifferent) tells us

$$\begin{aligned} A \succ B &\Leftrightarrow ru(100) + bu(0) + gu(0) > ru(0) + bu(100) + gu(0) \\ &\Leftrightarrow (r - b)\,[u(100) - u(0)] > 0 \\ &\Leftrightarrow r > b, \end{aligned}$$

    since Bernoulli functions must be increasing so $[u(100) - u(0)] > 0$.
    Likewise, choosing D over C tells us

$$\begin{aligned} D \succ C &\Leftrightarrow ru(0) + bu(100) + gu(100) > ru(100) + bu(0) + gu(100) \\ &\Leftrightarrow (b - r)\,[u(100) - u(0)] > 0 \\ &\Leftrightarrow r < b, \end{aligned}$$

    contradicting the earlier conclusion.
  - Therefore, the choice combination A, D violates SEU. A similar argument shows that $[B, C]$ also violates SEU.
  - Explanation? People dislike the ambiguity of making choices under uncertainty. They dislike the possibility that they might have the odds wrong, so they go

with betting for or against red, the known probability. Choice A is betting for red, choice D is betting against red. Of course, that tendency is not rational, and potentially could be exploited by a clever outsider.

- The second pair of choices constitutes the Allais Paradox (1953) *Econometrica*
  - The most common choice pair is: $[E, H]$, that is $E \succ F$ and $H \succ G$.
  - $[E, H]$ violates SEU. SEU implies (i) if $E \succ F$ then $G \succ H$ or (ii) if $F \succ E$ then $H \succ G$.
  - Here's why. Write payoffs in thousands, to avoid all the ,000s. Then

    $$E \succ F \Leftrightarrow 1U(100) > 0.10U(500) + 0.89U(100) + 0.01U(0)$$

    subtracting $0.89U(100)$ from both sides we have:

    $$0.11U(100) > \ 0.10U(500) + 0.01U(0) \Leftrightarrow G \succ H$$

    Thus, the most common choice pair $[E, H]$ violates SEU. A similar argument shows that $[F, G]$ also violates SEU.
  - The second may be thought of as a compound lottery of the first:
    * Adding a 89% chance of 100 and subtracting a 89 chance of 0 to both G and H results in E and F respectively.
  - Why do people choose $[E, H]$? Explanations include
    * over-emphasizing the "chance" of small probability events. The 1% chance of zero in F gets "too much" weight in the decision
    * There is a great psychological loss if zero is obtained in F when 100 could have been had for sure (greedy by trying for the 500 ?)
  - Also, note the following expected values:

    * Expected value for E is 100 while expected value for F is 139 . Thus a risk-neutral or risk-preferring person will always choose F. A risk-averse person may choose either depending on the curvature of their utility function. Strongly risk-averse people will choose E while those with a risk premium less than 39 will choose F.
    * Expected value for G is 11 while expected value for H is 50 . The same reasoning applies, thus a risk-neutral or risk-preferring person will always choose H and a risk-averse person may choose either. If G is chosen the person must be risk-averse. So, the most-common choice pair $[E, H]$ violates SEU and implies that risk preferences change (convex then concave at higher income levels).

- The final choice is an example of a framing anomaly.
  - The expected outcome is exactly the same 200 live, 400 die.
  - The most common choice is J. Choice I contains the phrase "$\frac{2}{3}$ will die." This makes most people uncomfortable and they choose the more optimistic phrase "potentially cures everyone."
  - Note that a risk-averse person, which is commonly assumed, would choose I, since it has a higher expected value for a concave utility function. A risk-neutral person will choose either and a risk-preferring person will choose J.
- In the final section, Behavioral Considerations, we will say more about how to think about such choice anomalies. In the meantime, we will focus on "rational" decision makers with worrying about realism.

3 Useful Bernoulli Functions and Metrics

3.1 Cardinality

Inspecting (2) and similar equations, you can easily see that lottery rankings are unchanged if you replace the Bernoulli function $u(x)$ by $v(x) = u(x) - 2$ or by $v(x) = 0.5u(x)$, or indeed by any positive affine transform $v(x) = au(x) + b$, with $a > 0$. Economists say that such v are cardinally equivalent to u, and that they represent the same *cardinal preferences*.

Recall that, in non-risky choice, preferences are considered ordinal — if they are represented by the utility function u then, for any any smooth increasing transformation h (so $h' > 0$), then those preferences are also represented by $v(x) = h(u(x))$. But this doesn't work for Bernoulli functions.

An Example.

- Consider lottery L $= ([10, 4, 0], [.5, 0, .5]$, i.e., a 50% chance at 10, and lottery M $=([10, 4, 0], [0, 1, 0]$, i.e., 4 for sure.
- Suppose risk preferences are represented by $u(x) = x^{0.4}$. Then
  - $E_L u = (.5)10^{0.4} + (.5)0 \approx 1.26$
  - $E_M u = 4^{0.4} \approx 1.74 > 1.26$
  - Hence this person prefers M over L
- Now suppose risk preferences are represented by $v(x) = x^{0.8}$. Then
  - $E_L v = (.5)10^{0.8} + (.5)0 \approx 3.15$
  - $E_M v = 4^{0.8} \approx 3.03 < 3.15$.
  - Hence this person prefers L over M

Thus u and v represent different preferences, even though the two functions are ordinally equivalent via the transformation $h(y) = y^2$, which is strictly increasing over the range $[0, 10]$.

For several decades, economists criticized expected utility theory because it is not ordinal, but they eventually got over it. When it comes to evaluating lotteries, it is not enough to say that, for example, \$10 is better than \$4 which is better than \$0. Ordinal is not enough; you really have to know *how much* better. On the other hand, it doesn't affect comparisons of expected utility if you add or subtract a fixed amount from every monetary outcome, or if you change the scale (say from dollars to cents, or to a foreign currency at fixed exchange rates).

3.2 Useful utility functions

The function family $u(x|r) = x^{1-r}/(1-r)$ is called **CRRA with parameter** r.

- We just saw two special cases, for $r = .6$ and $r = .2$. (Actually, we dropped the constant factor $1/(1-r) > 0$, but that doesn't matter for cardinal preferences.)
- This function family works for outcomes $x > 0$ and parameters $r > 0$.
- The CRRA family is often used in macroeconomics, where the parameter r is chosen so as to align the model with data.
- Using L'Hospital's rule, you can show that in the limiting case $r \to 1$, this function takes the form $u(x|1) = \ln(x)$, the function that Daniel Bernoulli originally proposed in 1738 !
- The next subsection will help explain where the name CRRA comes from.

The function family $u(x|a) = 1 - e^{-ax}$ is called **CARA with parameter** $a > 0$.

- It is also often used in applied work, where the parameter a is fitted to the data.
- A higher value of a is interpreted as greater risk aversion, or more cautious preferences, as we will see in the next subsection,
- which also will explain why this family got its name.

3.3 Measuring Risk Aversion

Given a twice continuously differentiable Bernoulli function u, the *coefficient of absolute risk aversion* at monetary outcome $x \in (-\infty, \infty)$ is

$$A(x) = \frac{-u''(x)}{u'(x)}. \tag{5}$$

It is straightforward to verify that $A(x) = a$ in the CARA function family, i.e., A is a constant, independent of x. Indeed, CARA is an acronym for constant absolute risk aversion.

The *coefficient of relative risk aversion* at $x > 0$ is

$$R(x) = \frac{-u''(x)}{u'(x)} x = xA(x). \tag{6}$$

It will not surprise you to hear (but check this anyway) that $R(x)$ is constant for all functions in the CRRA family, including $\ln x$. Can you now decipher the CRRA acronym?

You should also check that the functions $A(x)$ and $R(x)$ are unaffected by positive affine transformations, that is, they are the same for $v(x) = \alpha u(x) + b$ as they are for u if $\alpha > 0$. So these are valid cardinal measures of risk preferences.

To gain some intuition about these risk measures, consider the linear Bernoulli function $w(x) = \alpha x + b$ with positive slope $\alpha > 0$. Then w represents the same cardinal preferences as $u(x) = x$. Of course, expected utility for u is exactly the same thing as expected value, and w must give the same rankings. Thus according to expected utility theory, a person with increasing linear utility will always choose the lottery with highest expected value, irrespective of variance. Such a person is said to be **risk-neutral**.

A risk-neutral person has linear utility, so $u''(x) = 0$. By equations (5) and (6), such a person has $A(x) = R(x) = 0$.

Higher values of A and R indicate greater aversion to risk. Why? Look at equations (5) and (6) again. They are sign-adjusted measures of concavity, appropriately normalized, and greater concavity means less willingness to accept risk. To spell this out,

- recall that $u'' < 0$ for a concave function, so the numerators are $-u'' > 0$.
- The denominators u' normalize u'' so that positive affine transformations have no effect, as you checked with u and v in the previous subsection.
- Greater normalized curvature implies a larger drop in utility for a risky lottery relative to a non-risky lottery with the same expected value.

The last point is illustrated in Figure 2. It uses the notation $EL = E_L x$ for the expected value of lottery L. There would be no gap between $u(EL)$ and $E_L u$ if the utility function were linear, but the gap would be very large if the utility function were tightly curved in the neighborhood of EL. That is, the tighter the curvature, the greater the drop in expected utility when facing a risky lottery compared to a certain payment of the lottery's expected value.

3.4 A CRRA example

Insert an example here of how one reads off R from choices in a budget line task for given p and probs.

3.5 Measuring Risk

The risk aversion metrics A and R tell us something about Bernoulli functions, but not about the risks deciders face. Is there any nice metric for capturing the amount of risk that is inherent in a lottery L?

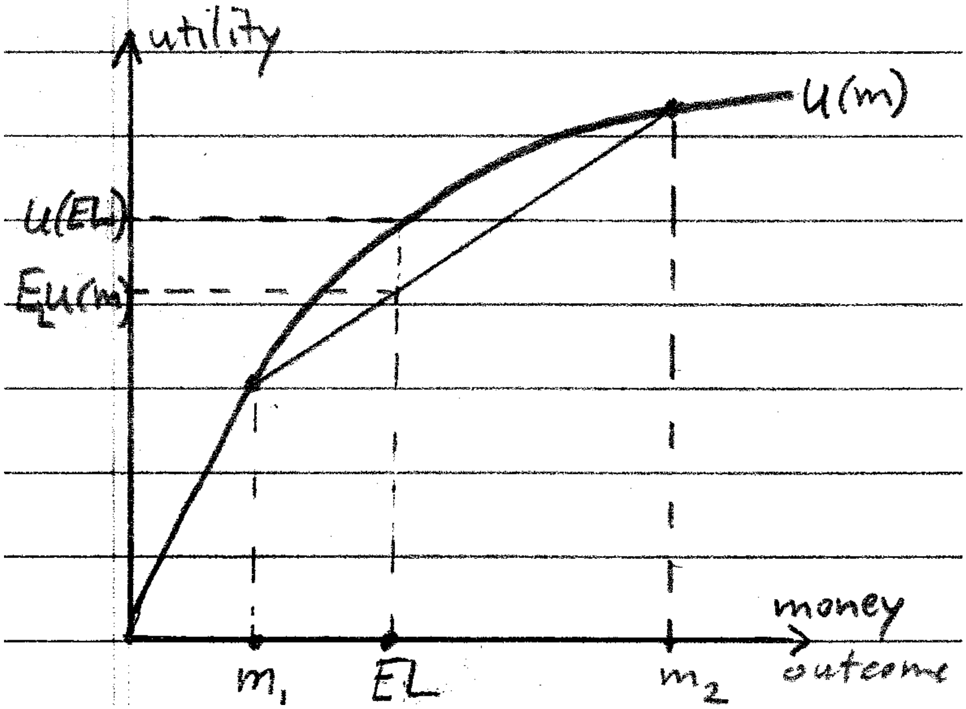

Figure 2: Expected Utility. Here the lottery prizes are $m_1 = 20$ and $m_2 = 90$ with probabilities $p_1 = .3$ and $p_2 = .7$; the expected value is $EL = 69$. The expected utility $E_L u = p_1 u(m_1) + p_2 u(m_2)$ of the lottery is the height of the point above EL on the line segment connecting the points $(m_1, u(m_1))$ and $(m_2, u(m_2))$ on the utility curve. The utility of getting (for sure) the expected value of the lottery is $u(EL)$. The loss of utility due to the risk of the lottery is the gap between $u(EL)$ on the u-curve and the expected utility $E_L u$ on the line segment. This gap is larger the greater the degree of concavity (curvature) in the utility function u.

1. The best known such metric is variance: $Var_L[x] = E_L(x - E_L x)^2 = \sum_{i=1}^{N} p_i (x_i - \bar{x})^2$, where $\bar{x} = E_L x = \sum_{i=1}^{N} p_i x_i$ is the lottery's expected value.

2. Sometimes it is convenient to use the closely related metric, the standard deviation $\sigma_L = \sqrt{Var_L[x]}$.

To see how these risk metrics work with our preference metrics, recall that by Taylor's Theorem, any smooth (continuously differentiable 3 times) utility function u can be expanded at any point z in its domain as a quadratic function plus remainder:

$$u(z+h) = u(z) + u'(z)h + \frac{1}{2}u''(z)h^2 + R^3(z,h), \tag{7}$$

where $R^3(z,h) = \frac{1}{6}u'''(y)h^3$ for some point y between z and $z+h$.

Given a lottery $L = (\mathbf{x}, \mathbf{p})$, put $z = \bar{x} = E_L x$ and $h = x - z$ (the deviation from the mean) in equation (7), and take the expected value of both sides. The second (linear) term disappears because $Eh = E(x - \bar{x}) = \bar{x} - \bar{x} = 0$. Since $Eh^2 = Var_L[x]$, the equation becomes

$$E_L u(x) = u(\bar{x}) + \frac{1}{2}u''(\bar{x})Var_L[x] + ER^3. \tag{8}$$

That is, expected utility of the lottery is equal to the utility of the mean outcome, plus a term proportional to the variance of the lottery and to the second derivative of u evaluated at the mean of the lottery, plus a remainder term.

- The next-to-last term in equation (8) is key. It says that variance reduces expected utility to the extent that u is concave, as measured by (unnormalized) $A(x)$. Otherwise put, a person with higher A will be more averse to variance than another rational person with A closer to zero. If $A(x) = 0$ for all x, then the Bernoulli function is linear and that person is risk neutral.

- We can ignore the remainder term if either

  1. h is small because all likely outcomes s are near $\bar{s}$, or

  2. $u'''(\bar{s})$ is small because u'' is almost constant in the neighborhood of $\bar{s}$.

  In other words, the approximation using the first two terms is reliable if either the risks are small, or the utility function is locally nearly quadratic. A separate argument shows that the approximation is exact if the probability distribution is Normal (aka Gaussian).

- Using a suitable affine transformation and ignoring the remainder term, expected utility takes the form $\bar{x} - cVar_L[x]$, or in other notation $\bar{x} - c\sigma_L^2$, where c is proportional to the coefficient of absolute risk aversion.

- A lot of finance literature assumes directly that utility takes this mean-variance form. In many important cases it is a good approximation, but it can be misleading for highly risky non-Normal lotteries. See homework problem I.2.

More sophisticated risk measures use the lottery's cumulative distribution function (cdf) $F : \Re \rightarrow [0,1]$. The function F is non-decreasing and $F(x) = \sum_{i=1}^{N} p_i I_{[s_i \leq x]}$ is the probability that the actual state does not exceed x. For a continuous range $[a,b]$ of possible outcomes with probability density $f(s)$, the cdf is $F(x) = \int_a^x f(s)ds$.

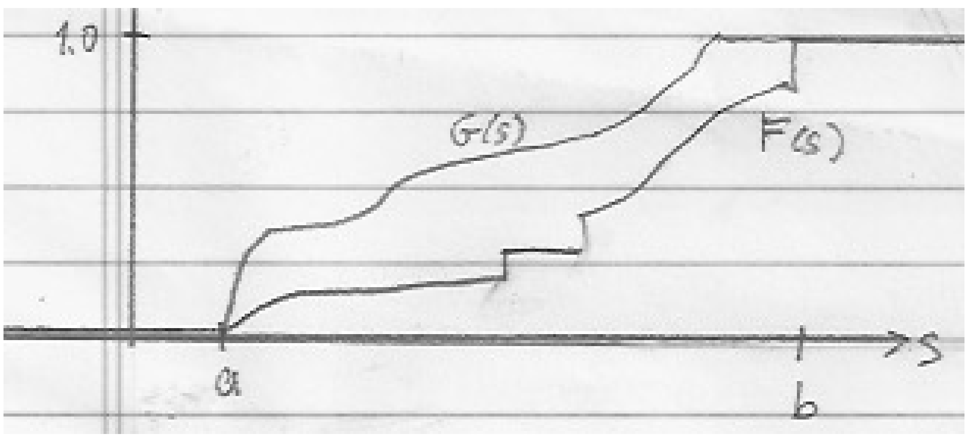

Figure 3: First order stochastic dominance. F is the cdf of a mix of a continuous random variable with a lottery on wealth increments $x = b$ and two other wealth increments in between a and b; while G is the cdf of a random variable with positive density over the interval $[a,c]$ with $c < b$. Every decider would prefer F because outcomes better than s are always at least as likely for F than G, for any $x \in \Re$. in next version of the figure, replace the s's by x's.

Suppose that lottery L has cdf F and lottery M has cdf G.

1. L is said to *first order stochastically dominate* M if $F(x) \leq G(x) \quad \forall x \in \Re$. In this case, write $L \succeq_{FOSD} M$.

- It is not especially difficult to show (e.g., Mas-Colell p. 195-199) that if $L \succeq_{FOSD} M$ then $E_L u \geq E_M u$ for any Bernoulli function u. That is, every coherent decider, no matter their risk preferences, prefers lottery L over M.

- The converse is also true: if expected utility rankings are the same irrespective of Bernoulli function, then the preferred lottery FOSD's the alternative.

- The inequality seems backwards at first, but makes sense when you think about it: a smaller value of the cdf at a given point means that there is a greater probability of larger monetary payoffs.

- Geometrically, FOSD means that the graph of the cdf for the better lottery is always below (or to the right of) that of the alternative lottery, as illustrated in Figure 3.

- Of course, if two cdf's cross each other, as they often do, then the two corresponding lotteries can't be ranked by FOSD.

2. Second order stochastic dominance (SOSD) can sometimes rank lotteries that are not ordered via FOSD.

   - Write $L \succeq_{SOSD} M$ if (a) $EL = EM$ and (b) $\int_{-\infty}^{x} G(t)dt \leq \int_{-\infty}^{x} F(t)dt \ \ \forall x \in \Re$.

   - That is, the lotteries have the same mean, but the area under the cdf for L up to x never exceeds that for M.

   - $L \succeq_{SOSD} M$ is equivalent to saying that the lottery M can be obtained from the lottery L via a "mean-preserving spread" that moves probability mass away from the mean while keeping the mean constant.

   - $L \succeq_{SOSD} M$ is also equivalent to the condition $E_L u \geq E_M u$ for any *concave* Bernoulli function u. That is, all risk-averse deciders prefer L to M. See Mas-Colell again.

3. Sometimes the worst possible outcome in a lottery L is better than the best possible outcome in some other lottery M. Although this is not standard terminology, we will say that L zeroth order dominates, and write $L \succ_{OOSD} M$. Some recent literature has referred to OOSD as "obvious domination." Of course, $L \succ_{OOSD} M$ automatically implies that $L \succ_{FOSD} M$.

4 Behavioral Considerations

What should we make of the empirical fact that the neoclassical models we study are sometimes inaccurate descriptions of actual human behavior?

The first point is that the models are "normative": they analyze what people *should* do in simple worlds.

But a second point is equally important. The models also have a "positive" role: they predict what people will do in actual situations.

Those predictions are useful for two reasons

- they sometimes are pretty accurate, and thus can guide policy, business plans, etc.
- other times they are inaccurate, but still provide an important benchmark.

That benchmark is important because it allows us to measure how large and consistent are deviations from prediction, enabling clearer formulation of new theory.

New theories arising this way generally go under rubrics of *behavioral economics* and *behavioral game theory.* Most chapters of these notes include a closing section with remarks and pointers to recent work analyzing the predictive accuracy of the models, and of behavioral models.

Later editions of these notes may include references to the fast-developing literature on choice anomalies, following up the discussion in Section 2.4 above.

5 Readings

[to be added in the next edition] new cites

Savage, Leonard J. 1954. The Foundations of Statistics. New York, Wiley.

Camerer, Colin. Behavioral Game Theory book.