The von Neumann–Morgenstern expected utility representation
Given our measurable space $(E, \mathcal{E})$, we now define some preliminary notions:

Definition. For each $x \in E$, the measure δ_x, defined by the rule

$$\delta_x(A) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

is called the Dirac point mass at x. A probability measure measure μ of the form

$$\mu = p_1 \delta_{x_1} + \ldots + p_n \delta_{x_n}$$

where $x_1, \ldots, x_n \in E$ and $p_1, \ldots, p_n > 0$ and $p_1 + \ldots + p_n = 1$ is called simple.

Proposition. Suppose that $U_0 : \mathcal{P} \to \mathbb{R}$ is affine. Define a function $U : E \to \mathbb{R}$ by $U(x) = U_0(\delta_x)$. Then for every simple probability measure μ we have the representation

$$U_0(\mu) = \int_E U(x)\mu(dx).$$

Proof. It follows from the definition of affine and induction that

$$\begin{aligned} U_0(\mu) &= p_1 U_0(\delta_{x_1}) + \ldots p_n U_0(\delta_{x_n}) \\ &= p_1 U(x_1) + \ldots p_n U(x_n) \\ &= \int_E U(x)\mu(dx). \end{aligned}$$

□

The following corollary provides a justification for the expected utility hypothesis:

Corollary. Let $\succ$ be a preference relation satisfying transitivity, completeness, independence and continuity. Then there exists a function $U : E \to \mathbb{R}$ such that for simple measures λ and μ we have

$$\lambda \succ \mu \text{ if and only if } \int_E U(x)\lambda(dx) > \int_E U(x)\mu(dx).$$

The conclusion is that in order to compare simple measures λ and μ, it enough to compare the integrals (expected values) of the *utility function* U. Hence, the von Neumann–Morgenstern axioms more-or-less justify the expected utility hypothesis. Note the above corollary is not quite everything we want since it only applies to simple measures.[1] Nevertheless, we will ignore this technicality and adopt the expected utility hypothesis from now on (unless we explicitly abandon it to look at other problems).

Properties of utility functions

We have seen that there exists that an axiomatic framework that justifies Bernoulli's idea that preferences are determined by expected utility. Now, we consider what properties we should insist that the function U have. In this section, we will let $E = \mathbb{R}$, so we can think of our experiment as a random payment of money at some fixed future date. Later in the course, E maybe a sequence space modelling a sequence of payments, or even E may be all (suitably regular) functions on $[0, \infty)$ when we have continuous streams of money.

Axiom 5. (Preference for more-to-less) For all $x > y$, we have $\delta_x \succ \delta_y$. In terms of the expected utility representation, this means $U(x) > U(y)$; i.e. U is increasing.

Axiom 6. (Risk aversion) For an integrable random variable X, we have[2]

$$\delta_{\mathbb{E}(X)} \succcurlyeq \mathbb{P} \circ X^{-1}.$$

In terms of expected utility, we have $U(x) \geq \mathbb{E}[U(X)]$; i.e. U is concave.[3]

Unless other stated, all utility functions in this course will be increasing and concave. Now to get some intuition about utility functions, we do a formal calculation. Suppose you have a certain amount x of money and you are offered two payments

- A certain payment of y units of money, or if y is negative, a certain loss of $|y|$ units of money.
- A payment of a random variable Y, where $\mathbb{E}(Y) = 0$.

[1]Suppose that E is separable metric space. Then the simple measures are dense in $\mathcal{P}$ with respect to the topology of weak converence. If we assume that U_0 is continuous and bounded, then we can extend the representation

$$U_0(\mu) = \int_E U(x)\mu(dx).$$

to all $\mu \in \mathcal{P}$. What must we assume about the preference relation $\succ$ to guarantee that the affine numerical representation U_0 is continuous? or bounded? In fact, things are slightly more subtle since we are really interested in cases where the utility function U is *unbounded.*

Fortunately, with there exist several alternative axioms, when added to the four other axioms, that yield the desired integral represenation of $\succ$, even when U is unbounded. See, for instance, Chapter 2 of *Stochastic Finance: An Introduction in Discrete Time* by Fölmer and Schied.

[2]A clearer way to write this is

$$\mathbb{E}(X) \succcurlyeq X$$

but this really is an abuse of notation. The idea is that given a choice between a random payment X and a certain payment of $\mathbb{E}(X)$, the certain payment is preferred.

[3]This supplements Bernoulli's argument of decreasing marginal utility.

How big, or small, must y be so that you are indifferent between these two options? We consider the case where y and Y are very small and do a Taylor series expansion:

$$\begin{aligned} U(x+y) &= \mathbb{E}[U(x+Y)] \\ U(x) + U'(x)y &\approx \mathbb{E}[U(x) + U'(x)Y + \tfrac{1}{2}U''(x)Y^2] \\ &= U(x) + \tfrac{1}{2}U''(x)\mathrm{Var}(Y) \end{aligned}$$

so the answer is given, roughly, by

$$y \approx \frac{1}{2}\frac{U''(x)}{U'(x)}\mathrm{Var}(Y).$$

Note that y is negative since U is increasing and concave. The point of this calculation is to justify the following definition:

Definition. Given a smooth utility function U, the Arrow–Pratt coefficient of absolute risk aversion is the function

$$= -\frac{U''(x)}{U'(x)}.$$

Examples.

(1) Constant absolute risk aversion (CARA):

$$U(x) = -e^{-\gamma x}$$

for a constant $\gamma > 0$.

(2) Constant relative risk aversion (CRRA):

$$U(x) = \frac{x^{1-R}}{1-R}$$

for a constant $R > 0$ and $R \neq 1$.

(3) The $R = 1$ case of CRRA

$$U(x) = \log x$$