The expected utility hypothesis

The aim of this section is to describe the so-called expected utililty hypothesis and give a plausible argument why expected utility maximisation is the right problem to study.

The Saint Petersburg paradox (usually attributed to Nicolas Bernoulli in 1713) is this. Suppose I offer you a choice:

(1) I will flip a coin repeatedly, until the first time it lands heads up. I will then pay you 2^n units of money if the coin first comes up heads on the nth flip.
(2) I will pay you x units of money.

How big[1] does x have to be before you would prefer (2) to (1)? You may first think to compute your expected payout in choice (1). However, since the probability of the first heads appearing on the nth flip is 2^{-n}, your expected payout is

$$\sum_{n\geq 1} 2^n \times 2^{-n} = \infty.$$

Does this mean you would always prefer choice (1) to choice (2)? This can't be right! It is probably safe to say most people would prefer (2) to (1) if x is sufficiently large. Indeed, if $x = 10^6$ then I certainly would prefer (2) to (1).

Daniel Bernoulli in 1738 argued that there is function U that measures the happiness, or utility, derived from his wealth. That is to say, a wealth of x corresponds to $U(x)$ units of utility. The function U is clearly increasing, but Bernoulli argued that the marginal utility

$$\frac{U(x+\varepsilon) - U(x)}{\varepsilon} \approx U'(x)$$

of a small amount ε should be decreasing in the level of wealth x; that is, the utility function U should be concave. Bernoulli went on to argue that his marginal utility is inversely proportional to level of wealth, so that

$$U'(x) = \frac{a}{x}$$

for some $a > 0$. Thus, Bernoulli's utility function is

$$U(x) = a \log x + b$$

logarithmic.

The next step of his argument is to invoke the *expected utility hypothesis*: an economic agent's preferences for a random payment are determined by the expected value of the utility of the payment. For someone with logarithmic utility, the expected utility of the payout[2] of

[1] In lecture I asked, 'how much would you be willing to pay for choice (1)?' This is nearly the same question. Indeed, if you are willing to pay x for choice (1), then you would be willing to forgo x units of certain money for the random payout. But please don't worry too much over the details here since this section is only intended for motivation.

[2] we are assuming here that he has no other wealth aside from the payout of the game

choice (1) is

$$\sum_{n\geq 1} a(\log 2^n)2^{-n}\log 2 + b = 2a\log 2 + b.$$

Now, Bernoulli's expected utility of receiving a certain payment of x is

$$a\log x + b.$$

Hence, if $x > 4$ Bernoulli would prefer choice (2), and if $x = 4$ he would be indifferent.

In this course, we will consider the problem of maximising the expected utility of an economic agent. There are (at least) two points of view:

(1) We might think that expected utility maximisation is a realistic description of economic behaviour, and hence the solution of this problem would lead to predictions about the economy.
(2) We might think that this constitutes *rational* economic behaviour, and hence the solution is merely a suggestion of what one ought to do.

This course will remain agnostic on this philosophical issue[3].

Before we begin studying utility maximisation problems, let's first think a little about Bernoulli's argument. Now, even if you are happy assuming people have utility functions which they consult to find out how happy they are with a fixed (non-random) payment, you might wonder why, when faced with a random payment, we should care about the expected value of the utility. At this point, the expected utility hypothesis seems a bit *ad hoc*. Fortunately, there is an axiomatic framework for preferences which shows that the expected utility hypothesis is natural. This is the subject of the next section.

The von Neumann–Morgenstern theorem

Consider a measurable space $(E, \mathcal{E})$, i.e. a set E and sigma-field $\mathcal{E}$ of subsets of E. The set E models the possible outcomes of a random experiment, and the sigma-field $\mathcal{E}$ models the measurable events. In most cases considered in this course (including the example from last section) the set of outcomes will be a set of possible payouts of an investment, so we can identitify E with the real line $\mathbb{R}$, and the sigma-field $\mathcal{E}$ with is the Borel sigma-field. But for the sake of generality, we will not impose any structure on E now.

Rather than considering preferences over possible outcomes of the experiment, i.e. elements of E, the idea here is to assign preferences over the collection $\mathcal{P}$ of *probability measures*[4] on the measurable space $(E, \mathcal{E})$.

First we introduce a symbol $\succ$ which should read as *is preferred to*. For instance, if $\lambda, \mu \in \mathcal{P}$ are probability measures, then

$$\lambda \succ \mu$$

means that the hypothetical person prefers the random outcome of the experiment drawn according the probability distribution λ to an outcome drawn from the distibution μ.

[3]However, it should be noted that though there is now much empirical evidence that view (1) is not entirely accurate. Google the phrase *experimental economics*

[4]In the preference literature, the terms *lottery* and *gamble* are often used in place of probability measure.

Now we introduce some more notation which is built from $\succ$.

$$\begin{array}{ll} \lambda \prec \mu & \text{means } \mu \succ \lambda \\ \lambda \not\prec \mu & \text{means that } \lambda \prec \mu \text{ is not true} \\ \lambda \not\succ \mu & \text{means that } \lambda \succ \mu \text{ is not true} \\ \lambda \sim \mu & \text{means that both } \lambda \not\prec \mu \text{ and } \lambda \not\succ \mu \text{ are true} \end{array}$$

Von Neumann and Morgenstern introduced the following axioms for the relation $\succ$.

Axiom 1 (Transitivity) $\lambda \succ \mu$ and $\mu \succ \nu$ implies $\lambda \succ \nu$.

Axiom 2 (Completeness) For two probability measures $\lambda, \mu \in \mathcal{P}$ we have exactly one of the following possibilities:

$$\lambda \succ \mu, \quad \mu \succ \lambda \text{ or } \lambda \sim \nu$$

Axiom 3 (Independence) If $\lambda \succ \mu$ then

$$p\lambda + (1-p)\nu \succ p\mu + (1-p)\nu$$

for any $0 < p < 1$ and $\nu \in \mathcal{P}$.

Axiom 4 (Continuity) If $\lambda \succ \mu \succ \nu$ then there exists a $0 < p < 1$ such that

$$\mu \sim p\lambda + (1-p)\nu$$

Our goal is to understand what structure these axioms give to the preference relation. To write down a clean theorem, let's now introduce some definitions:

Definition. A function $U_0 : \mathcal{P} \to \mathbb{R}$ is called a numerical representation of the preference relation $\succ$ iff

$$\lambda \succ \mu \text{ if and only if } U_0(\lambda) > U_0(\mu).$$

Definition. A function $U_0 : \mathcal{P} \to \mathbb{R}$ is called affine if

$$U_0(p\lambda + (1-p)\mu) = pU_0(\lambda) + (1-p)U_0(\mu)$$

Theorem. (von Neumann and Morgenstern 1947) A preference relation $\succ$ satisfies Axioms 1, 2, 3 and 4 if and only if there exists an affine numerical representation U_0 of $\succ$.

Furthermore, if an affine numerical representation U_0 of $\succ$ exists it is unique in the sense that if V_0 is another affine numerical representation then

$$V_0 = aU_0 + b$$

for some constants $a > 0$ and $b \in \mathbb{R}$.

Proof. Given an affine numerical representation U_0 of $\succ$, it is an exercise to show that $\succ$ satisfies Axioms 1, 2, 3 and 4.

So suppose that $\succ$ satisfies Axioms 1, 2, 3 and 4. We will need a lemma: If $\lambda \succ \mu \succ \nu$ then there exists a *unique* $0 < p < 1$ such that

$$\mu \sim p\lambda + (1-p)\nu.$$

The proof of this lemma is sketched on the example sheet.

We need to build a numerical representation U_0. If $\lambda \sim \mu$ for every pair $\lambda, \mu \in \mathcal{P}$ there is nothing to show. So fix two probability measures such that

$$\lambda_1 \succ \lambda_0.$$

Assign $U_0(\lambda_1) = 1$ and $U_0(\lambda_0)$. Now for every other probability measure μ, let

$$U_0(\mu) = \begin{cases} 1/p & \text{if } \mu \succ \lambda_1 \succ \lambda_0 \quad \text{and } \lambda_1 \sim p\mu + (1-p)\lambda_0 \\ 1 & \text{if } \mu \sim \lambda_1 \\ q & \text{if } \lambda_1 \succ \mu \succ \lambda_0 \quad \text{and } \mu \sim q\lambda_1 + (1-q)\lambda_0 \\ 0 & \text{if } \mu \sim \lambda_0 \\ -r/(1-r) & \text{if } \lambda_1 \succ \lambda_0 \succ \mu \quad \text{and } \lambda_0 \sim r\lambda_1 + (1-r)\mu \end{cases}$$

It is an exercise to verify that U_0 is an affine numerical representation of $\succ$.

Now suppose that V_0 is another affine numerical representation of $\succ$. Define constants a and b by

$$a = V_0(\lambda_1) - V_0(\lambda_0) \text{ and } b = V_0(\lambda_0).$$

Since $\lambda_1 \succ \lambda_0$ we have that $a > 0$. Note that if $\lambda_1 \succ \mu \succ \lambda_0$ then there is a $0 < q < 1$ such that $\mu \sim q\lambda_1 + (1-q)\lambda_0$. Hence

$$\begin{aligned} V_0(\mu) &= qV_0(\lambda_1) + (1-q)V_0(\lambda_0) \\ &= a[qU_0(\lambda_1) + (1-q)U_0(\lambda_0)] + b \\ &= aU_0(\mu) + b \end{aligned}$$

The other cases are left as exercises. □