ApEc 8001
Applied Microeconomic Analysis: Demand Theory

Lecture 10: Expected Utility Theory
(MWG, Ch. 6, pp.167-183)

I. Introduction

So far we have assumed that there is no uncertainty of any kind, but of course in the real world many things are uncertain. Economists have long known that consumers (and producers) face many kinds of risk and uncertainty. In the 1940s and 1950s they developed the **first formal theory of behavior under uncertainty**, which is known as **expected utility theory**. This lecture will explain this theory, and the next two lectures will cover additional topics concerning choice under uncertainty.

II. How Economists Define Risky Alternatives

To start, consider a "decision maker" who has to choose from among a number of different alternatives, some or all of which have uncertain outcomes.

Formally, **denote the set of all possible outcomes by C**. These "outcomes" could be almost any object of choice, not just bundles of goods and services.

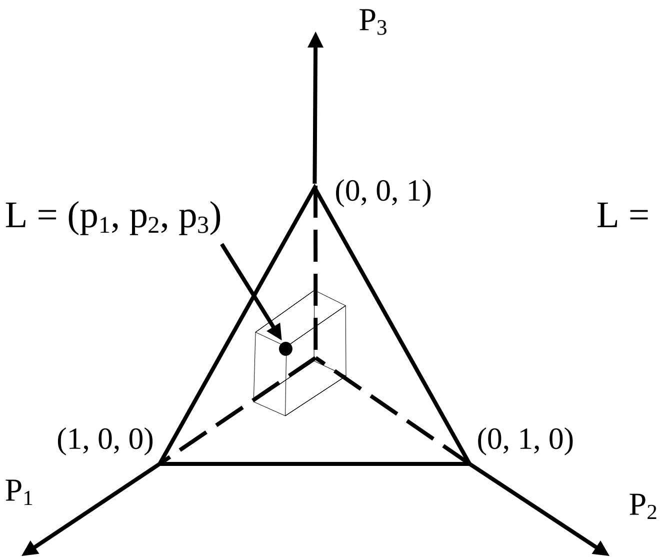
Assume that the number of possible outcomes in C is finite, and we assign an index to them of n = 1, 2, … N.

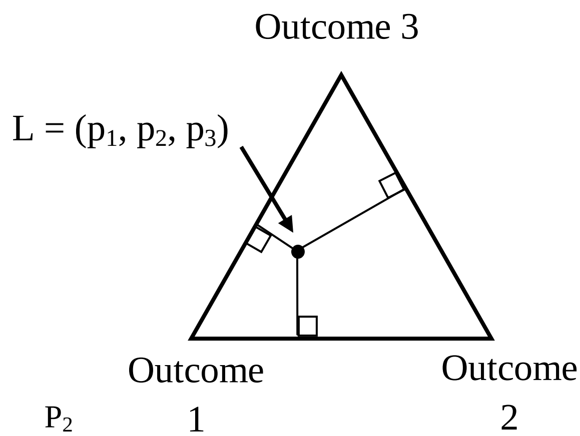
Each of these outcomes has a probability of occurring, and these probabilities are denoted by $p_1, p_2, \ldots p_N$. As will be seen shortly, the decision maker's choices may result in a different set of probabilities. An **important assumption** that we will maintain throughout this lecture is that these **probabilities are objectively known**. That is, all decision makers know the values of all the p's, and how their choices could lead to different values of the p's.

The **basic building block of expected utility theory** is the concept of a **lottery**, of which there are **two types**:

Definition 6.B.1: A **simple lottery L** is a list (**vector**) of the form $L = (p_1, p_2, \ldots p_N)$ with $p_n \geq 0$ for all n and $\sum_{n=1}^{N} p_n = 1$.

Each possible simple lottery can be represented geometrically as a point in an N-1 **dimensional simplex**, $\Delta = \{p \in \mathbb{R}_+^N : p_1 + p_2 + \ldots p_N = 1\}$. This can be seen for the N = 3 case in the following diagram:





The triangle in this diagram depicts all possible sets of probabilities for the N = 3 case (all values of p_1, p_2 and p_3 such that $p_1 + p_2 + p_3 = 1$). Note that the triangle itself is a two dimensional (N – 1) object.

Question: Suppose that the second outcome (n = 2) is a "sure thing" that will occur with 100% certainty. Where is the associated lottery L in this diagram? Or is it not in the diagram?

Sometimes it is useful to have some or all of the outcomes of a lottery to be simple lotteries. There are called compound lotteries, and they can be defined as follows:

Definition 6.B.2. Given K simple lotteries, **each of which is denoted by** $L_k = (p^k_1, p^k_2, \ldots p^k_N)$, k = 1, 2, … K, and probabilities $\alpha_k \geq 0$ with $\sum_{k=1}^{K} \alpha_k = 1$, the **compound lottery** $(L_1, L_2, \ldots L_K; \alpha_1, \alpha_2, \ldots \alpha_K)$ is the risky alternative that yields the simple lottery L_k with probability α_k for k = 1, 2, … K.

Example. Upon graduation you have a chance of getting one of 5 similar jobs. Each job has probabilities for: 1. Staying in Minnesota; 2. Moving to another state in the U.S.; and 3. Moving to another country.

For any compound lottery $(L_1, L_2, \ldots L_K; \alpha_1, \alpha_2, \ldots \alpha_K)$ we can work out the corresponding **reduced lottery** that is a simple lottery that generates the identical "ultimate" distribution of outcomes. The **value of each p_n in this reduced lottery** can be calculated by multiplying the probability that each lottery L_k has, α_k by the probability $p^k{}_n$ for outcome n in lottery L_k, and then adding these $p^k{}_n$'s over all values of k. That is, the probability of outcome n in the reduced lottery **is**:

$$p_n = \alpha_1 p^1{}_n + \alpha_2 p^2{}_n + \ldots + \alpha_K p^K{}_n \quad \text{for } n = 1, 2, \ldots N$$

Thus the reduced lottery L of any compound lottery $(L_1, L_2, \ldots L_K; \alpha_1, \alpha_2, \ldots \alpha_K)$ can be obtained by vector addition (each L is a vector of p's):

$$L = \alpha_1 L_1 + \alpha_2 L_2 + \ldots \alpha_K L_K \in \Delta$$

The expression "$\in \Delta$" indicates that, just as each simple lottery is a member of the N-1 "simplex", the reduced lottery that is a weighted average of K simple lotteries is also a member of this simplex.   For the N = 3 case, this can be seen in the following diagram:

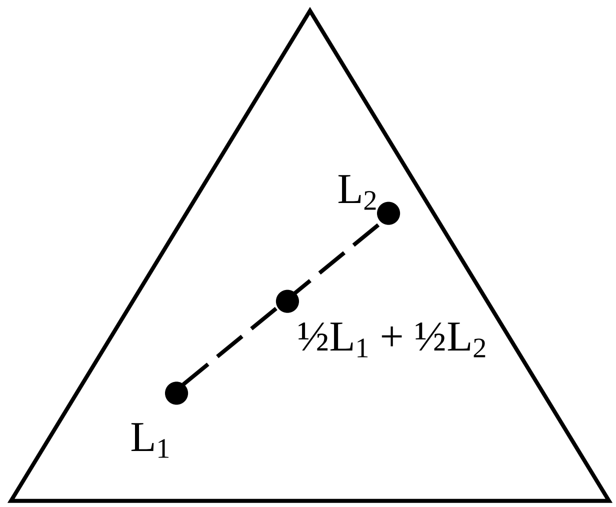


Figure 6.B.3 in MWG (p.170) gives an example of how 2 different compound lotteries lead to the same simple lottery.

III. Preferences over Lotteries

Now that we have a clear definition of risk (in terms of lotteries) we can define preferences over these types of risk, which is simply preferences over lotteries. To begin **assume** is that the **decision maker cares only about the reduced lottery** over the N possible outcomes. This assumption is called the **consequentialist hypothesis**.

To start, **let $\mathscr{L}$ denote the set of all simple lotteries over the set of outcomes C.** Assume that the **decision maker has a rational** (i.e. complete and transitive) **preference relation** $\succsim$ that applies to **all elements of the set $\mathscr{L}$**.

In general, this assumption of a rational preference relation is a stronger assumption than rationality in the "choice with no uncertainty" scenario that we have maintained in previous lectures. Roughly speaking, it is a stronger assumption because people probably find it harder to make decisions in the more complex scenario of uncertainty and lotteries.

In fact, to derive useful results we need to assume that preferences over lotteries are **continuous** and satisfy the **independence axiom**. So let's define those two concepts:

Definition 6.B.3. The preference relation $\succsim$ on the space of simple lotteries $\mathscr{L}$ is **continuous** if, for any L, L′ and L″ $\in \mathscr{L}$, the following two sets:

$$\{\alpha \in [0, 1]: \alpha L + (1-\alpha)L' \succsim L''\} \subset [0, 1]$$

$$\{\alpha \in [0, 1]: L'' \succsim \alpha L + (1-\alpha)L'\} \subset [0, 1]$$

are closed sets.

This definition is **rather opaque**, but what **it means** in practice is that **small changes in probabilities do not change the ordering between two lotteries**. This property rules out lexicographical preferences over lotteries.

More importantly, as in the case with no uncertainty, **continuity implies the existence of a utility function** over all simple lotteries, which can be expressed as $U: \mathscr{L} \to \mathbb{R}$, such that $L \succsim L'$ if an only if $U(L) \geq U(L')$.

The following property of preferences over lotteries has important implications for the utility function U( ):

Definition 6.B.4: The preference relation $\succsim$ on the space of simple lotteries $\mathscr{L}$ satisfies the **independence axiom** if, for all L, L′ and $L'' \in \mathscr{L}$ and all $\alpha \in (0, 1)$ the following holds:

$$L \succsim L' \text{ if and only if } \alpha L + (1 - \alpha)L'' \succsim \alpha L' + (1\text{-}\alpha)L''$$

That is, if we combine each of two lotteries with a third one, then the preference ordering of the two combined lotteries (which are compound lotteries) does not depend on (**is independent of**) the third lottery.

The independence axiom has important implications for the theory of choice under uncertainty, as we will soon see. It is also a strong assumption. For example, an **analogous assumption in the certainty scenario** would be **weak separability**, i.e. that a consumer's preferences for a subset of goods, e.g. goods 1 and 2, does not depend on the other goods consumed by the consumer. **Yet** note that in the uncertainty scenario that **independence does not assume** that the **decision maker will be "forced" to consume L or L′ together with L″**. In the end only one of those L's will be "consumed" by the decision maker.

Now turn to an **extremely important definition**.

Definition 6.B.5: The utility function U: $\mathscr{L} \to \mathbb{R}$, has an **expected utility form** if there is an assignment of "numbers" (u_1, u_2, …u_N) to the N possible outcomes such that for **every** lottery L = (p_1, p_2, … p_N) $\in \mathscr{L}$ we have:

$$U(L) = u_1p_1 + u_2p_2 + \ldots + u_Np_N$$

A utility function with this property is call a **von Neumann – Morgenstern (v.N-M) expected utility function**.

Note: The "little u's" can be interpreted as the utility of their associated outcomes. For any n we could choose a lottery that yields outcome n with probability 1 (all other outcomes would have probability 0); for this lottery, which we can call L_n, we have $U(L_n) = u_n$, which implies that u_n is a "utility".

This type of utility function implies that the **utility** obtained **from a compound lottery is a linear function of the utilities associated with each of the simple lotteries** that are a part of the compound lottery. That is:

Proposition 6.B.1: A utility function U: $\mathscr{L} \to \mathbb{R}$ has an expected utility form if and only if it is **linear**, that is if and only if it satisfies the following property:

$$U\left(\sum_{k=1}^{K} \alpha_k L_k\right) = \sum_{k=1}^{K} \alpha_k U(L_k)$$

for any K lotteries $L_k \in \mathscr{L}$, k = 1, 2, … K with probabilities $(\alpha_1, \alpha_2, \ldots \alpha_K) \geq 0$ such that $\sum_{k=1}^{K} \alpha_k = 1$.

Proof: We can write any $L = (p_1, p_2, \ldots p_N)$ as a convex combination of the "degenerate" lotteries $(L^1, L^2, \ldots L^N)$, that is $L = \sum_{n=1}^{N} p_n L^n$. Suppose that U( ) satisfies the linear property. Then $U(L) = U\left(\sum_{n=1}^{N} p_n L^n\right) = \sum_{n=1}^{N} p_n U(L^n) = \sum_{n=1}^{N} p_n u^n$, and so it has the expected utility form.

To go in the other direction, assume that U( ) has the expected utility form. Consider any compound lottery

$(L_1, L_2, \ldots L_K; \alpha_1, \alpha_2, \ldots \alpha_K)$, where $L_k = (p_1^{\,k}, p_2^{\,k}, \ldots p_N^{\,k})$. The associated reduced lottery is $L' = \sum_{k=1}^{K} \alpha_k L_k$. So we have (recalling the the u's are numbers, not functions):

$$U(\sum_{k=1}^{K} \alpha_k L_k) = \sum_{n=1}^{N} u_n(\sum_{k=1}^{K} \alpha_k p_n^{\,k}) = \sum_{k=1}^{K} \alpha_k(\sum_{n=1}^{N} u_n p_n^{\,k}) = \sum_{k=1}^{K} \alpha_k U(L_k).$$

Q.E.D.

Note that this expected utility function is a **cardinal** property of utility functions that are defined on the "space" of lotteries. That is, **the ranking of choices remains the same only under LINEAR transformations of the expected utility function**. The following proposition states this formally:

Proposition 6.B.2: Suppose that $U: \mathscr{L} \rightarrow \mathbb{R}$ is a v.N-M expected utility function for the preference relation $\succsim$ on $\mathscr{L}$. Then another function $\tilde{U}: \mathscr{L} \rightarrow \mathbb{R}$ is a v.N-M utility function for the same $\succsim$ if and only if there are scalars $\beta > 0$ and γ such that $\tilde{U}(L) = \beta U(L) + \gamma$ for every $L \in \mathscr{L}$.

Proof: See MWG pp.173-174.

An important consequence of this proposition is that **utilities (over lotteries) are not simply ordinal**, which means that the **differences in utilities** between different lotteries **affect choices**.

For example, suppose that there are four outcomes, each with their own utility (u_1, u_2, u_3 and u_4). The statement that the difference in the first two utilities is greater than the difference between the last two, i.e. $u_1 - u_2 > u_3 - u_4$, is equivalent to:

$$u_1/2 + u_4/2 > u_2/2 + u_3/2$$

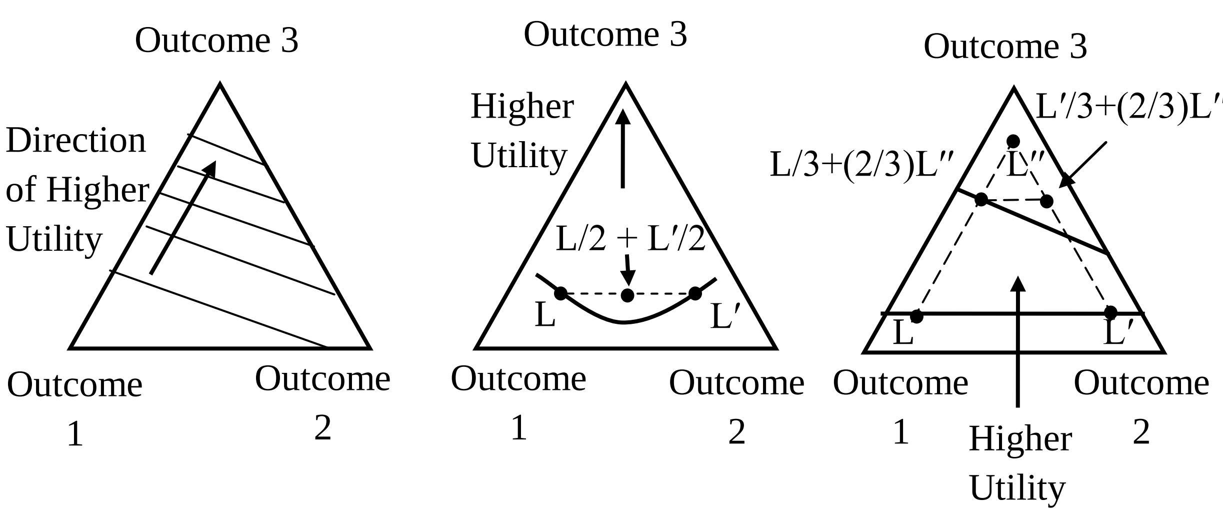
That is, the lottery L = (½, 0, 0, ½) is preferred to the lottery L′ = (0, ½, ½, 0). This ranking is preserved under all linear transformations of the v.N-M expected utility function, but is not necessarily preserved under nonlinear transformations. **Draw concave and convex utility functions (L's on horizontal axis).**

Note: **If** a preference relation $\succsim$ on $\mathscr{L}$ can be represented by **a utility function U( ) that has the expected utility form**, then since a linear utility function is continuous it follows that $\succsim$ is continuous on $\mathscr{L}$. Even more importantly, **the preference relation $\succsim$ must also satisfy the independence axiom**. The expected utility function, which we will now examine in detail, implies that the converse is true (a preference relation $\succsim$ that is continuous and satisfies the independence axiom can be represented by a utility function that has the expected utility form).

IV. The Expected Utility Theorem

The expected utility theorem is used very often in economic analysis to study behavior under uncertainty. Indeed, Mas-Colell et al. say that "the rest of the book bears witness to its usefulness".

Before presenting the theorem rigorously, it is useful to provide some visual intuition for it. Consider the case with only 3 outcomes (N = 3). For the triangle that depicts all possible sets of lotteries (probabilities) for these outcomes, we can draw indifference curves to show the decision maker's preferences for all possible sets of lotteries. This is shown in the left triangle in the following diagram:



Notice that the indifference curves are straight, parallel lines. Why can't they be curved? This is explained in the middle triangle. Indifference curves will be straight lines if, for every pair of lotteries L and L′ for which L ~ L′, it is also the case that $\alpha L + (1-\alpha)L' \sim L$ for all $\alpha \in [0, 1]$. In the middle triangle this is shown **not** to be true for the case where $\alpha = 1/2$. That is, $L'/2 + L/2 \succ L/2 + L/2$. But this can be shown to contradict the independence axiom (possible homework).

Finally, why not straight lines that are not parallel? This is shown in the triangle on the right. In that figure, we have $L \succsim L'$ (since $L \sim L'$). But in that figure the relation $L/3 + (2/3)L'' \succsim L'/3 + (2/3)L''$ does NOT hold, which also violates the independence axiom. Thus the lines must also be parallel.

With this intuition, we are now ready to state the expected utility theorem:

Proposition 6.B.3 (Expected Utility Theorem): Suppose that the **rational preference relation** $\succsim$ on the space of lotteries $\mathscr{L}$ **satisfies** the **continuity and independence** axioms. **Then** $\succsim$ **can be represented by a utility representation of the expected utility form**. That is, we can assign a number u_n to each outcome n = 1, 2, … N in such a manner that for any two lotteries $L = (p_1, p_2, \ldots p_N)$ and $L' = (p_1', p_2', \ldots p_N')$, we have:

$$L \succsim L' \quad \text{if and only if} \quad \sum_{n=1}^{N} u_n p_n \geq \sum_{n=1}^{N} u_n p_n'$$

The proof is rather long and is given on pp.176-178 of Mas-Colell et al.

V. Discussion of the Theory of Expected Utility

Expected utility theory has many implications, but it is important to think carefully about whether it is an accurate description of behavior. In particular, the independence axiom is perhaps the key assumption that one could challenge.

Example: Expected utility can show you how to choose between two lotteries that are similar (MWG, pp.178-9)

In a situation with 3 possible outcomes (N = 3), suppose you have 3 choices that are all on the same line in a 2-dimensional (triangular) simplex. Suppose that 2 are close together, but the third is farther away. If you can't make up your mind between the first two, but you can decide for the one farther away, then that tells you how to choose between the 2 close together ones (**Draw picture**).

The Allais Paradox

A famous paper by Allais (1953) presents an example which may make you question expected utility theory. It is quite simple. It has only 3 possible outcomes (N = 3):

First prize	Second prize	Third prize
\$2,500,000	\$500,000	\$0

First, the decision maker is asked to choose between the following two lotteries:

$$L_1 = (0, 1, 0) \qquad L_1' = (0.10, 0.89, 0.01)$$

Which would you choose? (Let's vote!)

Second, the decision maker is asked to choose between the following two lotteries:

$$L_2 = (0, 0.11, 0.89) \qquad L_2' = (0.10, 0, 0.90)$$

Which would you choose? (Vote again!)

Quite a few people choose $L_1 \succ L_1'$ and $L_2' \succ L_2$. Intuitively, people choose $L_1 \succ L_1'$ to "avoid uncertainty" even though the expected value of the L_1 (\$500,000) is well below the expected value of L_1' (\$695,000). The intuition for the second choice is that in an unavoidably risky world, it is OK to accept a little more risk of getting nothing in exchange for a bigger “grand prize” if you happen to “hit the jackpot”.

Yet **these two choices are not consistent with expected utility theory**. To see why, denote the utilities from these three outcomes by $u_{2.5}$, $u_{0.5}$ and u_0. The choice $L_1 \succ L_1'$ implies that:

$$u_{0.5} > 0.10 \times u_{2.5} + 0.89 \times u_{0.5} + 0.01 \times u_0$$

But if you add $0.89 \times u_0 - 0.89 \times u_{0.5}$ to both sides you get:

$$0.11 \times u_{0.5} + 0.89 \times u_0 > 0.10 \times u_{2.5} + 0.90 \times u_0$$

which means that you should choose $L_2 \succ L_2'$ in the second lottery.

One response to this is that people choose $L_1 \succ L_1'$ because if they did choose L_1' and end up with 0, they have feelings of regret because they could have had \$500,000 as a “sure thing”. This leads to **regret theory**.

See pp.180-183 of Mas Colell et al. for further discussion.

VI. Introduction to Prospect Theory (Barberis, 2013)

Psychologists also have studied decision-making under uncertainty, and in the past 30 years or so economists have been interacting with them. Perhaps the theory that has received the most attention from economists is **prospect theory**, which made its debut in economics in a 1979 paper in *Econometrica* by Kahneman and Tversky. Here I use the notation of that paper, which is somewhat different from Mas Colell et al. notation.

Consider the following "gamble":

$$(x_{-m}, p_{-m}; x_{-m+1}, p_{-m+1}; \ldots ; x_0, p_0; \ldots ; x_{n-1}, p_{n-1}; x_n, p_n)$$

where each x is an amount of money, each p is its associated probability, the x's are arranged in order from lowest to highest ($x_i < x_j$ if $i < j$), and $x_0 = 0$.

Under expected utility theory a decision maker values this gamble as:

$$\sum_{i=-m}^{n} p_i U(W + x_i)$$

where W is wealth and U( ) is increasing and concave.

In contrast, (cumulative) prospect theory evaluates it as:

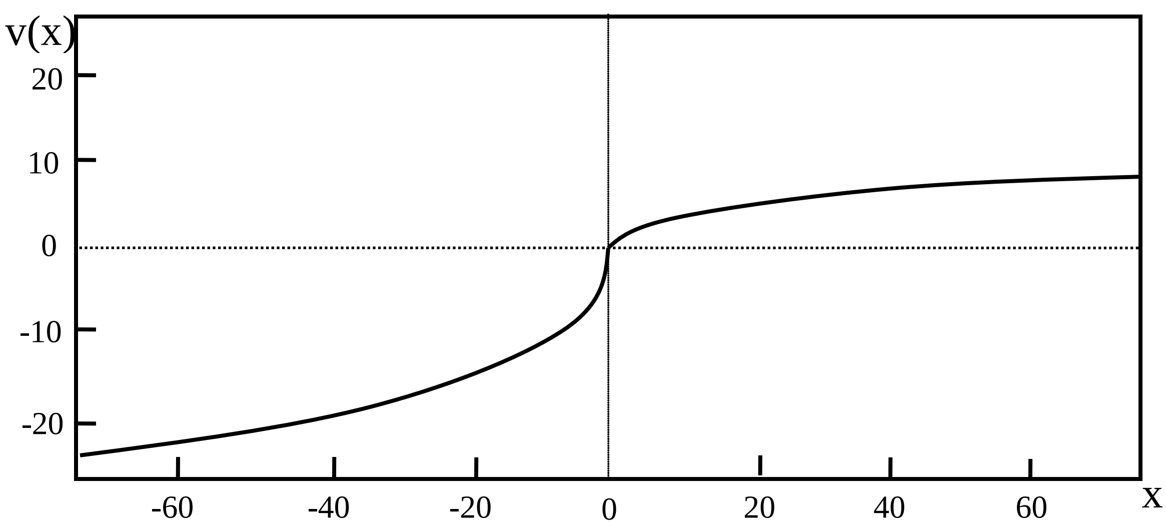
$$\sum_{i=-m}^{n} \pi_i v(x_i)$$

where the π's are "decision weights" and and v( ) is a "value function" that is increasing in x and has v(0) = 0.

This set-up illustrates the **4 elements of prospect theory**:

1. **Reference dependence.** People derive utility from gains or losses *measured with respect to some reference point* (often their current situation) instead of deriving utility from specific values of money (wealth). That is people are more sensitive to ***changes*** in their situations than they are to levels/magnitudes. Hence no W in v( ).

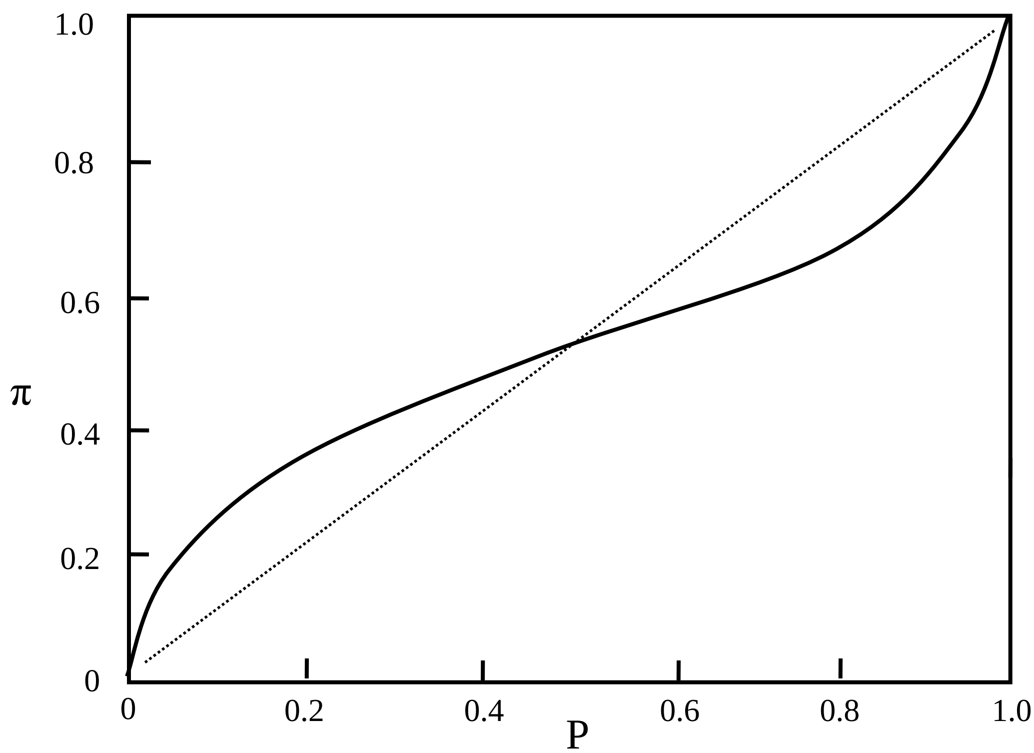
2. **Loss aversion.** People are more sensitive to losses, even very small ones, than they are to gains of the same magnitude. This can be depicted as follows:



The loss of "value" from losing $5 is much larger than the gain in value from receiving $5. For example, most people will decline a gamble that pays $5.50 with 50% probability but costs -$5 with 50% probability.

3. **Diminishing sensitivity to gains *and* losses.** With respect to the "status quo" (no change in income/wealth), the value function is convex for negative x but concave for positive x. This is also seen in the diagram on p.18. For example, the increase in value going from a $10 gain to a $20 gain is larger than the increase in value going from a $100 gain to a $110 gain. That is, people are risk averse in gains; they prefer $500 for sure relative to a 50% change of gaining $1000. **However**, they are ***risk seeking*** in losses: they prefer a 50% chance of losing $1000 than a certain loss of $500.

4. **Probability weighting.** People do not weigh events by their actual probabilities but apply probability weights that may not equal actual probabilities. The probability weights are **systematic** in that they **give greater weight to rare events**. This is seen in the next figure. An example of this is that even risk averse people will prefer a 0.1% chance of obtaining $5000 to a certain $5. Perhaps this explains why societies spend billions of $ to reduce deaths from rare events (plane crash, terrorist attack) but ignore common events (car accidents).



There are, of course, some **complications with applying prospect theory**:

1. What are people's reference points, and how stable are they over time?

2. Do "laboratory experiments" that support prospect theory tell us how people behave in the "real world?

3. Do we need all 4 of the above elements, or just some of them?

4. In what "domains" of economics does it apply?