Utility and Decision Making [1]

Suppose that a casino offers to keep flipping a coin until it comes up heads for the first time. If it comes up heads on the 1st flip, you win \$1. If it comes up heads for the first on the 2nd flip, you win \$2. On the 3rd flip, \$4. Indeed, if the coin initially comes up tails k times in a row before coming up heads, you will win \$$2^k$.

If you play the game, you're guaranteed to win at least a little bit of money. And if you're lucky, you might win a bundle! For example, if the coin comes up tails 10 times in a row before finally coming up heads—unlikely, but not impossible—you'll walk away with $2^{10} = \$1024$. The question is: how much money should you rationally be willing to pay to enter the game?

If this is the first time you've encountered this hypothetical, the naïve answer may surprise you. Let's follow the reasoning of Swiss mathematician Gabriel Cramer, in his letter of 28 May 1728 written to Nicolas Bernoulli, the man who originally posed the question:

> I know not if I deceive myself, but I believe to hold the solution of the singular case that you have proposed to Mr. de Montmort in your letter of 9 September. . . . In order to render the case more simple I will suppose that A throw in the air a piece of money, B undertakes to give him a coin, if the side of the Heads falls on the first toss, 2, if it is only the second, 4, if it is the 3rd toss, 8, if it is the 4th toss, etc. The paradox consists in this: that the calculation gives for the equivalent that A must give to B an infinite sum, which would seem absurd, since no person of good sense, would wish to give 20 coins. [2]

[2] *Correspondence of Nicolas Bernoulli concerning the St. Petersburg Game*, available at `www.cs.xu.edu/math/Sources/Montmort/stpetersburg.pdf`.

An infinite sum?! To understand Cramer's point, let's compute the expected value of the game. If X is the random variable representing your potential winnings, then

$$
\begin{aligned}
E(X) &= \sum_{k=0}^{\infty} x_k \cdot p(x_k) \\
&= \sum_{k=0}^{\infty} 2^k \cdot (1/2)^{k+1} \\
&= 1(1/2) + 2(1/4) + 4(1/8) + \ldots \\
&= 1/2 + 1/2 + 1/2 + \ldots \\
&= \infty
\end{aligned}
$$

The expected value of the game is indeed infinite. If you are trying to maximize your expected value, you should play the game no matter how much the casino charges. Very large wins may happen with very small probability, but their potentially enormous amount compensates you for the risk.

Yet as a practical matter, very few people express a desire to pay even $10 or $20 enter into the game, even though it would seem to be to their advantage to do so. Is this behavior irrational? Not at all! Here's Cramer again, in the same letter to Bernoulli:

> One asks the reason for the difference between the mathematical calculation and the vulgar estimate. I believe that it comes from this that the mathematicians estimate money in proportion to its quantity, and men of good sense in proportion to the usage that they may make of it. This which renders the mathematical expectation infinite, this is the prodigious sum that I am able to receive, if Heads falls only very late, the 100th or 1000th toss. Now this sum, if I reason as a sensible man, is not more for me, does not make for more pleasure for me, does not engage me more to accept the game, than if it would be only 10 or 20 million coins. [3]

[3] *ibid.*

In other words, even if the potential rewards (2, 4, 8, 16, 32 ...) don't reach any ceiling in terms of face value, they do in terms of practical value. This simple fact makes the game much less attractive.

Utility functions

Cramer's answer is, essentially, the modern-day economist's answer: *people don't keep score with money.* Rather, they keep score according to a more private, idiosyncratic scale of value. This scale need not match up in any linear way with actual money or material goods, and is entirely subjective. It might also be characterized by something like a law of diminishing returns. As the famous statistician and philosopher Harold Jeffreys put it,

> ... the pleasures to me of two dinners on consecutive nights seem to be nearly independent, while those of two dinners on the same night are definitely not.[4]

[4] *The Theory of Probability* (1961), Chapter I, page 32.

Or, to invoke a more crass example: having $20 million might give you twice as big a bank account as having $10 million, but would it mean twice as much happiness or pleasure? For many people, the answer is no. (And if you answered yes: would $20 billion give you twice as much pleasure as $10 billion?)

This private scale of relative satisfaction is called *utility*. The hypothetical situation discussed by Bernoulli and Cramer came to be known as the *St. Petersburg paradox*,[5] and it is the genesis of modern utility theory. The game is considered a paradox in the sense that naïvely applying the idea of expected value leads to a silly recommendation: that you should be willing to pay any finite sum of money to play, a willingness that no one exhibits in real life.

[5] So named after the 1738 presentation of the problem by Daniel Bernoulli (Nicolas's cousin) in a collection of scientific and mathematical essays called *Commentaries of the Imperial Academy of Science of Saint Petersburg*.

The paradox can be made to disappear by invoking the notion of *expected utility*. This is a mathematical concept, not a psychological one, and therefore it takes some sort of justification if we are going to equate the two. We now sketch the outlines of one such justification.

Subjective utilities

Utilities, like probabilities, are yours and yours alone. They are assignments to value to different possible outcomes, and can be defined in terms of your preferences between gambles.

For example, here's a simple decision tree, depicting a choice between a sure thing ($5 million) and a gamble with an equivalent expected value. If you have a preference either way, then your utility for $10 million is not equal to twice your utility for $5 million.

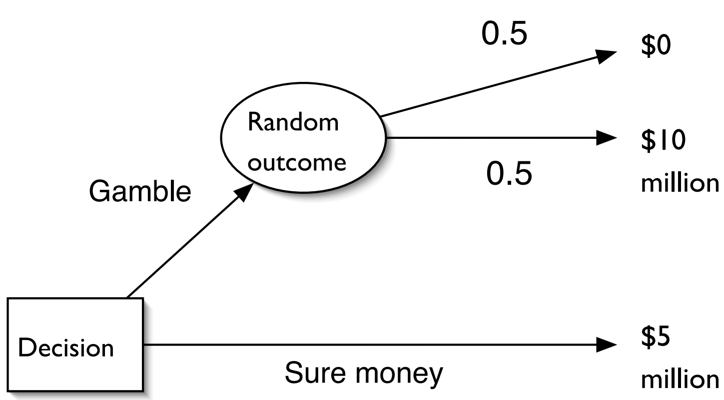


Figure 1: In a decision tree, all random outcomes are depicted with an oval, with the possible outcomes and their probabilities attached. These are called *stochastic nodes.* All points where you must make a decision among different courses of action are marked with rectangles. These are called *decision nodes*.

If you prefer the sure thing, you are are risk-averse. If you prefer the gamble, you are risk-seeking. If you are indifferent, you are risk-neutral.

To formalize this notion, we must define the notion of a *simple* *lottery*—that is, a probability distribution with just two possible outcomes, A (which happens with probability w) and B (probability $1 - w$). In the decision tree above, the branch leading to a coin flip between \$10 million and \$0 is a simple lottery. So is the "sure \$5 million" branch of the tree, which is an especially simple lottery since one of the two events occurs with probability 1.

A compound lottery is a lottery with more than one possible outcome: $P = w_1 A_1 + \cdots w_n A_n$, where the probabilities $w_1, \ldots, w_n$ sum to 1. Any compound lottery can be reduced recursively to a tree of simple lotteries. For example, the compound lottery $P = w_1 A_1 + w_2 A_2 + w_3 A_3$ is equivalent to the simple lottery

$$P = w_1 A_1 + (1 - w_1) Q \,,$$

where Q is itself a simple lottery between A_2 and A_3. If you are indifferent between a sure thing P and a lottery Q, then P is your *certainty equivalent* for the uncertain gamble Q. The difference between the expected value of Q and its certainty equivalent P is called a *risk premium*.

The von Neumann–Morgenstern theorem of expected utility

Your utilities are defined in terms of your preferences between different lotteries. We'll use the notation "$P \succ Q$" to express mathematically the idea that you'd rather play lottery P than Q. The binary relation "$\succ$" is called your *preference relation*.

We will start by assuming that your preferences among lotteries obey four simple rules of rationality, which are called the *von Neumann–Morgenstern rules*. What the principle of coherence is to probability theory, these rules are to utility theory: some very simple, common-sense constraints on your preferences from which many beautiful results can be derived. The rules are:

(1) *Your preferences are complete.* That is, for any two lotteries P and Q, you are able to state whether $P \succ Q$, $Q \succ P$, or $P \sim Q$ (the third option indicating your indifference between the two.) You're not allowed to say, "I don't know."

(2) *Your preferences are transitive.* That is, if $P \succ Q$ and $Q \succ R$, then $P \succ R$.

(3) *You are willing to "split the difference" between favorable and unfavorable options.* This one is often called the rule of continuity. Suppose there

are three lotteries P, and Q, and R, such that P is your favorite choice, R is your least favorite, and Q is somewhere in the middle ($P \succ Q \succ R$). If your preferences satisfy the rule of continuity, then there exists some probability w such that $Q \sim wP + (1-w)R$. In other words, there must be some probability w where you indifferent between a w-weighted coin flip for P and R, and the "split the difference" option Q.

(4) *Your preference for P or Q remains unchanged in the face of a third option.* Formally, if $P \succ Q$, w is a probability, and R is any third lottery, then

$$wP + (1-w)R \succ wQ + (1-w)R\,.$$

That is, you prefer P to Q, then you also prefer a lottery involving P to the same lottery involving Q.

The significance of these rules is that they make possible the construction of something called a *utility function,* a fact realized in 1944 by John von Neumann and Oskar Morgenstern. In their book entitled *Theory of Games and Economic Behavior*, von Neumann and Morgenstern proved the following lovely result, often called the *expected-utility theorem*:

Theorem 1. *Suppose that an agent's preferences among outcomes satisfy Rules 1–4. Then there exists a utility function $u(A_i)$ that assigns a real number to each possible outcome A_i, and that satisfies two properties:*

1. *For any two lotteries P and Q and any probability w,*

$$u\{wP + (1-w)Q\} = wu(P) + (1-w)u(Q)$$

2. *For any two lotteries P and Q, $E\{u(P)\} > E\{u(Q)\}$ if and only if $P \succ Q$.*

In the theorem, $E\{u(P)\}$ denotes the agent's expected utility in lottery P, recalling that P is a probability distribution over outcomes: $P = w_1 A_1 + \cdots + w_n A_n$. Just as the expected value is a weighted sum of possible values, the expected utility is the weighted sum of possible utilities:

$$E\{u(P)\} = w_1 u(A_1) + \cdots w_n u(A_n)\,.$$

A proof of this theorem can be found in just about any textbook on utility theory, including von Neumann and Morgenstern's original work on the topic.

The beauty of the von Neumann–Morgenstern theorem is its demonstration that it is both philosophically meaningful and mathematically precise to codify preferences using real numbers called utilities (subject to the assumption that preferences obey certain rules). These utilities are on an arbitrary scale, but can be ordered just like ordinary numbers. It is therefore sensible to compare a single individual's utilities across different outcomes, though not to compare utilities across different people.

You'll also notice that the theorem makes no assumption about the kind of events involved in the lotteries. They could be monetary or hedonistic, but they need not be. Indeed, they could be personally significant outcomes—like getting married, having the respect of one's peers, or going for a swim in the afternoon—whose worth defies easy quantification in terms of money. Regardless of the events themselves, their utilities can all be placed on a common scale, as long as your preferences among the events behave the basic rules of rationality (1–4).

Invoking the notion of utility resolves the St. Petersburg paradox: a person's expected utility for the bet can be small, even if the expected monetary value of the bet is infinite. Expressed mathematically, if $u(x)$ denotes a person's utility for having x dollars, having a small expected utility for the St. Petersburg game would mean that

$$E\{u(X)\} = \sum_{k=0}^{\infty} u(x_k) \cdot p(x_k) = c\,,$$

for some small number c, even though

$$E\{X\} = \sum_{k=0}^{\infty} x_k \cdot p(x_k) = \infty\,.$$

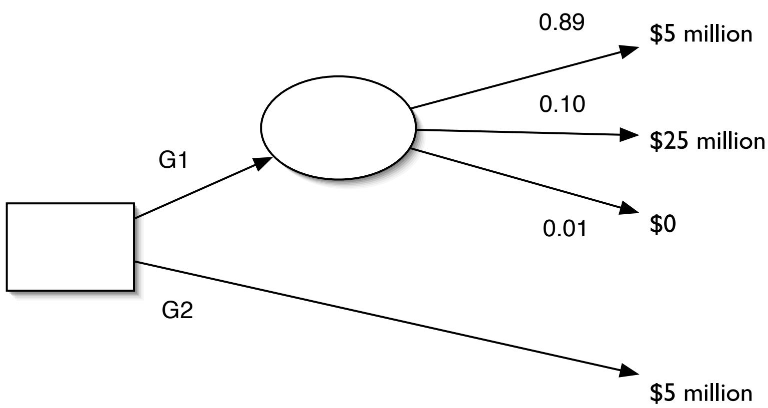
This would happen whenever the utility of a reward increases more slowly than the numerical size of a reward itself.

This leads us to the notion of *risk aversion* and a categorization of people according to their risk tolerance. Your utility function $u(x)$ marks you as **risk averse** if $\mathrm{E}\{u(X)\} \leq u\{\mathrm{E}(X)\}$; **risk neutral** if $\mathrm{E}\{u(X)\} = u\{\mathrm{E}(X)\}$; or **risk seeking** if $\mathrm{E}\{u(X)\} \geq u\{\mathrm{E}(X)\}$. (These classifications are subject to the assumption that these inequalities hold for all probabilities and random variables.) In layman's terms: a risk-averse person prefers a certain outcome over a lottery with an equivalent expected value, while a risk-seeking person prefers the opposite.

The Allais paradox

The von Neumann–Morgenstern rules strike most people as eminently sensible requirements of a rational actor. Nonetheless, they have some teeth! Indeed, it is quite common for people's declared preferences to be inconsistent with one or more of the axioms.

One such example is the *Allais paradox*. Take this lottery:



The expected value of G_1 is

$$E(G_1) = 0.89 \cdot 5 + 0.1 \cdot 25 + 0.01 \cdot 0 \approx 6.95$$

in units of $1 million, which is clearly higher than the expected value of a guaranteed $5 million. Yet many people prefer G_2, reasoning that $5 million is a life-changing amount of money, and that the risk of getting nothing—even if only 1%—simply isn't worth the 10% shot at $25 million. This is perfectly sensible from an expected-utility point of view. Invoking the von Neumann–Morgenstern rules, for a rational agent to choose G_2 over G_1 simply means that her utilies are such that $E\{u(G_2)\} > E\{u(G_1)\}$. In turn, this means that, in multiples of $1 million,

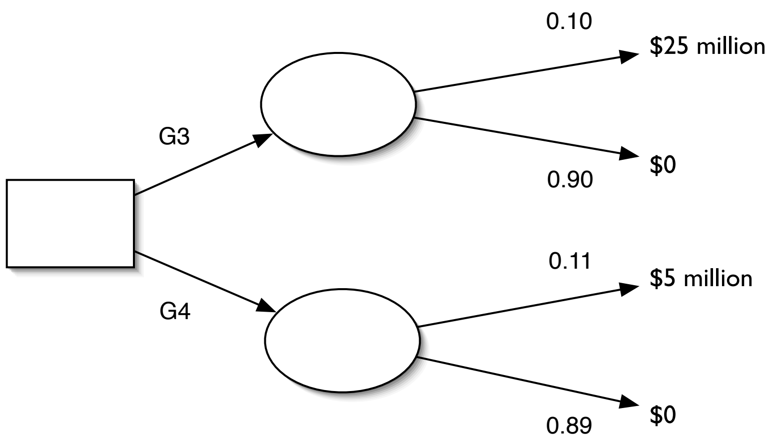
$$u(5) > 0.89u(5) + 0.1u(25) + 0.01u(0)\,.$$

If we assume that $u(0) = 0$, then this is equivalent to having utilities $u(5)$ and $u(25)$ such that

$$u(5) > \left(\frac{0.10}{0.11}\right) u(25) \approx 0.909u(25)\,. \qquad (1)$$

That is to say: as long as your utility for $5 million is greater than 90.9% of your utility for $25 million, then G_2 will maximize your expected utility, and the bottom branch of the tree is the rational choice.

So far so good. But what about this lottery?



Many people state a strong preference for G_3 here, reasoning that the probability of winning G_3 is only slightly smaller, but the reward much larger, than for G_4. This is also perfectly sensible from an expected-utility point of view, as long as

$$0.11u(5) < 0.10u(25)\,,$$

or equivalently, that

$$u(5) < \left(\frac{0.10}{0.11}\right) u(25) \approx 0.909u(25)\,. \qquad (2)$$

This set of preferences is also fine, on its own. But it flatly contradicts the constraint on $u(25)$ and $u(5)$ implied by the choice $G_2 \succ G_1$, in Equation (1).

It is therefore impossible to hold a consistent set of utilities for $5 million and $25 million such that $G_2 \succ G_1$ and $G_3 \succ G_4$ are both true. Even though your utilities are subjective, the notion of expected utility requires that either G_2 or G_3 must be the "wrong" choice, no matter what your utilities are! Someone who takes both of these gambles must be violating one of the four rules.

The Allais paradox is just one of many examples where many people's declared preferences put them at odds with the consistency requirements of formal utility. Simply put: in practice, most people don't behave like utility maximizers at all.

For a wonderful discussion of this, try reading up on the work of Daniel Kahneman and Amos Tversky, who won a Nobel Prize in economics for their work characterizing the cognitive underpinnings—the biases, the fears, the hasty generalizations, the faulty heuristics—for many human "errors" of precisely this type.