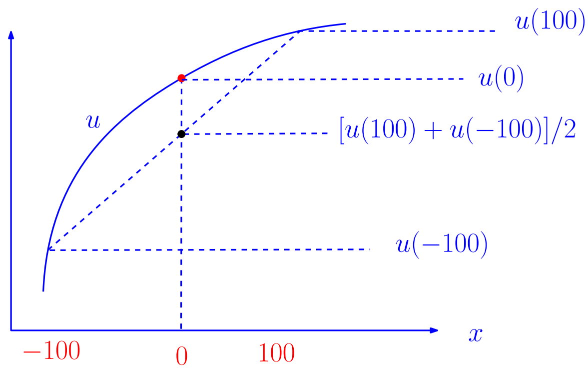
Notes on Decisions Under Risk [1]

Motivation and intuition

A coin which falls Heads or Tails with equal probability (a half each) is called fair. If you ever wondered, it is called fair because if you bet x Euros on Heads against me the coin does not favor either of us, both have zero expected gain: $0.5 * x + 0.5 * (-x) = 0$. That is the most the coin can do, but zero isn't necessarily the worth we attach to the bet. In other words expected value as a valuation criterion doesn't always reflect preferences. The starkest example to see the point is the St. Petersburg "paradox": suppose you are offered the prospect of winning 2^k Euros with probability 2^{-k}, for $k \geq 1$;[2] The expected value of such a bet is $1 + 1 + \cdots = \infty$, but I am pretty sure you prefer 10 million Euros for sure to such a bet. If so you are not evaluating bets according to their expected values.

Indeed, going back to the simpler coin case: if I invite you to bet a hundred Euros on a fair coin, would you accept? Most likely not - you would prefer not to bet, that is getting zero for sure. Said otherwise, your valuation of that zero-expectation bet is *less* than zero. The easiest explanation of this "risk averse" behavior is that money gives utility but the utility of money - say u - does not grow as much as the money itself - in other words it has a decreasing derivative, that is it is *concave*. In this case when you take expected *utility* rather than expected value, that is you evaluate the bet by computing $0.5 * u(x) + 0.5 * u(-x)$, you get a number smaller than $u(0)$. Picture:



Making precise the connection between risk aversion (properly defined) and concave utility of money and its basic implications is one of the principal goals of these notes.[3]

---

[1] Salvatore Modica 2023. Based on Wakker *Prospect Theory*, CUP

[2] Those probabilities do sum to 1 (geometric series).

[3] For the sake of curiosity: if in the St. Petersburg bet you take a logarithmic utility of money, $u(Euro) = \ln(Euro)$, and compute expected utility you get $\sum_{k \geq 1} 2^{-k} \ln 2^k = \ln 4 \approx 1.4$ (pretty far from infinity). Indeed $\ln 2$ factors out, and the remaining sum is 2 because

$$\sum_{k \geq 1} k \cdot 2^{-k} = \sum_{k \geq 1} 2^{-k} + \sum_{k \geq 1} (k-1) \cdot 2^{-k} = 1 + \frac{1}{2} \sum_{k \geq 1} (k-1) \cdot 2^{-(k-1)} = 1 + \frac{1}{2} \sum_{k \geq 1} k \cdot 2^{-k}.$$

1 Setup and basic definitions

The model A measurable space $(\Omega, \mathcal{F}, \mathsf{P})$ is given. Objects of choice are simple random variables (henceforth r.v.)

$$\xi = \sum\nolimits_{i=1}^{n} x_i \mathbf{1}(A_i)$$

where the x_i are distinct reals, $\mathbf{1}(A_i)$ is the indicator of set $A_i \subseteq \Omega$ and $\{A_1, \ldots, A_n\}$ is a partition of Ω; here $A_i = \{\omega \colon \xi(\omega) = x_i\}$. For any $x \in \mathbb{R}$ the constant r.v. $x\mathbf{1}(\Omega)$ will be denoted by x.

A preference $\succcurlyeq$ on r.v.'s characterizes the decision maker, where $\xi \succcurlyeq \eta$ reads "ξ is preferred to η". It is assumed that $\succcurlyeq$ is a *weak order*, that is

1. Complete: for any pair ξ, η either $\xi \succcurlyeq \eta$ or $\eta \succcurlyeq \xi$ or both, and
2. Transitive: $\xi \succcurlyeq \eta \succcurlyeq \zeta \Longrightarrow \xi \succcurlyeq \zeta$.

We write $\xi \sim \eta$ if $\xi \succcurlyeq \eta$ and $\eta \succcurlyeq \xi$, and $\xi \succ \eta$ if $\xi \succcurlyeq \eta$ and not $\eta \succcurlyeq \xi$.

Recall that a simple r.v. ξ induces a probability distribution P_ξ on $\mathbb{R}$ defined by

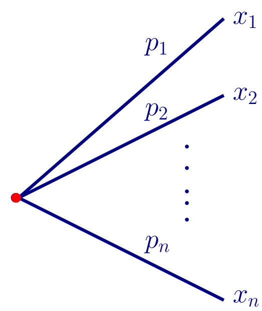
$$\mathsf{P}_\xi(x_i) = \mathsf{P}(A_i) = \mathsf{P}\{\omega \colon \xi(\omega) = x_i\}.$$

Reduction The fundamental behavioral assumption we make is that if $\mathsf{P}_\xi = \mathsf{P}_\eta$ then $\xi \sim \eta$. This implies that the decision maker is only concerned with the distributions on $\mathbb{R}$ induced by the various r.v.'s (inelegantly speaking, money) rather than with the r.v.'s themselves. The space $(\Omega, \mathcal{F}, \mathsf{P})$ plays no role in decision making. It is therefore natural to define preferences directly on those distributions, letting $\mathsf{P}_\xi \succcurlyeq \mathsf{P}_\eta$ if $\xi \succcurlyeq \eta$ (we use the same $\succcurlyeq$ symbol to save notation). This new preference clearly inherits the transitivity property of the original relation.

To speak of completeness we need to make another assumption. Observe that the distribution P_ξ induced by $\xi = \sum_i x_i \mathbf{1}(A_i)$ is characterized by its finite *support* $X = \{x_1, \ldots, x_n\}$ and by the probability vector $p = (p_1, \ldots, p_n)$ where $p_i = \mathsf{P}_\xi(x_i)$. In other words it may be represented as a pair $P = (X, p)$. The (technical) assumption we need is that the space $(\Omega, \mathcal{F}, \mathsf{P})$ is rich enough that *any* such P is induced by some r.v. on Ω.[4] At this point we have a relation $\succcurlyeq$ defined on all the finite-support distributions on $\mathbb{R}$, and we assume it is a weak order.

It is convenient to visualize $P = (\{x_1, \ldots, x_n\}, (p_1, \ldots, p_n))$ as in the following picture:

[4]Note that such an Ω cannot be finite. Fortunately the space $([0,1), \mathcal{B}[0,1), \mathrm{Leb})$ is sufficient for the purpose. If you don't know what that is just ignore the remark.



The degenerate distribution $P = (\{x\}, 1)$ giving x for sure will be also written as x. We may also occasionally use the handy notation $x_p y$ for the two-outcome distribution assigning probability p to x and $1-p$ to y.

Note on terminology: we shall use interchangeably the terms distribution, lottery and prospect.

Expectation Recall that for r.v. $\xi = \sum_{i=1}^{n} x_i 1(A_i)$ and $\varphi \colon \mathbb{R} \to \mathbb{R}$, letting $p_i = \mathsf{P}_\xi(x_i)$ we have

$$E\varphi(\xi) = \sum\nolimits_{i=1}^{n} \varphi(x_i) \, \mathsf{P}_\xi(x_i) = \sum\nolimits_{i=1}^{n} \varphi(x_i) \, p_i.$$

For $P = (\{x_1, \ldots, x_n\}, (p_1, \ldots, p_n))$ we now define

$$E\varphi(P) = \sum\nolimits_{i=1}^{n} \varphi(x_i) \, p_i.$$

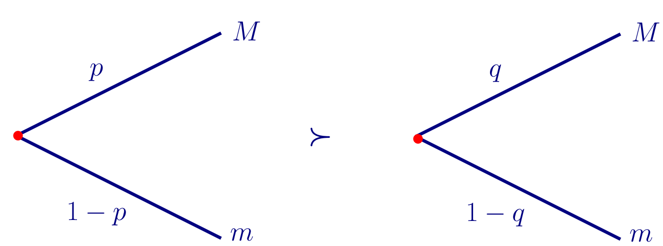
Note that if P is induced by ξ (so that $P = \mathsf{P}_\xi$) we are just defining $E\varphi(P)$ as $E\varphi(\xi)$. In the case of identity function $\varphi(\xi) = \xi$ this reduces to $E(P) = \sum_i x_i p_i$. In the sequel we most often write EP for $E(P)$ to ease reading.

2 The expected utility theorem

Expected Utility We say that EU holds for $\succcurlyeq$ if there exists a strictly increasing $u \colon \mathbb{R} \to \mathbb{R}$ such that for any pair of distributions P, Q one has $P \succcurlyeq Q \iff Eu(P) \geq Eu(Q)$. The function u is called a vonNeumann-Morgenstern (vNM) utility.

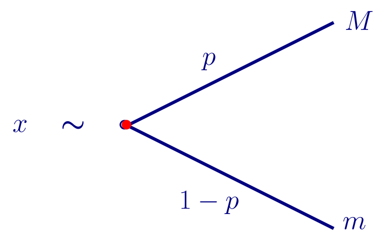
We next state three axioms on $\succcurlyeq$ which are necessary and sufficient for EU to hold.

A. Monotonicity For $M > m$ and $p > q$ the following strict preference holds:



This is pretty self-explanatory: you prefer winning more money with higher probability. Observe that for $p = 1$ and $q = 0$ the axiom says that $M \succ m$ (here the comparison is between degenerate distributions); thus we have the direct implication $M > m \Longrightarrow M \succ m$ - more money is better.

B. Continuity For any numbers $m < x < M$ there exists $p \in (0, 1)$ such that the following indifference holds:



The monotonicity axiom implies that such a p is unique (exercise).

C. Consistency To state the last axiom we need to define the mixture of two distributions. Start with two distributions $P = (\{x_1, \ldots, x_n\}, (p_1, \ldots, p_n))$ and $Q = (\{y_1, \ldots, y_m\}, (q_1, \ldots, q_m))$. For $\alpha \in \mathbb{R}$ let $p(\alpha)$ be the probability of α under P and $q(\alpha)$ the probability of α under Q, that is

$$p(\alpha) = \begin{cases} p_i & \textit{if } \alpha = x_i \textit{ some } i = 1, \ldots, n \\ 0 & \textit{otherwise} \end{cases} \qquad q(\alpha) = \begin{cases} q_j & \textit{if } \alpha = y_j \textit{ some } j = 1, \ldots, m \\ 0 & \textit{otherwise}. \end{cases}$$

Given a number $0 \leq \lambda \leq 1$, the *mixture* $P_\lambda Q$ is defined as follows: its support is $X \cup Y$, and the probability assigned to $\alpha \in X \cup Y$ is $\lambda p(\alpha) + (1 - \lambda) q(\alpha)$. These probabilities sum to 1 (proof in footnote) so the mixture is well defined.[5]

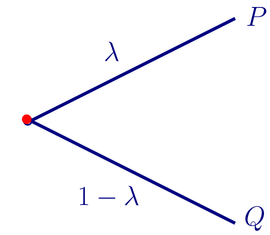
$P_\lambda Q$ obtains if you get P with probability λ and Q with probability $1 - \lambda$. The picture is this:

[5] For $\alpha \in Y \setminus X$ we have $p(\alpha) = 0$ and for $\alpha \in X \setminus Y$ we have $q(\alpha) = 0$. Therefore, since $X \cup Y = X \cup (Y \setminus X) = Y \cup (X \setminus Y)$, we have

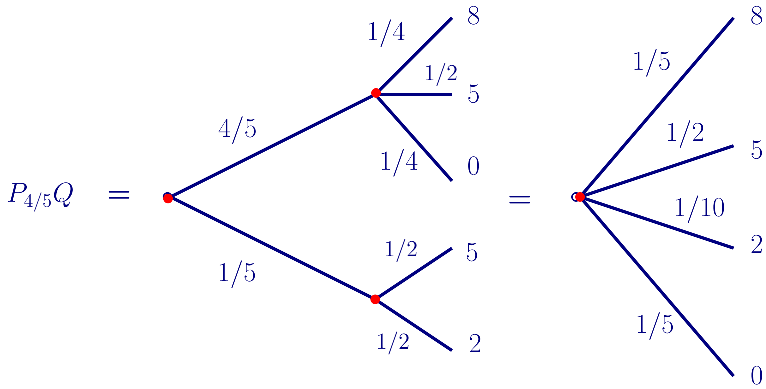
$$\sum\nolimits_{\alpha \in X \cup Y} p(\alpha) = \sum\nolimits_{\alpha \in X} p(\alpha) = 1 \quad \textit{and} \quad \sum\nolimits_{\alpha \in X \cup Y} q(\alpha) = \sum\nolimits_{\alpha \in Y} q(\alpha) = 1$$

so that

$$\sum\nolimits_{\alpha \in X \cup Y} \lambda p(\alpha) + (1 - \lambda) q(\alpha) = \lambda \sum\nolimits_{\alpha \in X} p(\alpha) + (1 - \lambda) \sum\nolimits_{\alpha \in Y} q(\alpha) = 1.$$

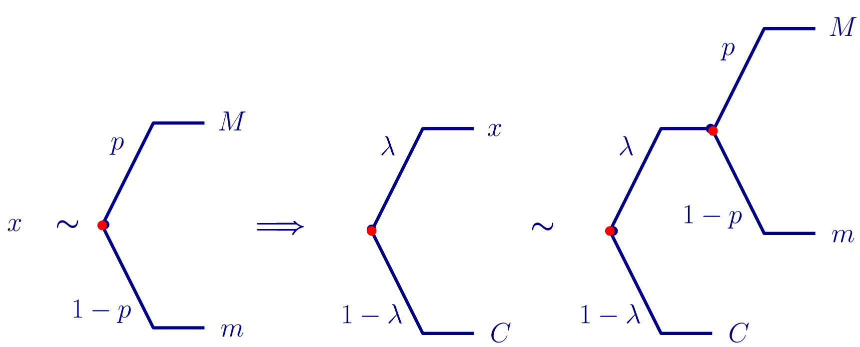


Example: let $P = (\{8, 5, 0\}, (1/4, 1/2, 1/4))$, $Q = (\{5, 2\}, (1/2, 1/2))$ and $\lambda = 4/5$. Then



Exercise Verify the following linearity property of the expectation operator, precisely that $E\varphi(P_\lambda Q) = \lambda E\varphi(P) + (1-\lambda)E\varphi(Q)$. Solution in footnote.[6]

The consistency axiom states that for all numbers x, m, M, all probabilities p and λ and distributions C the following implication holds:



[6]This is exercise 2.6.6 in Wakker. We use (as in footnote 5) the fact that for $\alpha \in Y \setminus X$ we have $p(\alpha) = 0$ and for $\alpha \in X \setminus Y$ we have $q(\alpha) = 0$. From this we get

$$\begin{aligned} E\varphi(P_\lambda Q) &= \sum\nolimits_{\alpha \in X \cup Y} [\lambda p(\alpha) + (1-\lambda) q(\alpha)]\, \varphi(\alpha) \\ &= \lambda \sum\nolimits_{\alpha \in X} p(\alpha)\varphi(\alpha) + (1-\lambda) \sum\nolimits_{\alpha \in Y} q(\alpha)\varphi(\alpha) \\ &= \lambda E\varphi(P) + (1-\lambda) E\varphi(Q). \end{aligned}$$

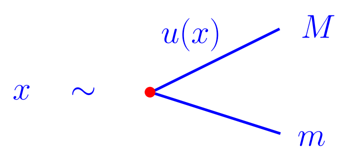
The expected utility theorem EU holds for $\succcurlyeq$ if and only if $\succcurlyeq$ is a weak order satisfying monotonicity, continuity and consistency.

Uniqueness of vNM utility The following result holds. Under EU, the vNM utility is unique up to linear transformations. That is, if u is a vNM utility for $\succcurlyeq$ then any other vNM utility for $\succcurlyeq$ is of the form $\sigma u + \tau$ with $\sigma > 0$ and $\tau \in \mathbb{R}$.

Important exercise Prove the easy directions of the above assertions. First, show that if EU holds then the axioms are satisfied (the hard part is to show the reverse implication). Second, show that if u is a vNM utility for $\succcurlyeq$ then so also is $\sigma u + \tau$ with $\sigma > 0$ and $\tau \in \mathbb{R}$ (the hard part here is to show that there are no others). Solution of the easy direction in footnote.[7]

Existence of a certainty equivalent A certainty equivalent for a distribution P is a number $CE(P)$ such that $CE(P) \sim P$. By monotonicity if such a number exists it is unique. Under EU it does exist for all P if and only if the vNM utility for $\succcurlyeq$ is continuous. This is Exercise 2.6.5 in Wakker.

The key idea for the proof of the EU theorem Notation: if P assigns probability p to x and $1-p$ to y we write $P = x_p y$. Fix two values $m < M$; *suppose EU holds*; take a vNM utility u such that $u(m) = 0$ and $u(M) = 1$; then if for an $x \in (m, M)$ the number p solves $x \sim M_p m$ (existence of such p guaranteed by the continuity axiom) we have $u(x) = p * u(M) + (1-p) * u(m) = p$. So in the search for a vNM utility between m and M we are *forced to define* $u(x)$ by the equivalence



[7]Let u be a vNM utility for $\succcurlyeq$, let $u^* = \tau + \sigma u$ and take any $P = (\{x_1, \ldots, x_n\}, (p_1, \ldots, p_n))$. Then

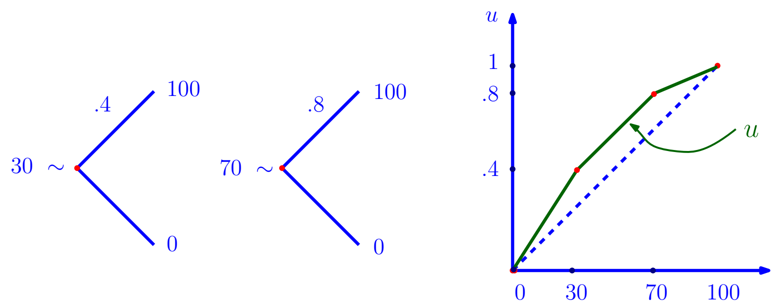
$$\begin{aligned} Eu^*(P) = \sum\nolimits_{i=1}^{n} p_i u^*(x_i) &= \sum\nolimits_{i=1}^{n} p_i \left[\tau + \sigma u(x_i)\right] \\ &= \sum\nolimits_{i=1}^{n} p_i \tau + \sum\nolimits_{i=1}^{n} p_i \sigma u(x_i) \\ &\tau + \sigma Eu(x) \end{aligned}$$

From this it follows directly that for any P, Q we have

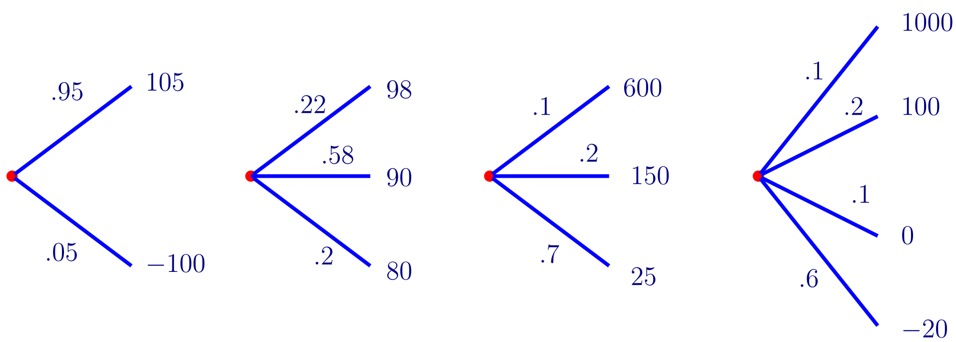
$$P \succcurlyeq Q \iff Eu(P) \geq Eu(Q) \iff Eu^*(P) \geq Eu^*(Q).$$

because we know that if there is a vNM it must be it. So the proof starts from this definition, and repeatedly using consistency and then monotonicity shows that EU holds for distributions with values in an interval $[m, M]$. The procedure is illustrated in Appendix 1 where we give the details of the proof and show how the extension from $[m, M]$ to all of $\mathbb{R}$ and uniqueness are obtained.

Building some intuition Take $m = 0,\ M = 100$ and $u(m) = 0,\ u(M) = 1$. Applying the definition of u as above, to find for example $u(30)$ and $u(70)$ we must find probabilities in the indifferences below (in the picture it is assumed they are .4 and .8) and then utility is as drawn in the right diagram (with some interpolation to get an idea).



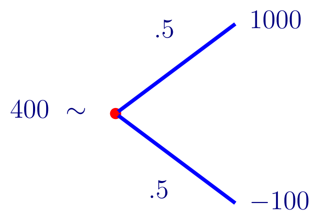
Assessing vNM utility the other way around[8] Suppose you have to choose one of these four lotteries:



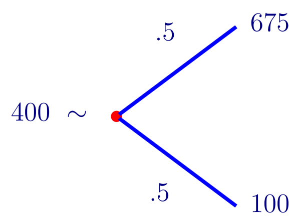
It does not seem an easy task at all, but assuming you subscribe to the expected utility axioms the problem "reduces" to assessing your vNM utility for the outcomes, for then you just compute expected utility and choose the one for which that is highest. To derive u we start by normalizing at the extremes: $u(1000) = 1, u(-100) = 0$. We want to get some utility values and then interpolate to get an acceptable estimate. We assume existence of certainty equivalents. We have just seen one way to do it: derive probabilities to get utility of given values. An alternative is to start with given probabilities - say 50-50 two-valued prospects, which are the simplest to understand - and derive utility, as follows. Start with $1000_{1/2}(-100)$

[8] From David Kreps *Notes on the theory of choice.*

and give the value which gives indifference in the comparison below; here we suppose the value is 400.



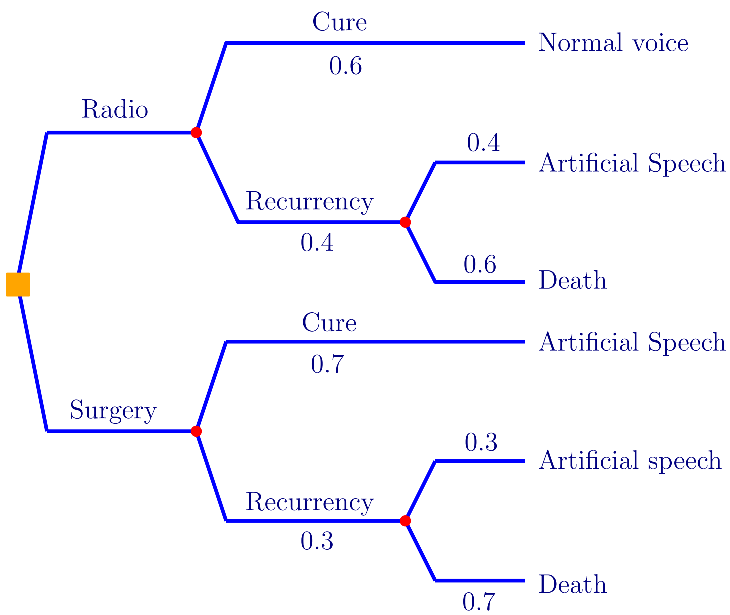
Then you know $u(400) = 1/2 \cdot u(1000) + 1/2 \cdot u(-100) = 1/2$. Given this, to get x such that $u(x) = 0.75$ we can then use the lottery $1000_{0.5}400$. Supposing you assert that $1000_{0.5}400 \sim 675$ then $u(675) = 0.5 \cdot (1 + 1/2) = 0.75$. For 0.25 we can use $400_{0.5}(-100)$; assuming you say $400_{0.5}(-100) \sim 100$ then $u(100) = 0.25$. You can go further, or just interpolate somehow between these four values to draw a continuous u between -100 and 1000. In fact you can also perform some consistency check. For example, it should now be the case that



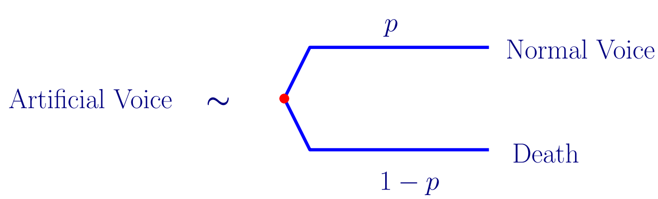
because the prospect on the right has now expected utility $0.5 \cdot u(675) + 0.5 \cdot u(100) = 0.5(0.75 + 0.25)$. If you give a number different from 400 here then you should go back and think better. In the end you should end up with a consistent assessment. If you reach that you are done, just choose the lottery with the highest expected utility.

Application: medical decisions

This is from Wakker chapter 2, which the reader may consult for accurate description of the situation. We will be brief. A patient with throat cancer must decide between Radiotherapy and Surgery. The situation is depicted in the self-explaining figure below. Note the use of squares for decision nodes and circles for chance nodes. Probabilities come from data on previous cases. The outcomes are not monetary, but rest assured that the theory can cover cases like this as well. We just assume that EU holds, with vNM u such that $u(\textit{normal voice}) > u(\textit{artificial voice}) > u(\textit{death})$.



The tree may be simple for us though most likely not for a person with cancer. But the whole point is that we don't need the patient to examine it. By normalizing $u(normal\ voice) = 1$ and $u(death) = 0$ (which we know it can always be done) we only need to elicit the patient's $u(artificial\ voice)$. And given 0-1 utility at the extremes we know that $u(artificial\ voice)$ is the probability p which solves the indifference



Given $u(artificial\ voice)$ then we compute expected utility of the two choices:

$$\begin{aligned}
Eu(Surgery) &= (.7 + .3 * .3) * u(artificial\ voice) + (.3 * .7) * 0 \\
&= .79 * u(artificial\ voice) \\
Eu(Radiotherapy) &= .6 * 1 + (.4 * .4) * u(artificial\ voice) + (.4 * .6) * 0 \\
&= .6 + .16 * u(artificial\ voice)
\end{aligned}$$

therefore the patient prefers Radiotherapy iff

$$\begin{aligned}
.79 * u(artificial\ voice) &< .6 + .16 * u(artificial\ voice) \\
u(artificial\ voice) &< .6/.63 \approx 0.95.
\end{aligned}$$

This is quite a high value, so we can expect to be able to advice a majority of patients to try radiotherapy. Still, eliciting whether $u(\textit{artificial voice})$ is above or below the threshold may not be always simple. It may be easy to find that the patient prefers the lottery with 95% probability to complete recovery to artificial voice (so that $u(\textit{artificial voice}) < 0.95$). But if the threshold were say 0.6, when confronted with the alternatives

it may be more problematic to have a definite answer. In any case it is quite clear that a decision theorist can be of huge help in situations like this.

A case study: the coach problem

This is from the book "*Thinking strategically: the competitive edge in business, politics, and everyday life*" by Avinash K. Dixit and Barry J. Nalebuff, W.W. Norton 1991.

9. Case Study #2: The Tale of Tom Osborne and the 1984 Orange Bowl

In the 1984 Orange Bowl the undefeated Nebraska Cornhuskers and the once-beaten Miami Hurricanes faced off. Because Nebraska came into the Bowl with the better record, it needed only a tie in order to finish the season with the number-one ranking.

But Nebraska fell behind by 31–17 in the fourth quarter. Then the Cornhuskers began a comeback. They scored a touchdown to make the score 31–23. Nebraska coach Tom Osborne had an important strategic decision to make.

In college football, a team that scores a touchdown then runs one play from a hash mark 2½ yards from the goal line. The team has a choice between trying to run or pass the ball into the end zone, which scores two additional points, or trying the less risky strategy of kicking the ball through the goalposts, which scores one extra point.

Coach Osborne chose to play it safe, and Nebraska successfully kicked for the one extra point. Now the score was 31–24. The Cornhuskers continued their comeback. In the waning minutes of the game they scored a final touchdown, bringing the score to 31–30. A point conversion would have tied the game and landed them the title. But that would have been an unsatisfying victory. To win the championship with style, Osborne recognized that he had to go for the win.

The Cornhuskers went for the win with a two-point conversion attempt. Irving Fryer got the ball, but failed to score. Miami and Nebraska ended the year with equal records. Since Miami beat Nebraska, it was Miami that was awarded the top place in the standings.

Put yourself in the cleats of Coach Osborne. Could you have done better?

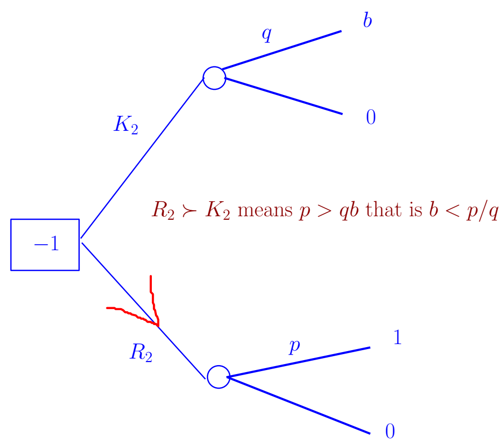
As you may have guessed a touchdown gives six points. The problem is to translate this real world problem into a decision tree which we can analyze with the tools we have learned. We let p the probability of scoring a $RUN \equiv R$ (2 points) and $q > p$ the probability of scoring a $KICK \equiv K$ (1 point).

First observation: at 31-23 you need another touchdown to draw or win. So let us put ourselves in the coach's position and assume we score a touchdown. With the touchdown we are at $-8+6=-2$ from the opposing team. So we need 2 points to draw and 3 to win.

There are two decisions to be made. K or R now - at -2 - and K or R after next touchdown. We call these decisions K_1, R_1 and K_2, R_2. These are choices of lotteries.

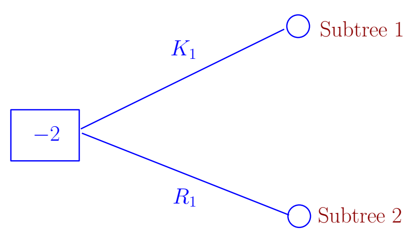
Next, possible outcomes: lose, draw, win. So let us take $u(\ell)=0, u(w)=1$ and $u(d)=b$ where $0<b<1$.

The last choice, represented below,

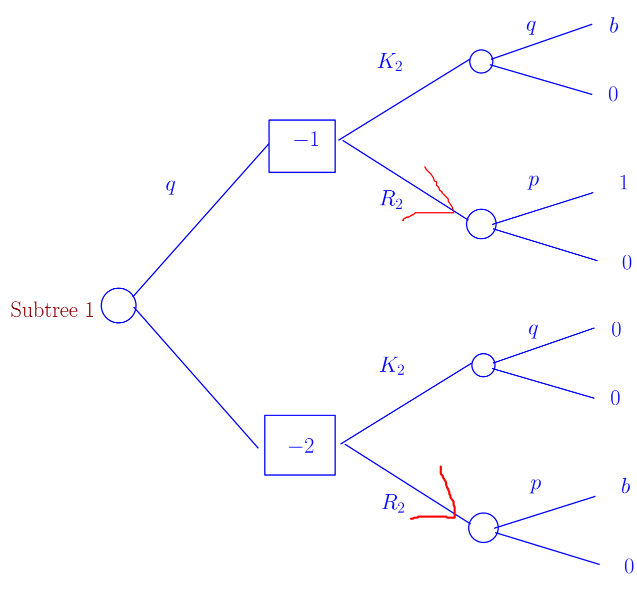


depends on preferences, hence it cannot be criticized. We have to examine the coach's first choice, between K_1 and R_1.

The first choice is at -2:



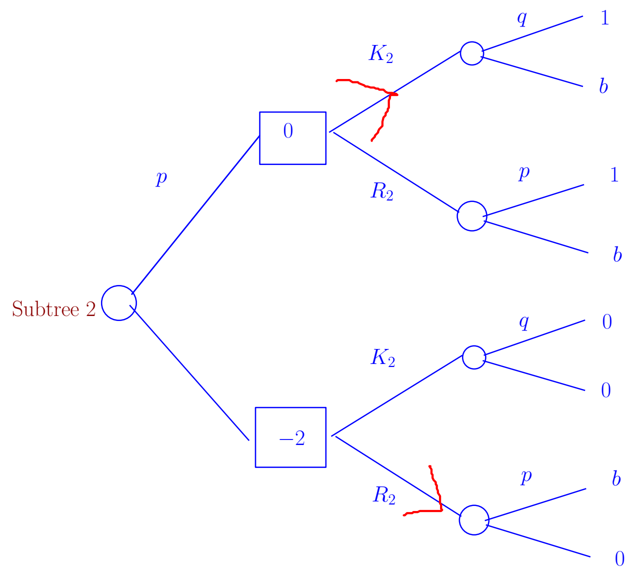
Then we split the tree. After choice K_1 we get



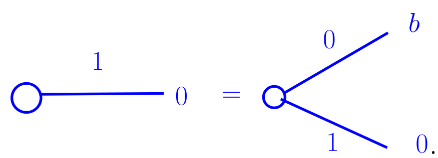
The arrows indicate the subsequent choices (R_2 at the top node because of preferences, and at the bottom node by monotonicity). So K_1 gives expected utility

$$Eu(K_1) = q * p * 1 + (1-q) * p * b = p\,[q + (1-q)b]$$

Now R_1, depicted below:



where K_2 at node 0 follows from monotonicity and R_2 at node -2 also because K_2 gives zero for sure so we can write it as



So R_1 gives expected utility

$$Eu(R_1) = p[q + (1 - q)b] + (1 - p)pb$$

Now let us compare them:

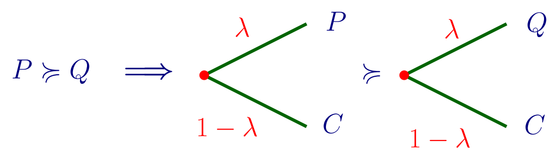
$$Eu(R_1) = Eu(K_1) + (1 - p)pb > Eu(K_1).$$

So the mistake was to choose K_1. The coach should have chosen to Run when they were at -8. The moral of the story is: look ahead to choose what to do now.

3 Caveat: the Allais experiment

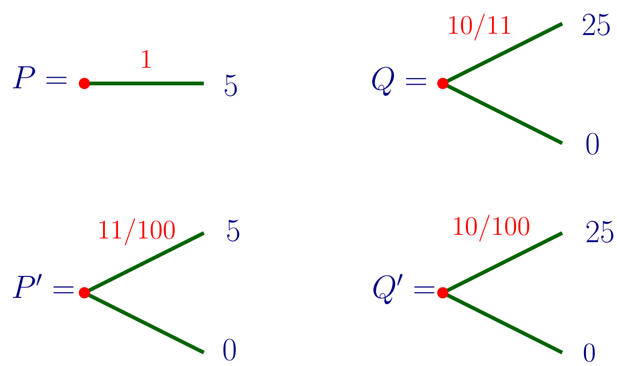
Maurice Allais (Nobel 1988) produced an example of "reasonable" choices incompatible with EU. Incompatibility is most easily seen through violation of the following axiom which is implied by EU.

The independence axiom This axiom says that for any distributions P, Q, C and number $0 \le \lambda \le 1$ the following implication holds:



To see that EU implies independence it suffices to recall that $Eu(P_\lambda C) = \lambda Eu(P) + (1-\lambda)Eu(C)$, same with P replaced by Q.

The example The preferences displayed in this example violate independence, hence are incompatible with EU. We are confronted with the two pairs of lotteries drawn below:



First we imagine the numbers are Euros. Then we make our choices if they represent millions of Euros. Usually the choices are different. For the smaller amounts the typical choices are the following:

$$P \prec Q \quad \text{and} \quad P' \prec Q'$$

But for the larger amounts many express the following preferences:

$$P \succ Q \quad \text{and} \quad P' \prec Q'$$

which are "reasonable" since it is hard to give up 5 million Euros for sure. However we show presently that the latter choices violate expected utility, for they violate independence. To see this let C be the lottery which gives zero for sure and observe that $P' = P_{11/100}C$ and $Q' = Q_{11/100}C$: the former is clear, and for the latter note that the support of $Q_{11/100}C$ is $\{25, 0\}$ and the probability of zero is $\frac{11}{100} \cdot \frac{1}{11} + \frac{89}{100} \cdot 1 = \frac{90}{100}$. Independence then implies $P \succ Q \Longrightarrow P' \succ Q'$.

4 Attitudes towards risk for general, monotone weak orders $\succcurlyeq$

Recall that monotonicity implies that for any pair of numbers x, y we have $x > y \iff x \succ y$. We assume this, and in addition that for any P the certainty equivalent $CE(P) \sim P$ (unique by monotonicity) exists.

Risk aversion We say that $\succcurlyeq$ is *risk averse* (RA) if $EP \succ P$ for any non-degenerate P.

That is, a risk averse decision maker strictly prefers the sure amount $EP = \sum_i x_i p_i$ to the distribution P. For example, she would refuse to enter a bet on a fair coin.

The degree of risk aversion of different preferences can be compared. We say that $\succcurlyeq_2$ is *more risk averse than* $\succcurlyeq_1$ if $x \sim_1 P \Rightarrow x \succ_2 P$, for any $x \in \mathbb{R}$ and P.

Risk loving There is no universally accepted term for this: risk attraction, risk loving, etc. The obvious definition is that $\succcurlyeq$ is *risk loving* if $P \succ EP$ for any non-degenerate P.

A risk lover would, for example, pay to bet on a fair coin. It is a fairly uncommon case, but we do observe risk attraction in some cases: indeed people entering casinos or buying lottery tickets enter unfavorable bets. Casinos are a special case in that people may derive utility from being there. The case of lotteries is more intriguing, and well studied: lotteries involve long shot bets which could change the individual's life at a typically small cost, and can be rationalized. But keep in mind that it is usually the case that the same individuals would not accept to bet non trivial amounts of money on a fair coin, so may be regarded as risk averse after all.

Risk neutrality You can guess the definition: $\succcurlyeq$ is *risk neutral* if $EP \sim P$ for any P.

Observe that risk neutrality implies that $\succcurlyeq$ is represented by expected value: $P \succcurlyeq Q$ iff $EP \geq EQ$. *Therefore* a risk neutral decision maker is in fact an EU maximizer, with linear vonNeumann utility $u(x) = x$ (indeed any $\tilde{u}(x) = \sigma x + \tau$ with $\sigma > 0, \tau \in \mathbb{R}$).

A firm maximizing expected profits is in fact assumed to be risk neutral; this is a fairly common assumption. However keep in mind that for a generic decision maker risk neutrality incurs in the St. Petersburg paradox (Wakker sec. 2.5).

Risk premium The *risk premium* of P is the number $\rho(P)$ such that $EP - \rho(P) \sim P$.

The following are direct consequences of the definitions:

1. $\rho(P) = EP - CE(P)$
2. $\succcurlyeq$ is risk averse $\Longrightarrow CE(P) \sim P \prec EP \Longrightarrow$ (monotonicity) $CE(P) < EP \Longrightarrow \rho(P) > 0$
3. $\succcurlyeq_2$ is more risk averse than $\succcurlyeq_1 \iff CE_2(P) < CE_1(P) \iff \rho_2(P) > \rho_1(P)$.[9]

[9] Taking $x = CE_1(P)$ in the definition $CE_1(P) \succ_2 P$, whence the result by monotonicity.

Of course if $\succcurlyeq$ is risk lover everything is reversed: $CE(P) > EP$ and $\rho(P) < 0$. And for risk neutral preference $CE(P) = EP$ and $\rho(P) = 0$.

Application: risk premium and insurance Suppose a risk averse individual faces a risk of an accident causing damage $d > 0$ with probability π. This means she "owns" the distribution P which assigns probability $1-\pi$ to zero and π to $-d$; of course $EP = -\pi d < 0$ (to fix ideas you may take $EP = -50$). By risk aversion her certainty equivalent $CE(P) = EP - \rho(P) < EP$ will be less than EP (again to fix ideas suppose $\rho(P) = 5$ so $CE(P) = -55$). This means she is willing to pay $-CE(P) = -EP + \rho(P) = \pi d + \rho(P)$ (with the above numbers this is 55) to get rid of P and eliminate the risk - that is to acquire full insurance. Suppose an insurance company insures N identical customers paying that price. If their accidents are independent of one another, by the Law of Large Numbers with probability 1 the fraction of accidents will be very close to π for N large so the cost per customer will be approximately $(N\pi \cdot d)/N = \pi d$ ($= 50$ in our example) and the company's profit is $\pi d + \rho(P) - \pi d = \rho(P)$ (in the example $50 + 5 - 50 = 5$). In other words, the risk premium is approximately the insurance company's profit per customer. This is why it's called risk premium.[10]

Pareto improvement via transfer of risk Generalizing the previous example, if $\succcurlyeq_2$ is more risk averse than $\succcurlyeq_1$ then a non-degenerate lottery P can be sold from Mr.2 to Mr.1 for any price p such that $C_2(P) < p < C_1(P)$. This is a general principle: it is always advisable to transfer risks from more risk averse agents to less risk averse ones.

4.1 Financial markets with risk averse traders[11]

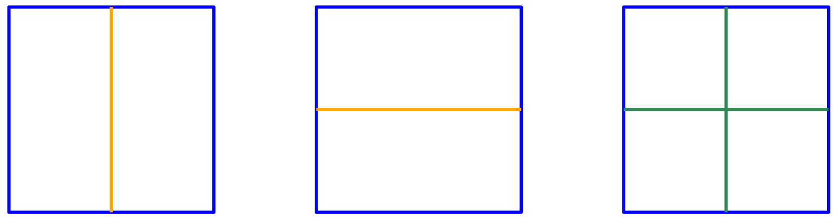
The purpose of this section is to show that when traders in a financial market are risk averse, no trader can have strictly positive expected gains.

Premise: meet of information partitions On a set Ω a partition $\mathcal{P}$ is a family of nonempty disjoint sets whose union is Ω. A partition can be interpreted as information: you know in which element of $\mathcal{P}$ the realized ω lies. Given two partitions $\mathcal{P}$ and $\mathcal{Q}$ their *meet* $\mathcal{P} \wedge \mathcal{Q}$ is defined as

$$\mathcal{P} \wedge \mathcal{Q} = \{A \cap B \neq \emptyset \colon A \in \mathcal{P}, B \in \mathcal{Q}\}.$$

It corresponds to the aggregate information in $\mathcal{P}$ and $\mathcal{Q}$: you know which element of $\mathcal{P}$ *and* which element of $\mathcal{Q}$ are true. Picture:

[10] If you have not studied the Law of Large Numbers before you may read about it on Wikipedia.
[11] From Jean Tirole, *Econometrica* 1982



$\mathcal{P}$ $\quad$ $\mathcal{Q}$ $\quad$ $\mathcal{P} \wedge \mathcal{Q}$

The meet of partitions $\mathcal{P}_1, \ldots, \mathcal{P}_I$ is defined inductively. It reflects the information aggregated over the given partitions. If $\mathcal{P}_i$ is individual i's information the meet results if all the individuals share their information among one another.

The market There are $i = 1, \ldots, I$ individuals trading an asset ξ which is a random variable on a space $(\Omega, \mathcal{F}, P)$; if the asset price is p and i trades x_i units (positive if she buys, negative if she sells) her gain is $G_i = x_i(\xi - p)$. Traders are expected utility maximizers, with concave vNM utility on gains. Trader i has information partition $\mathcal{A}_i$, and we let $\mathcal{A}$ denote the meet of the individual partitions. The market clears when $\sum_i x_i = 0$. Note that when markets clear $\sum_i G_i = 0$ as well.

Rational Expectations Equilibrium The idea is that price reveals information, and rationality consists of correctly using this information when choosing x_i. It is assumed that the asset price depends on the aggregate information $\mathcal{A}$ according to a function $\Phi : \mathcal{A} \to \mathbb{R} \ni p$. The information contained in p is the counterimage $\Phi^{-1}(p) = \{A \in \mathcal{A} \colon \Phi(A) = p\}$, which we denote by $S(p)$. Trader i's information then consists of $(A_i, S(p))$ where A_i is the known element of $\mathcal{A}_i$ and $S(p)$ is the public information which the rational trader takes into account.

A *Rational Expectation Equilibrium* is a function Φ and trades $x_i(p, A_i, S(p))$, $i = 1, \ldots, I$ such that: $x_i(p, A_i, S(p))$ maximizes expected utility of i conditional on $(A_i, S(p))$ and markets clear, that is $\sum_i x_i(p, A_i, S(p)) = 0$. In equilibrium trader i's expected gain is $E(G_i \mid A_i, S(p))$.

Proposition (***Impossibility of speculation***)**.** *$E(G_i \mid A_i, S(p)) = 0$ for each i if traders are not risk lovers.*

Proof. For risk averse traders it must be $E(G_i \mid A_i, S(p)) > 0$, and for risk neutral ones $E(G_i \mid A_i, S(p)) \geq 0$; so without risk lovers we must have $E(G_i \mid A_i, S(p)) \geq 0$ for all i. By iterating conditional expectation we then get $E(G_i \mid S(p)) = E\left[E\left(G_i \mid A_i, S(p)\right) \mid S(p)\right] \geq 0$. But from $\sum_i G_i = 0$ we deduce $\sum_i E(G_i \mid S(p)) = E(\sum_i G_i \mid S(p)) = 0$; so since each term is non-negative, each must be zero. But $E\left[E(G_i \mid A_i, S(p)) \mid S(p)\right] = 0$ and $E(G_i \mid A_i, S(p)) \geq 0$ imply $E(G_i \mid A_i, S(p)) = 0$. ☐

5 Risk attitudes for EU Preferences

We now turn to the analysis of risk attitudes in the important case where EU holds. We assume that all vNM utilities possess derivatives of all orders.

Risk aversion Risk aversion is characterized by the shape of the vNM utility:

Proposition. *For EU preferences risk aversion is equivalent to concavity of vNM utility.*

Proof. Recall that a function u is concave if for any $x_1, x_2 \in \mathbb{R}$ and $0 \leq p_1, p_2$ with $p_1 + p_2 = 1$ it is $u(p_1 x_1 + p_2 x_2) > p_1 u(x_1) + p_2 u(x_2)$. It can be shown by induction that this is equivalent to: for any $x_1, \ldots, x_n \in \mathbb{R}$ and $0 \leq p_1, \ldots, p_n$ with $\sum p_i = 1$ it is $u(\sum p_i x_i) > \sum p_i u(x_i)$. This says exactly that for distribution P (values x_i, probabilities p_i) $u(EP) > Eu(P)$ that is (by definition) $EP \succ P$.[12] $\square$

It is instructive to visualize the situation considering a distribution P concentrated on the two points $\alpha < \beta$, with probability p on α (and $1 - p$ on β); its expected value is $EP = p\alpha + (1 - p)\beta$. Note that $Eu(P) = pu(\alpha) + (1 - p)u(\beta)$ is the value on the line at EP (see Appendix 2 on straight lines). All the relevant quantities appear in the picture below.

[12] Sketch of the induction step: suppose it is true for two-outcome lotteries. Take a three-outcome lottery $x = (p_1 x_1, p_2 x_2, p_3 x_3)$; write it as

$$x = (p_1 x_1, (1 - p_1)\left(\frac{p_2}{1 - p_1} x_2, \frac{p_3}{1 - p_1} x_3\right))$$

and apply the result for two-outcome lotteries; you get

$$u(p_1 x_1 + p_2 x_2 + p_3 x_3) \geq p_1 u(x_1) + (1 - p_1) u(\left(\frac{p_2}{1 - p_1} x_2, \frac{p_3}{1 - p_1} x_3\right))$$

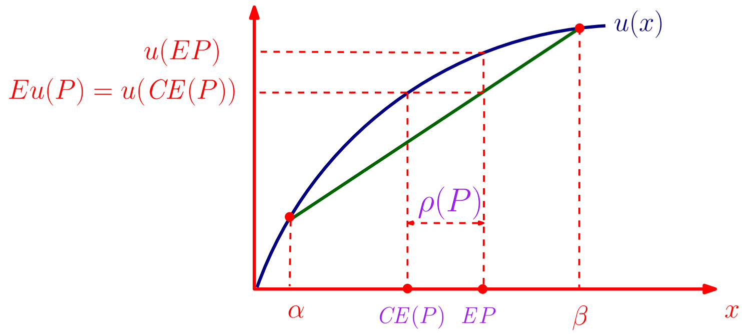
but $\frac{p_2}{1-p_1} x_2, \frac{p_3}{1-p_1} x_3$ is a two-outcome lottery because $\frac{p_2}{1-p_1} + \frac{p_3}{1-p_1} = 1$ so we can re-apply the two-outcome lottery result

$$\begin{aligned} p_1 u(x_1) + (1 - p_1) u(\left(\frac{p_2}{1 - p_1} x_2, \frac{p_3}{1 - p_1} x_3\right)) &\geq p_1 u(x_1) + (1 - p_1)\left(\frac{p_2}{1 - p_1} u(x_2) + \frac{p_3}{1 - p_1} u(x_3)\right) \\ &= p_1 u(x_1) + p_2 u(x_2) + p_3 u(x_3) \end{aligned}$$

so we see that

$$u(p_1 x_1 + p_2 x_2 + p_3 x_3) \geq p_1 u(x_1) + p_2 u(x_2) + p_3 u(x_3)$$

which means the result is true for three-outcome lotteries. Then analogously go from $n - 1$ to n, and complete the proof.



Of course risk attraction corresponds to convex utility. Risk neutrality we have already observed that corresponds to linear utility.

Exercise Consider the prospect $P = 100_{1/2}0$. Calculate its risk premium for $u(x) = 1 - e^{-x/100}$ and $u(x) = \sqrt{x}$. Answers: for the first $\rho(P) \approx EP/4$, for the second $\rho(P) = EP/2$. Solution in footnote.[13]

Application: choice of insurance A risk averse individual with initial wealth w may incur a loss $d \leq w$ with probability π. Now assume she can buy insurance at price p; this means if she buys α units of insurance she pays $p\alpha$ in advance and gets α in case the loss occurs. Thus her wealth will be $w - p\alpha$ with probability $1 - \pi$ and $w - p\alpha - d + \alpha = w - d + (1-p)\alpha$ with probability π. Observe that the insurer, that we may assume to be risk neutral, bears an expected cost π for each unit sold, so she must sell at $p \geq \pi$. Observe also that if $\alpha = d$ the individual's income is $w - pd$ for sure; thus full insurance is $\alpha = d$.

Let u denote the individual's vNM utility; she chooses α to maximize expected utility, so taking into account the budget constraint $p\alpha \leq w$ she solves

$$\max_{0 \leq \alpha \leq w/p} (1-\pi)u(w - p\alpha) + \pi u(w - d + (1-p)\alpha).$$

Let $V(\alpha)$ denote the function to be maximized. We have $V'(\alpha) = -p(1-\pi)u'(w-p\alpha) + (1-p)\pi u'(w-d+(1-p)\alpha)$; and $V''(\alpha) = p^2(1-\pi)u''(w-p\alpha) + (1-p)^2\pi u''(w-d+(1-p)\alpha) < 0$, so V' is decreasing. Let α^* be the optimal choice. At $p = 1$ we have $V'(0) < 0$, while at $p = \pi$ it is $V'(0) = \pi(1-\pi)\left[-u'(w) + u'(w-d)\right] > 0$. Therefore $\alpha^* = 0$ for p close enough to 1, otherwise $\alpha^* > 0$. At $\alpha = d$ we have $V'(d) = u'(w-pd)\left[-p(1-\pi) + (1-p)\pi\right]$ so if $p = \pi$ then $V'(d) = 0$ and $\alpha^* = d$: if $p = \pi$ the optimal choice is full insurance. If on the other hand $p > \pi$ then $V'(d) < 0$ (the expression $-p(1-\pi) + (1-p)\pi$ is decreasing in p) so $\alpha^* < d$: for

[13] For the first, $Eu(P) = (1 - e^{-1})/2 \approx 0.316$; next, $CE(P)$ solves $1 - e^{-x/100} = Eu(P) \approx 0.316$ so $CE(P) = -100 \cdot \ln(1 - Eu(P)) \approx 37.99$ so $\rho(P) = 50 - CE(P) \approx 12.01$ which is almost $EP/4$. For the second, proceeding in the same way we get: $Eu(P) = 5$; the equation for $CE(P)$ is $\sqrt{x} = 5$ which gives $CE(P) = 25$; and finally $\rho(P) = 50 - 25 = 25 = EP/2$.

any $p > \pi$ the individual chooses to bear some risk (because $w - p\alpha - d + \alpha < w - p\alpha$ that is wealth in the bad state is lower than in the good state).

Comparative risk aversion Turning to comparative risk aversion, the idea is that more RA must correspond to u "more concave". This means u' goes down faster i.e. u' steeper i.e. $-u''$ larger. But this does not work because u can be rescaled. The solution is to divide by u' (for then the rescaling factors cancel out) to get the *risk aversion coefficient at* $x \in \mathbb{R}$

$$r(x) = -u''(x)/u'(x).$$

The coefficient r is called the *Arrow-Pratt index.* This is invariant to linear transformations of u, that is it remains the same if u is replaced by $\hat{u} = \sigma u + \tau$ with $\sigma > 0$ and $\tau \in \mathbb{R}$ (exercise). Note that if u is linear, $u(x) = \sigma x + \tau$ then $r(x) = 0$ for all x.

Proposition. *Under EU, $\succcurlyeq_2$ is more risk averse than $\succcurlyeq_1$ iff $r_2(x) > r_1(x)$ for all x.*

Proof. Consider a general $P = (\{x_1, \ldots, x_n\}, (p_1, \ldots, p_n))$. First we claim that the hypothesis amounts to $u_2 = \phi \circ u_1$ for a ϕ increasing concave. Since $u_2 = \left(u_2 \circ u_1^{-1}\right) \circ u_1$ we must show $\succcurlyeq_2$ is more risk averse than $\succcurlyeq_1$ iff $\phi = u_2 \circ u_1^{-1}$ is concave (it is clearly increasing since both u_2 and u_1^{-1} are). Now $x \sim_1 P \Rightarrow x \succ_2 P$ says $u_1(x) = \sum p_i u_1(x_i) \Rightarrow u_2(x) > \sum p_i u_2(x_i)$; this in turn means $u_2\left(u_1^{-1}\left(\sum p_i u_1(x_i)\right)\right) > \sum p_i u_2(x_i)$, since $u_1(x) = \sum p_i u_1(x_i) \iff x = u_1^{-1}\left(\sum p_i u_1(x_i)\right)$. The inequality can be written as $\phi\left(\sum p_i u_1(x_i)\right) > \sum p_i \phi\left(u_1(x_i)\right)$, and letting $z_i = u_1(x_i)$ in turn as $\phi\left(\sum p_i z_i\right) > \sum p_i \phi\left(z_i\right)$; this shows that is ϕ concave.

Now differentiate $u_2 = \phi \circ u_1$. We obtain $u_2' = \phi'(u_1) \cdot u_1'$ and $u_2'' = \phi'' \cdot (u_1')^2 + \phi'(u_1) \cdot u_1''$. Thus with $\phi' > 0$ and $\phi'' < 0$ (that is ϕ increasing concave) we get

$$r_2(x) = -\frac{u_2''}{u_2'} = -\frac{u_1''}{u_1'} - \frac{\phi''}{\phi'} \cdot u_1' > -\frac{u_1''}{u_1'} = r_1(x)$$

as was to be shown. □

Application: a simple portfolio choice You are risk averse, and you have to allocate your wealth w between two assets. The first is risk free: it pays 1 Euro for sure for each Euro invested; the second is risky. Its return is a random variable P, but it has the advantage that its expectation is larger than 1: $EP > 1$. By construction, if you invest α in the latter and β in the safe asset you get the (random) return $\alpha P + \beta$. You have to choose $\alpha, \beta \geq 0$ to maximize expected utility $Eu(\alpha P + \beta)$. Since it must be $\alpha + \beta = w$ the problem may be written in terms of α alone:

$$\max_{0 \leq \alpha \leq w} Eu(w + \alpha(P - 1)).$$

To apply calculus we want to differentiate with respect to α, and we *assume* that the derivative of an expectation is equal to the expectation of the derivative.[14] Then letting $V(\alpha) = Eu(w + \alpha(P-1))$ we get $V'(\alpha) = E\left[(P-1)u'(w+\alpha(P-1))\right]$ and $V''(\alpha) = E\left[(P-1)^2 u''(w+\alpha(P-1))\right]$; observe that $u'' < 0$ implies $V'' < 0$, so V' is decreasing and an interior optimum is characterized by the first order condition.

Our first observation is that the optimal $\alpha^* > 0$, that is you choose to bear some risk. This simply follows from $V'(0) > 0$ and $EP > 1$. Note that this choice is analogous to that of bearing some risk in the insurance choice whenever $q > \pi$.

The other point we make is on how choice depends on the degree of risk aversion. It should be obvious that a risk neutral individual chooses $\alpha^* = w$ - please verify it. We shall next check that the more risk averse you are the less you will buy of the risky asset. We have to show that if u_2 is more risk averse than u_1 then $\alpha_2^* < \alpha_1^*$. Assuming interior solutions, the optima are characterized by $V_1'(\alpha_1^*) \equiv E\left[(P-1)u_1'(w+\alpha_1^*(P-1))\right] = 0$ and $V_2'(\alpha_2^*) \equiv E\left[(P-1)u_2'(w+\alpha_2^*(P-1))\right] = 0$. Since V_2' is decreasing it suffices to show that $V_2'(\alpha_1^*) < 0$. From the proof on page 20 we know that $u_2 = \phi(u_1)$ with ϕ concave hence ϕ' decreasing; and by composition $u_2'(x) = \phi'(u_1(x)) \cdot u_1'(x)$, so

$$V_2'(\alpha_1^*) \equiv E\left[(P-1)\phi'\left(u_1(w+\alpha_1^*(P-1))\right) u_1'(w+\alpha_1^*(P-1))\right].$$

On the other hand by multiplying the equality $V_1'(\alpha_1^*) = 0$ by $\phi'\left(u_1(w)\right)$ we get

$$0 = E\left[(P-1)\phi'\left(u_1(w)\right) u_1'(w+\alpha_1^*(P-1))\right].$$

But $\phi'\left(u_1(w+\alpha_1^*(P-1))\right) \lesseqgtr \phi'\left(u_1(w)\right) \iff P \gtreqless 1$, so going from $\phi'\left(u_1(w)\right)$ to $\phi'\left(u_1(w+\alpha_1^*(P-1))\right)$ lowers the integrand in both ranges, making $V_2'(\alpha_1^*) < 0$ as we wanted.

An approximate expression for the risk premium Again we consider a general $P = (\{x_1,\ldots,x_n\},(p_1,\ldots,p_n))$. There is a basic relation between r, ρ and the variance $\sigma^2(P)$ of P, which is defined as $E(P-EP)^2 = \sum_i (x_i - EP)^2 p_i$.

Proposition. $\rho(P) \approx \frac{1}{2}\sigma^2(P) \cdot r(EP)$

Proof. Let $z_i = x_i - EP$, so that $\sigma^2(P) = \sum p_i z_i^2$ (notice that $\sum p_i z_i = 0$). By definition $u\left(EP - \rho(P)\right) = u(CE(P)) = Eu(P) = \sum p_i u(x_i) = \sum p_i u(EP + z_i)$. Now expand around EP. Left side: $u\left(EP - \rho(P)\right) \approx u\left(EP\right) - \rho(P)u'(EP)$; right side:

$$\begin{aligned}\sum p_i u(EP+z_i) &\approx \sum p_i \left[u(EP) + z_i u'(EP) + \left(z_i^2/2\right) u''(EP)\right] \\ &= \sum p_i \left[u(EP) + \left(z_i^2/2\right) u''(EP)\right] = u(EP) + u''(EP)\sigma^2(P)/2.\end{aligned}$$

[14] This is certainly true if P is discrete, but it is true in much more general cases.

Collecting terms we get $-\rho(P)u'(EP) \approx u''(EP)\sigma^2(P)/2$, which is what we wanted. $\square$

Application: risk sharing Consider two prospects P, Q with $EP = EQ = \mu$ and $Var(P) = \sigma^2$, $Var(Q) = \sigma^2/n$. Recalling that $P \sim EP - \rho(P)$, to a first approximation the proposition above gives $P \sim \mu - \frac{1}{2}\sigma^2 \cdot r(\mu)$ and $Q \sim \mu - \frac{1}{2n}\sigma^2 \cdot r(\mu)$, so if you have a sure income z_0 with $\mu - \frac{1}{2}\sigma^2 \cdot r(\mu) < z_0 < \mu - \frac{1}{2n}\sigma^2 \cdot r(\mu)$ then $Q \succ z_0 \succ P$. If n identical individuals facing *independent* prospects $P_i = P$ share the risk, in the sense of investing z_0/n on each P_i and getting $S_i = P_i/n$ (outcomes $1/n$ of those of P_i) for each i, then each individual faces n identical prospects S_i, each with $ES_i = \mu/n$ and $Var\, S_i = \sigma^2/n^2$, that is in total each spends z_0 and gets the prospect $S = \sum_i S_i$ which has $ES = \mu$ and $Var(S) = \sigma^2/n$ which is equal to Q and is therefore preferred to z_0. In practice, by this risk sharing agreement each member gets Q at price z_0.

6 Exponential and Power vNM utilities

These are the two most prominent families of vNM utilities use in theory and in applications.

The exponential family The exponential family is $u(x) = -e^{-\theta x}$ where we assume $\theta > 0$. You can verify that $r(x) = \theta$ for all x in this case. It is convenient to rescale the function to $u(x) = (1 - e^{-x\theta})/\theta$ so that as $\theta \to 0$ we get $u(x) = x$. The family is called CARA - constant absolute risk aversion.

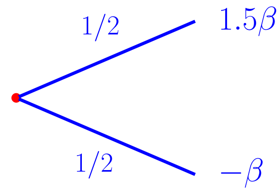
Relative risk aversion To introduce the power family we need a last new concept: *relative* risk aversion. We can take proportional increments of wealth tx with t close to 1 instead of absolute increments $x + h$ with h small. To obtain a measure analogue to r in this setting we let $\tilde{u}(t) = u(tx)$ and consider the Arrow-Pratt index for $\tilde{u}$, that us $-\tilde{u}''(t)/\tilde{u}'(t)$ at $t = 1$. We have $\tilde{u}'(t) = xu'(tx), \tilde{u}''(t) = x^2u''(tx)$ so for $t = 1$ we obtain the index

$$\tilde{r}(x) = -\frac{\tilde{u}''(1)}{\tilde{u}'(1)} = -x \cdot \frac{u''(x)}{u'(x)} = xr(x)$$

Note that if $r(x) > 0$ and $\tilde{r}(x)$ remains bounded above as x becomes large then $r(x)$ tends to zero.

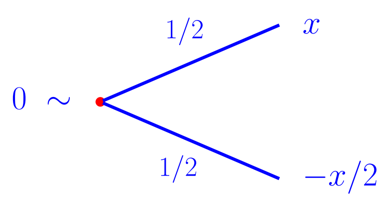
The power family The power family is $u(x) = x^\theta$, with $x > 0$ and $0 < \theta < 1$. We have $u' = \theta x^{\theta-1}, u'' = \theta(\theta - 1)x^{\theta-2}$ so that $\tilde{r}(x) = 1 - \theta$ for all x. For $\theta = 1$ we get risk neutrality; and as θ goes down relative risk aversion increases. By rescaling to $u(x) = (x^\theta - 1)/\theta$ we see that as $\theta \to 0$ we get $u(x) = \ln x$. The family is called CRRA - constant relative risk aversion.

CARA utility: investing Assume vNM $u(x) = 1 - e^{-x\theta}$. Suppose you can invest the amount β in a stock whose yield is represented by the prospect



How much do you invest to maximize expected utility? The result is $\beta \approx 0.16/\theta$ (solution in footnote):[15] optimal investment is inversely proportional to risk aversion.

CARA utility: assessing the parameter If you reckon your risk aversion is constant over the range relevant to a decision problem (and subscribe to the expected utility axioms), then discovering your utility amounts to determining the parameter θ. This is "relatively" easy, because it is equivalent to finding x such that the indifference below holds:



Indeed, with $u(x) = -e^{-\theta x}$ (the constant does not matter) the above indifference says $e^{-\theta x} + e^{\theta x/2} = 2$, and this is solved by $x \approx 1/\theta$ because $e^{-1} + e^{1/2} \approx 2$.

7 Practice on decision trees: initial investments

We now study variations of a rather common problem firms often face. The basic structure is that the level of an investment must be chosen before its productivity is known - and even if the investment turns out to be productive, further uncertainty comes from the state of the market demand for the firm's product. We frame the problem in terms of an R&D investment by a pharmaceutical firm to be concrete, but it should be clear that the structure is common to a fairly large range of situations business firms may find themselves in. *We assume expected value maximization* (risk neutrality) to keep computations simple.

[15] You want to maximize $0.5 \cdot u(1.5\beta) + 0.5 \cdot u(-\beta) = 0.5 * (1 - e^{-1.5\theta\beta} + 1 - e^{\beta\theta})$. Equivalently, minimize $e^{-1.5\theta\beta} + e^{\beta\theta}$. Set derivative equal to zero:

$$0 = -1.5\beta \cdot e^{-1.5\theta\beta} + \beta e^{\beta\theta} \iff 1.5 = e^{2.5 \cdot \beta\theta} \iff 2.5 \cdot \beta\theta = \ln 1.5$$

whence $\beta = \ln 1.5/2.5\theta \approx 0.16/\theta$ since $\ln 1.5 \approx 0.4$. This is the solution (the second derivative is positive).

Version 1 A pharmaceutical firm - you, to ease exposition - faces a decision concerning the introduction of a new drug. The first step is to choose a level of R&D investment, say High or Low (H or L). Under H there is probability p of successful development; in case of failure - probability $1-p$ - you lose 200 (payoff -200) and that's the end of the story. Under L probability of success is $p-\delta$ and failure results in losing 100. In case of success (under either H or L) you have to make another choice: market or not market the new product - M or N. After H the situation is as follows: if you don't (i.e. choose N) you lose 200; if you do (choice M) you face an uncertain state of demand: either good, G, with probability $s = 0.6$ in which case you get 2000, or bad B in which case you lose 600. After L the numbers are: if you choose N you lose 100, if you choose M you get 2100 in state G (same probability s) and -500 in state B. The payoffs are consistent with interpreting investment as costing 200 or 100 if high or low. Show that you choose H if $\delta > 100/1160 \equiv \delta_1$. I invite you to label DH and DL the decision nodes after success if the initial decision was respectively H or L.

Version 2 Suppose now that if the investment is successful you can choose to buy a survey on the state of demand at cost c before choosing on M, N. The outcome of the survey is either g (the state looks good) or b (the state looks bad). Of course the test is not perfect, its result carries some uncertainty: $P(g \mid G) = 0.95$ and $P(g \mid B) = 0.2$.

Now the decision *after success* changes: for example DH becomes say DH', where you can either not buy the survey - choice *NS* - in which case the expected payoff of the continuation remains the same as in Version 1; or buy the survey - choice S - in which case you first face the uncertain outcome of the survey - probabilities $P(g)$ and $P(b)$ to be computed using Bayes rule - and then in either case you then choose whether to market the product or not - choices M or N as before. If you choose N you lose $200 + c$; if you choose M you face uncertain demand, but now of course the probability of good state is not s but $p(G \mid g)$ or $P(G \mid b)$ according to the result of the survey (again computed via Bayes rule). The payoffs are as before but with c subtracted, that is $2000 - c$ and $-600 - c$ in good and bad states. The node DL changes analogously to DL', you should now be able to easily draw the corresponding tree.

Show first that in both DH' and DL' you choose to buy the survey if $c < 67.2$, and assume in the sequel that this is the case. If for some reason you did not reach that result *just assume* the choice is S both in DH' and DL'. Next show that in this case the initial investment decision is H if $\delta > 100/(1227.2 - c) \equiv \delta_2$.

Observe that $c < 67.2$ implies that $\delta_2 < 100/1160 = \delta_1$. Thus if the cost of the survey is low enough that it is optimal to buy it, the possibility of getting some information on the state of demand expands the range of parameters for which it is optimal to choose the high investment level H at the initial node.

Version 3 The basic structure is the same as in Version 1: first investment decision H or L; then chance node success/failure, then M/N marketing choice if success, then chance node G/B demand if M chosen. But we change the nature of the investments: the low investment choice leads to a less ambitious project, with scaled down costs and revenues, but with higher probability of success: $p+\delta$ (it was $p-\delta$ in the previous versions).

We assume for simplicity $p = s = 0.5$. So at the last chance node demand is in good or bad state with equal probabilities, and after the first H/L decision node the probability of success is $1/2$ for H and $1/2+\delta$ for H (with $0 < \delta < 1/2$).

Under initial choice H we assume loss of 50 if failure and if N chosen at the marketing choice node; and for demand we assume you get 1000 in good state and -900 in bad state (with $s = 1/2$ this has expected value of 50). Under initial choice L all payoffs are divided by 10 (so they become -5, 100 and -90). It is then easy to conclude that L is chosen at the initial node now. Again let's denote by DH and DL the marketing decision nodes corresponding to successful H and L.

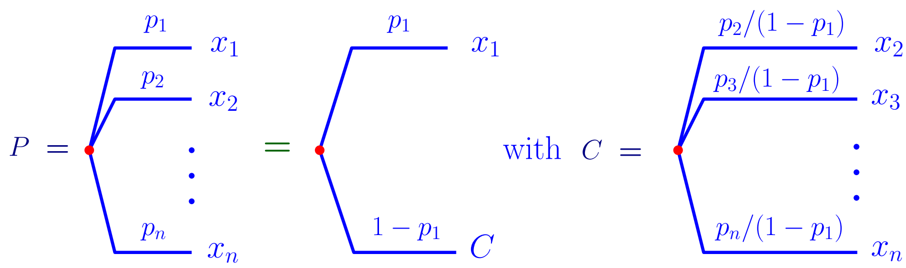
Now we again introduce the possibility of buying a demand survey at cost c if investment is successful, giving as before g or b answer. Of course, given that revenues under L are much smaller than under H it may be the case that it is worth taking the survey only under the larger project. We shall confirm that there is a range of c's for which this is actually the case. We assume $P(g \mid G) = 0.9$ and $P(g \mid B) = 0.2$. Again the marketing nodes become DH' and DL' to include the survey choice S/NS. And again you have to update probabilities using Bayes rule.

Show that at the DH' node you choose S if $c \leq 289.4$ while at DL' you choose S if $c \leq 28.94$. Assume that $28.94 < c < 289.4$ - so that you choose S only under H. If you did not find the above bounds on c, just assume you choose S only under H.

Then show that for $28.94 < c < 279.4$ you choose H at the initial node for all $0 < \delta < 1/2$. Now the bigger scale works in favor of the high investment even if the success probability is smaller than in the low investment case. The survey's cost becomes small relative to the scale of the larger project, and with the results on hand you enter the market only if the (updated) probability of good state is high.

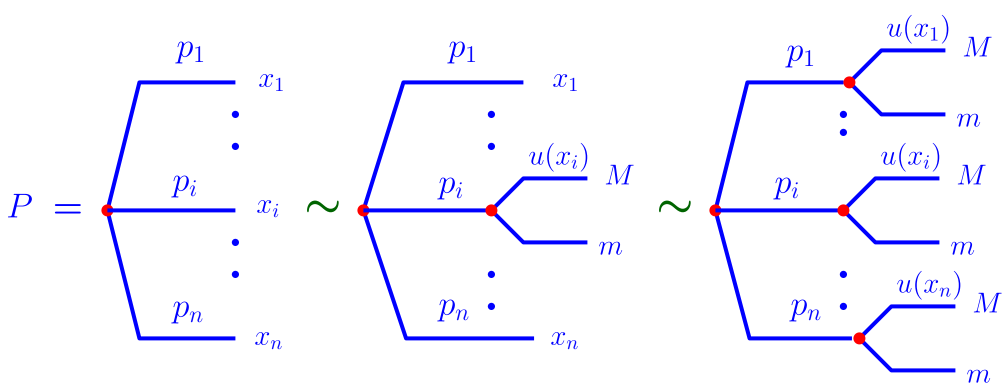
Appendix 1: proof of the EU theorem

The splitting step in the proof We can write P as a mixture as follows:

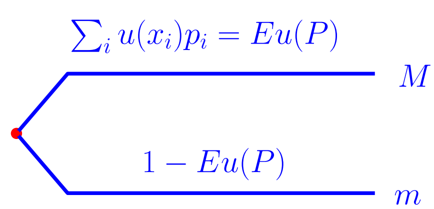


You can easily check that the equality (the second, in green) is true. In the proof a splitting like this is applied repeatedly.

The proof on $[m, M]$ The result of the procedure - for distributions with values in the interval $[m, M]$ - is the following, where we use consistency and the definition of u at each step:



But the latter is the two-outcome distribution



so for any P we have $P \sim M_{Eu(P)}m$. By transitivity we get $P \succcurlyeq Q$ iff $M_{Eu(P)}m \succcurlyeq M_{Eu(Q)}m$; and then from monotonicity we deduce $P \succcurlyeq Q \iff Eu(P) \geq Eu(Q)$.

Extension and uniqueness Extension from $[m, M]$ to $\mathbb{R}$. Call m_0, M_0 the starting pair above and u_0 the vNM utility obtained on $[m_0, M_0]$, with $u_0(m_0) = 0, u_0(M_0) = 1$. We repeat the process for a sequence of pairs m_n, M_n with $m_n \to -\infty$ and $M_n \to \infty$, each time selecting u_n on $[m_n, M_n]$ such that $u_n(m_0) = 0, u_n(M_0) = 1$. The lemma below implies that $u_{n+1} = u_n$ on $[m_n, M_n]$, for each n. Therefore the u_n's are extensions of each other, and since $\cup_n [m_n, M_n] = \mathbb{R}$ the procedure gives a vNM utility u on $\mathbb{R}$, with $u(m_0) = 0, u(M_0) = 1$. The same lemma establishes that any other vNM utility u^* for $\succcurlyeq$ must be an increasing linear transformation of u, that is of the form $u^* = \tau + \sigma u$ with $\tau \in \mathbb{R}$ and $\sigma > 0$.

Lemma. *(i) Let u be a vNM utility for $\succcurlyeq$ on a domain $A \subseteq \mathbb{R}$ which includes the interval $[m, M]$. Then $u(x)$ is uniquely determined by the two values $u(m)$ and $u(M)$, for each $x \in A$. (ii) Suppose $u(m) = 0, u(M) = 1$; then there are $\tau \in \mathbb{R}$ and $\sigma > 0$ such that any other vNM u^* is of the form $u^* = \tau + \sigma u$.*

Proof. (i) For $x \in [m, M]$ let p solve $x \sim M_p m$; then $u(x) = pu(M) + (1-p)u(m)$. For $x > M$ let q solve $M \sim x_q m$; then $u(M) = qu(x) + (1-q)u(m)$ so that $u(x) = [u(M) - (1-q)u(m)]/q$; for $x < m$ let r solve $m \sim M_r x$; then $u(m) = ru(M) + (1-r)u(x)$ so that $u(x) = [u(m) - ru(M)]/(1-r)$.

(ii) Suppose $u^*(m) = \tau, u^*(M) = \tau + \sigma$. On $[m, M]$ we know that u is defined by $x \sim M_{u(x)} m$, so $u^*(x) = u(x) * (\tau + \sigma) + (1 - u(x)) * \tau = \tau + \sigma u(x)$. Next, part (i) implies that for $x > M$, with q solving $M \sim x_q m$ we have $u(x) = 1/q$ and $u^*(x) = [\tau + \sigma - (1-q)\tau]/q = [q\tau + \sigma]/q = \tau + \sigma u(x)$; and that for $x < m$ we have $u(x) = -r/(1-r)$ and $u^*(x) = [\tau - r(\tau + \sigma)]/(1-r) = \tau + \sigma u(x)$ again. □

Appendix 2: Straight Lines

The equation for a line through (x_0, y_0) is

$$r(x) = y_0 + m(x - x_0) \tag{1}$$

where m is the *slope* of r. Note that

$$r(x+h) - r(x) = [y_0 + m(x + h - x_0)] - [y_0 + m(x - x_0)] = mh \tag{2}$$

In other words, *if* $\Delta x = h$ *then* $\Delta r = mh$. In particular, for any h we get

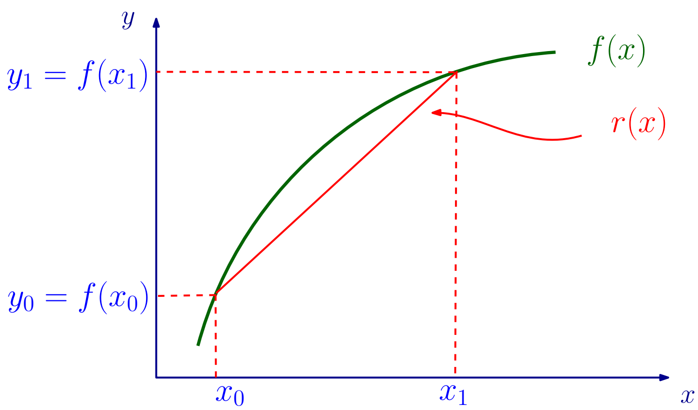
$$m = \frac{r(x+h) - r(x)}{h}. \tag{3}$$

The other point to note is that for any x_1, x_2 and $0 \le p \le 1$ we have

$$r(px_1 + (1-p)x_2) = pr(x_1) + (1-p)r(x_2). \tag{4}$$

To see this just write $y_0 = py_0 + (1-p)y_0$ and $x_0 = px_0 + (1-p)x_0$.

Now consider the line through $(x_0, f(x_0))$ and $(x_1, f(x_1))$ as in the figure below.



From (3) we see that the slope is

$$m = \frac{y_1 - y_0}{x_1 - x_0} = \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

so from (1) we have

$$r(x) = f(x_0) + \frac{f(x_1) - f(x_0)}{x_1 - x_0}(x - x_0).$$

If z is between x_0 and x_1, say $z = px_0 + (1-p)x_1$ with $0 \le p \le 1$ (please check that with these values for p we actually have $x_0 \le z \le x_1$) then from 4

$$\begin{aligned} r(px_0 + (1-p)x_1) &= pr(x_0) + (1-p)r(x_1) \\ &= pf(x_0) + (1-p)f(x_1). \end{aligned}$$