Supplemental notes on utility and investing

James G. Scott

This supplement summarizes two topics not discussed in the course packet: the utility of money, and bootstrap resampling for modeling financial assets.

1 The utility of wealth

Recall that your utility function is something that encodes your preferences between different options or 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*. If your preference relation satisfies the *von Neumann–Morgenstern rules*, then you have a *utility function,* which 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,

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

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

You are probability wondering: what does my utility function look like? Unfortunately, we can't just plot it—the function takes arbitrary outcomes in life, both monetary and non-monetary, as an input. For example, we might have $u(\$7) = 1$ and $u(\text{burrito}) = 1.1$, in which case you'd prefer the burrito to have an extra \$7 in your pocket. (Conventionally, the units of utility are called "utils," but I've always felt pretty dumb saying this.) The point is: it's hard to imagine how we could plot a function whose domain includes burritos and cash, in addition to all other non-cash, non-burrito things.

However, we *can* plot your utility function $u(w)$ if we restrict its domain—that is, the allowable inputs w—to encompass only outcomes denominated in a single unit. In what follows, we'll let w denote wealth, but more generally it could be units of any particular good or service you might be interested in consuming (like burritos). The function $u(w)$ is called your utility of wealth. It encode your attitude to risk in making financial decisions. It satisfies two properties:

1. Non-satiation: $u(w)$ increases with w, i.e. $u'(w) > 0$.

2. Risk aversion: $u(w)$ increases more slowly as w gets larger. In other words, it increases at a decelerating rate, so that $u''(w) < 0$.

The principle of expected utility says that, when when faced with a choice of different investment options, you should take the option that maximizes your expected utility of wealth.

Arbitrariness of scale. The scale of a utility function is arbitrary. To see this, suppose that Alice's utility function is $u(w)$, and that Bob's is $\tilde{u}(w) = a + b \cdot u(w)$, where $b > 0$. That is, Bob's utility differs from Alice's only by a rescaling and the addition of a constant. As we'll now show, Bob's utility is actually the same as Alice's, in the sense that the two of them will both prefer exactly the same things.

Suppose that X and Y are two gambles, and let's say that Bob prefers X to Y. From this, we can deduce that $E\{\tilde{u}(X)\} > E\{\tilde{u}(Y)\}$. This can be true if and only if

$$\begin{aligned} E\{a + b \cdot u(X)\} &> E\{a + b \cdot u(Y)\} \\ a + bE\{u(X)\} &> a + bE\{u(Y)\} \\ bE\{u(X)\} &> bE\{u(Y)\} \\ E\{u(X)\} &> E\{u(Y)\}. \end{aligned}$$

But if this last line holds, then Alice prefers X to Y as well. Because this argument works for any gambles X and Y, Alice's and Bob's utility functions are operationally equivalent: Bob prefers X to Y if and only if Alice does. We can therefore rescale utility functions to make the units convenient for us to think about, without changing the underlying ordering of preferences.

Logarithmic utility. What's a "good" utility function to use? The honest answer is that, subject to these basic properties of non-satiation and risk aversion, your utility function can be anything.

Nevertheless, there are some standard choices that economists tend to assume about people, and that might plausibly describe the utility function for a typical investor plucked at random off the street. A very common example is logarithmic utility:

$$u(w) = \log w .$$

Remember, log means the natural logarithm.

Let's use the log utility function to work through a simple decision problem where you, as an investor, must decide how much of your wealth to allocate between a risky asset and a riskless asset. The riskless asset carries no risk, but no return—for example, stuffing cash under your mattress. The risky asset will return at rate r_g with probability p (a good outcome), and rate r_b with probability $1 - p$ (a bad outcome). If your beginning wealth is W_0 and if you invest some fraction c of your wealth in the risky asset, where c could be anything between 0 and 1, your possible outcomes $W_0 \cdot (1 + c r_g)$ or $W_0 \cdot (1 + c r_g)$.

First, suppose that $p = 0.5$, $r_g = 0.1$, and $r_b = -0.1$. Regardless of your current wealth W_0, you will allocate nothing to the risky asset. It has an expected return of zero (just like the riskless asset), but introduces risk. Because you are risk averse, you want no part of it. To show this mathematically, let W denote the random variable for your wealth after the asset has yielded its return. Then

$$E\{u(W)\} = 0.5 \log\{W_0 \cdot (1 + 0.1c)\} + 0.5 \log\{W_0 \cdot (1 - 0.1c)\},$$

which is maximized at $c = 0$. (To see this, just plot the function.)

Let's try a slightly more interesting example. Now suppose that your current wealth is $W_0 = 100$; that $r_g = 0.1$ and $r_b = -0.1$; and that $p = 0.52$. Then

$$E\{u(W)\} = 0.52 \log\{W_0 \cdot (1 + 0.1c)\} + 0.48 \log\{W_0 \cdot (1 - 0.1c)\},$$

Again, you could either plot this or use calculus to conclude that the optimal allocation is

$$c^\star = \left[\frac{(0.52 - 1)(-0.1) - 0.52(0.1)}{-0.1 \cdot 0.1} \right] = 0.4 \, .$$

So you should stuff \$60 under your mattress and put \$40 into the risky asset, so that you stand to gain \$4 or lose \$4. In general, you can use calculus to show that the optimal allocation is

$$c^\star = \frac{(p-1) r_b - p r_g}{r_g \cdot r_b} \, .$$

An interesting thing to note is that this is independent of your initial wealth. This property is called *constant relative risk aversion*: regardless of how wealthy you are, you always want to place the same percentage of your wealth in the risky asset.

Connection with the Kelly rule. As an aside, consider the special case where $r_b = 1$ and $r_g = -1$, so that you stand to either double your investment in the risky asset, or lose it all. In this case, the optimal allocation reduces to

$$c^\star = p - (1 - p),$$

which is exactly the same allocation as the one we derived using the Kelly rule in class. In general, Kelly-rule investing is equivalent to investing with logarithmic utility. Thus everything that applies to the Kelly rule also applies to the log utility function: although it entails some risk aversion, it is still very aggressive. In particular, "it marks the boundary between aggressive and insane investing." (This quip is from the book *Fortune's Formula* by William Poundstone.)

Isoelastic utility. Another very commonly assumed utility function is the following:

$$u(w) = \frac{w^{1-\gamma} - 1}{1 - \gamma},$$

where $\gamma > 1$ is a tunable parameter that encodes a person's risk aversion. Higher values of γ entail more risk aversion. This is called the isoelastic class of utility functions. In the limiting case as $\gamma \to 1$, the isoelastic class reduces to logarithmic utility. (You can demonstrate this using L'Hopital's rule, if you remember that from differential calculus.)

We won't talk about isoelastic utility very much, or any of the myriad other "off the shelf" utility functions out there. The point is: your utility function is yours and yours alone, but one of these functions might be a good place to start in the process of introspection necessary to know your own utilities.

2 Bootstrap resampling

Basic idea. Recall that the idea of Monte Carlo simulation is to approximate complicated probability distributions via computer simulations. The key equation is Monte Carlo simulation is the following. If X is a random variable with probability distribution $P(X)$, and we are interested in computing the expected value of some function $f(X)$, then

$$E[f(X)] \approx \frac{1}{N} \sum_{i=1}^{N} f(X^{(i)}),$$

where each $X^{(i)}$ is a simulated draw from the distribution $P(X)$. This statement says that we can approximate a population mean (the expected value on the left) with the corresponding sample mean (the average on the right). The number N is the Monte Carlo sample size, and the Monte Carlo error is the discrepancy between the left-hand side and the approximation on the right. Larger values of N will lead to smaller Monte Carlo error. We've used this fact repeatedly throughout the semester (albeit without explicitly writing it down as an equation). For example, in calculating bootstrapped sampling distributions and permutation tests, we relied upon this fact.

Joint distributions. The Monte Carlo method works for joint distributions, too. Suppose that $X_1, \dots, X_D$ are D correlated random variables with joint distribution $P(X_1, \dots, X_D)$. For example, the random variables might be the returns on a single day of all the assets in a financial portfolio. If we're interested in some nonlinear function $f(X_1, X_2, \dots, X_D)$ of these random variables, then we can use the same basic equation as above:

$$E[f(X_1, X_2, \dots, X_D)] \approx \frac{1}{N} \sum_{i=1}^{N} f(X_1^{(i)}, X_2^{(i)}, \dots, X_D^{(i)}),$$

where each set $(X_1^{(i)}, X_2^{(i)}, \dots, X_D^{(i)})$ is a single draw of all variables from their joint distribution.

A natural application of this principle is in calculating expected utilities for financial portfolios. Suppose that we're trying to decide the optimal allocation among D assets. Let $X_{j,t}$ be the random variable denoting the return of asset j during time period t, and that we're investing over a horizon from $t = 1$ to $t = T$. Let θ denote the set of returns for all assets across all time periods: $\theta = \{X_{jt} : j = 1, \dots, D \text{ and } t = 1, \dots T\}$. Finally, let $W(\theta)$ be the final wealth of our portfolio as a function of all these random variables. Naturally, if our utility of wealth is $u(w)$, then we want to choose our portfolio so that our expected utility $E(u\{W(\theta)\})$ is as large as possible.

Since θ is an extremely complicated random variable, we can use Monte Carlo simulation here. Suppose that your initial wealth is W_0, and that we distribute it so that your holdings in each asset at $W_{j,0}$. To evaluate the expected utility of a particular portfolio allocation rule, we could repeat the following simulation many times.

(1) For $t = 1, \ldots, T$:

   a. Simulate $X_{1t}, X_{2t}, \ldots, X_{Dt}$ from the joint distribution of asset returns.

   b. Update $W_{j,t}$, the value of your holdings in each asset at step t.

   c. Optionally, rebalance your portfolio to the target allocation.

(2) After T time periods, calculate final wealth W_T by summing the final holdings in each asset:

$$W_T = \sum_{j=1}^{D} W_{j,T} \, .$$

(3) Evaluate the utility of your final wealth, $u(W_T)$. Notice that implicitly, W_T is a function of θ, the asset returns $X_{j,t}$ across all periods.

At the end of many simulations, you will have a collection of Monte Carlo samples $u(W_T^{(i)})$ for the utility of your final wealth, allowing you to approximate the expected utility as

$$E[u(W_T)] \approx \frac{1}{N} \sum_{i=1}^{N} u(W_T^{(i)}) \, .$$

Sampling complicated joint distributions. Thus we see why it's important to be able to sample from complicated joint distributions. In the special case of two assets (e.g. stocks and bonds), we can imagine a bivariate normal model used in Step 1(a) of the above process.

However, this strategy breaks down easily. In general, using parametric probability models (like the bivariate normal) to describe complicated joint distributions is dicey. A joint distribution is typically very complicated mathematically. We might be oversimplifying drastically by assuming something like a bivariate normal distribution (or its generalization, called the multivariate normal).

To avoid the oversimplification of parametric models, a very practical technique is bootstrap resampling. Suppose we have M past samples of the random variables of interest, stacked in a matrix or spreadsheet:

$$X = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1D} \\ X_{21} & X_{22} & \cdots & X_{2D} \\ \vdots & & & \\ X_{M1} & X_{M2} & \cdots & X_{MD} \end{pmatrix}$$

where X_{ij} is the ith sample of the jth variable. For example, the ith row of this spreadsheet might give the returns/interest rates of D correlated assets on a single day.

The key idea of bootstrap resampling is the following. We may not be able to describe what the joint distribution $P(X_1, \ldots, X_D)$ is, but *we do know that every row of this X matrix is a sample from this joint distribution.* Therefore, instead of sampling from the joint distribution, we will sample from the sample—i.e. we will bootstrap the past data. Thus every time we need a new draw from the joint distribution $P(X_1, \ldots, X_D)$, we randomly sample (with replacement) a single row of X. This would entail a modification of the above algorithm:

(1) For $t = 1, \ldots, T$:

   a. Take an approximate sample from the joint distribution of $X_{1t}, X_{2t}, \ldots, X_{Dt}$ by resampling one set of past returns from data collected at the appropriate time scale (e.g. daily if the time period t is measured in days, yearly if the time period is years).

   b. Update $W_{j,t}$, the value of your holdings in each asset at step t.

   c. Optionally, rebalance your portfolio to the target allocation.

(2) As before.

(3) As before.