

Foundations of Reinforcement Learning with Applications in Finance

Ashwin Rao, Tikhon Jelvis

Part I.

Modeling Financial Applications

1. Utility Theory

1.1. Introduction to the Concept of Utility

This chapter marks the beginning of Module II, where we cover a set of financial applications that can be solved with Dynamic Programming or Reinforcement Learning Algorithms. A fundamental feature of many financial applications cast as Stochastic Control problems is that the *Rewards* of the modeled MDP are Utility functions in order to capture the tradeoff between financial returns and risk. So this chapter is dedicated to the topic of *Financial Utility*. We begin with developing an understanding of what *Utility* means from a broad Economic perspective, then zoom into the concept of Utility from a financial/monetary perspective, and finally show how Utility functions can be designed to capture individual preferences of “risk-taking-inclination” when it comes to specific financial applications.

[Utility Theory](#) is a vast and important topic in Economics and we won't cover it in detail in this book - rather, we will focus on the aspects of Utility Theory that are relevant for the Financial Applications we cover in this book. But it pays to have some familiarity with the general concept of Utility in Economics. The term *Utility* (in Economics) refers to the abstract concept of an individual's preferences over choices of products or services or activities (or more generally, over choices of certain abstract entities analyzed in Economics). Let's say you are offered 3 options to spend your Saturday afternoon: A) lie down on your couch and listen to music, or B) baby-sit your neighbor's kid and earn some money, or C) play a game of tennis with your friend. We really don't know how to compare these 3 options in a formal/analytical manner. But we tend to be fairly decisive (instinctively) in picking among disparate options of this type. Utility Theory aims to formalize making choices by assigning a real number to each presented choice, and then picking the choice with the highest assigned number. The assigned real number for each choice represents the “value”/“worth” of the choice, noting that the “value”/“worth” is often an implicit/instinctive value that needs to be made explicit. In our example, the numerical value for each choice is not something concrete or precise like number of dollars earned on a choice or the amount of time spent on a choice - rather it is a more abstract notion of an individual's “happiness” or “satisfaction” associated with a choice. In this example, you might say you prefer option A) because you feel lazy today (so, no tennis) and you care more about enjoying some soothing music after a long work-week than earning a few extra bucks through baby-sitting. Thus, you are comparing different attributes like money, relaxation and pleasure. This can get more complicated if your friend is offered these options, and say your friend chooses option C). If you see your friend's choice, you might then instead choose option C) because you perceive the “collective value” (for you and your friend together) to be highest if you both choose option C).

We won't go any further on this topic of abstract Utility, but we hope the above example provides the basic intuition for the broad notion of Utility in Economics as preferences over choices by assigning a numerical value for each choice. For a deeper study of Utility Theory, we refer you to [The Handbook of Utility Theory](#) (Barberà, Seidl, and Hammond 1998). In this book, we focus on the narrower notion of *Utility of Money* because money is

what we care about when it comes to financial applications. However, Utility of Money is not so straightforward because different people respond to different levels of money in different ways. Moreover, in many financial applications, Utility functions help us determine the tradeoff between financial return and risk, and this involves (challenging) assessments of the likelihood of various outcomes. The next section develops the intuition on these concepts.

1.2. A Simple Financial Example

To warm up to the concepts associated with Financial Utility Theory, let's start with a simple financial example. Consider a casino game where your financial gain/loss is based on the outcome of tossing a fair coin (HEAD or TAIL outcomes). Let's say you will be paid \$1000 if the coin shows HEAD on the toss, and let's say you would be required to pay \$500 if the coin shows TAIL on the toss. Now the question is: How much would you be willing to pay upfront to play this game? Your first instinct might be to say: "I'd pay \$250 upfront to play this game because that's my expected payoff, based on the probability of the outcomes" ($250 = 0.5(1000) + 0.5(-500)$). But after you think about it carefully, you might alter your answer to be: "I'd pay a little less than \$250." When pressed for why the fair upfront cost for playing the game should be less than \$250, you might say: "I need to be compensated for taking the risk."

What does the word "risk" mean? It refers to the degree of variation in the outcomes (\$1000 versus -\$500). But then why would you say you need to be compensated for being exposed to this variation in outcomes? If -\$500 makes you unhappy, \$1000 should make you happy, and so, shouldn't we average out the happiness to the tune of \$250? Well, not quite. Let's understand the word "happiness" (or call it "satisfaction") - this is the notion of utility of outcomes. Let's say you did pay \$250 upfront to play the game. Then the coin toss outcome of HEAD is a net gain of $\$1000 - \$250 = \$750$ and the coin toss outcome of TAIL is a net gain of $-\$500 - \$250 = -\$750$ (i.e., net loss of \$750). Now let's say the HEAD outcome gain of \$750 gives you "happiness" of say 100 units. If the TAIL outcome loss of \$750 gives you "unhappiness" of 100 units, then "happiness" and "unhappiness" levels cancel out, and in that case, it would be fair to pay \$250 upfront to play the game. But it turns out that for most people, the "happiness"/"satisfaction" levels are asymmetric. If the "happiness" for \$750 gain is 100 units, then the "unhappiness" for \$750 loss is typically more than 100 units (let's say for you it's 120 units). This means you will pay an upfront amount X (less than \$250) such that the difference in Utilities of \$1000 and X is exactly the difference in the Utilities of X and -\$500. Let's say this X amounts of \$180. The gap of \$70 ($\$250 - \180) represents your compensation for taking the risk, and it really comes down to the asymmetry in your assignment of utility to the outcomes.

Note that the degree of asymmetry of utility ("happiness" versus "unhappiness" for equal gains versus losses) is fairly individualized. Your utility assignment to outcomes might be different from your friend's. Your friend might be more asymmetric in assessing utility of the two outcomes and might assign 100 units of "happiness" for the gain outcome and 150 units of "unhappiness" for the loss outcome. So then your friend would pay an upfront amount X lower than the amount of \$180 you paid upfront to play this game. Let's say the X for your friend works out to \$100, so his compensation for taking the risk is $\$250 - \$100 = \$150$, significantly more than your \$70 of compensation for taking the same risk.

Thus we see that each individual's asymmetry in utility assignment to different out-

comes results in this psychology of “I need to be compensated for taking this risk.” We refer to this individualized demand of “compensation for risk” as the attitude of *Risk-Aversion*. It means that individuals have differing degrees of discomfort with taking risk, and they want to be compensated commensurately for taking risk. The amount of compensation they seek is called *Risk-Premium*. The more Risk-Averse an individual is, the more Risk-Premium the individual seeks. In the example above, your friend was more risk-averse than you. Your risk-premium was \$70 and your friend’s risk-premium was \$150. But the most important concept that you are learning here is that the root-cause of Risk-Aversion is the asymmetry in the assignment of utility to outcomes of opposite sign and same magnitude. We have introduced this notion of “asymmetry of utility” in a simple, intuitive manner with this example, but we will soon embark on developing the formal theory for this notion, and introduce a simple and elegant mathematical framework for Utility Functions, Risk-Aversion and Risk-Premium.

A quick note before we get into the mathematical framework - you might be thinking that a typical casino would actually charge you a bit more than \$250 upfront for playing the above game (because the casino needs to make a profit, on an expected basis), and people are indeed willing to pay this amount at a typical casino. So what about the risk-aversion we talked about earlier? The crucial point here is that people who play at casinos are looking for entertainment and excitement emanating purely from the psychological aspects of experiencing risk. They are willing to pay money for this entertainment and excitement, and this payment is separate from the cost of pure financial utility that we described above. So if people knew the true odds of pure-chance games of the type we described above and if people did not care for entertainment and excitement value of risk-taking in these games, focusing purely on financial utility, then what they’d be willing to pay upfront to play such a game will be based on the type of calculations we outlined above (meaning for the example we described, they’d typically pay less than \$250 upfront to play the game).

1.3. The Shape of the Utility function

We seek a “valuation formula” for the amount we’d pay upfront to sign-up for situations like the simple example above, where we have uncertain outcomes with varying payoffs for the outcomes. Intuitively, we see that the amount we’d pay:

- Increases as the Mean of the outcome increases
- Decreases as the Variance of the outcome (i.e., Risk) increases
- Decreases as our Personal Risk-Aversion increases

The last two properties above enable us to establish the Risk-Premium. Now let us understand the nature of Utility as a function of financial outcomes. The key is to note that Utility is a non-linear function of financial outcomes. We call this non-linear function as the Utility function - it represents the “happiness” / “satisfaction” as a function of money. You should think of the concept of Utility in terms of *Utility of Consumption* of money, i.e., what exactly do the financial gains fetch you in your life or business. This is the idea of “value” (utility) derived from consuming the financial gains (or the negative utility of requisite financial recovery from monetary losses). So now let us look at another simple example to illustrate the concept of Utility of Consumption, this time not of consumption of money, but of consumption of cookies (to make the concept vivid and intuitive). Figure 1.1 shows two curves - we refer to the lower curve as the marginal satisfaction (utility)

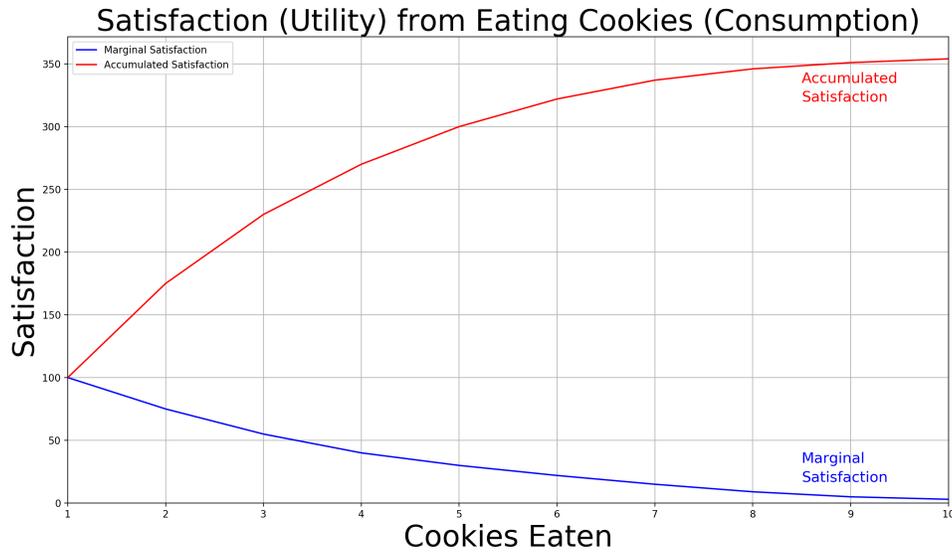


Figure 1.1.: Utility Curve

curve and the upper curve as the accumulated satisfaction (utility) curve. Marginal Utility refers to the *incremental satisfaction* we gain from an additional unit of consumption and Accumulated Utility refers to the *aggregate satisfaction* obtained from a certain number of units of consumption (in continuous-space, you can think of accumulated utility function as the integral, over consumption, of marginal utility function). In this example, we are consuming (i.e., eating) cookies. The marginal satisfaction curve tells us that the first cookie we eat provides us with 100 units of satisfaction (i.e., utility). The second cookie provides us 80 units of satisfaction, which is intuitive because you are not as hungry after eating the first cookie compared to before eating the first cookie. Also, the emotions of biting into the first cookie are extra positive because of the novelty of the experience. When you get to your 5th cookie, although you are still enjoying the cookie, you don't enjoy it as nearly as much as the first couple of cookies. The marginal satisfaction curve shows this - the 5th cookie provides us 30 units of satisfaction, and the 10th cookie provides us only 10 units of satisfaction. If we'd keep going, we might even find that the marginal satisfaction turns negative (as in, one might feel too full or maybe even feel like throwing up).

So, we see that the marginal utility function is a decreasing function. Hence, accumulated utility function is a concave function. The accumulated utility function is the Utility of Consumption function (call it U) that we've been discussing so far. Let us denote the number of cookies eaten as x , and so the total "satisfaction" (utility) after eating x cookies is referred to as $U(x)$. In our financial examples, x would be amount of money one has at one's disposal, and is typically an uncertain outcome, i.e., x is a random variable with an associated probability distribution. The extent of asymmetry in utility assignments for gains versus losses that we saw earlier manifests as extent of concavity of the $U(\cdot)$ function (which as we've discussed earlier, determines the extent of Risk-Aversion).

Now let's examine the concave nature of the Utility function for financial outcomes with another illustrative example. Let's say you have to pick between two situations:

- In Situation 1, you have a 10% probability of winning a million dollars (and 90% probability of winning 0).

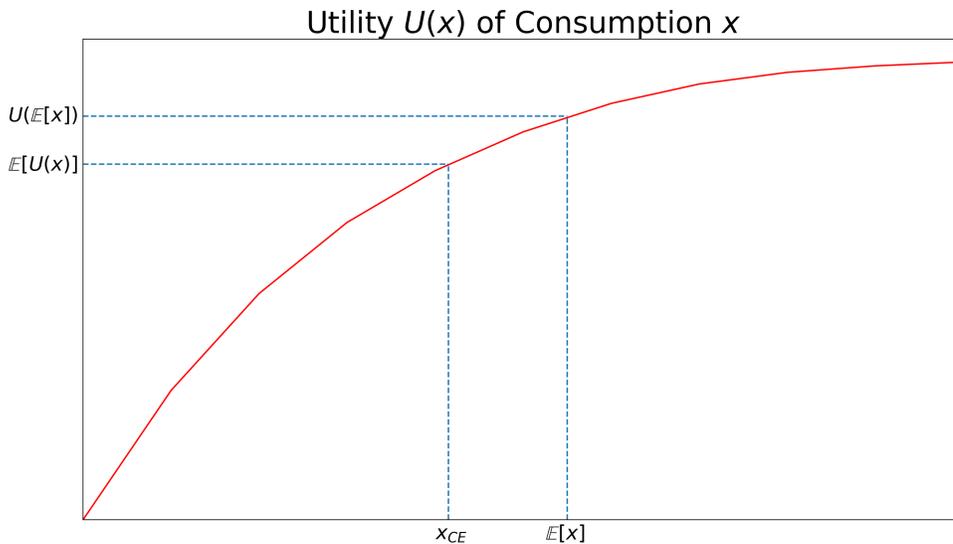


Figure 1.2.: Certainty-Equivalent Value

- In Situation 2, you have a 0.1% probability of winning a billion dollars (and 99.9% probability of winning 0).

The expected winning in Situation 1 is \$100,000 and the expected winning in Situation 2 is \$1,000,000 (i.e., 10 times more than Situation 1). If you analyzed this naively as winning expectation maximization, you'd choose Situation 2. But most people would choose Situation 1. The reason for this is that the Utility of a billion dollars is nowhere close to 1000 times the utility of a million dollars (except for some very wealthy people perhaps). In fact, the ratio of Utility of a billion dollars to Utility of a million dollars might be more like 10. So, the choice of Situation 1 over Situation 2 is usually quite clear - it's about Utility expectation maximization. So if the Utility of 0 dollars is 0 units, the Utility of a million dollars is say 1000 units, and the Utility of a billion dollars is say 10000 units (i.e., 10 times that of a million dollars), then we see that the Utility of financial gains is a fairly concave function.

1.4. Calculating the Risk-Premium

Note that the concave nature of the $U(\cdot)$ function implies that:

$$\mathbb{E}[U(x)] < U(\mathbb{E}[x])$$

We define *Certainty-Equivalent Value* x_{CE} as:

$$x_{CE} = U^{-1}(\mathbb{E}[U(x)])$$

Certainty-Equivalent Value represents the certain amount we'd pay to consume an uncertain outcome. This is the amount of \$180 you were willing to pay to play the casino game of the previous section.

Figure 1.2 illustrates this concept of Certainty-Equivalent Value in graphical terms. Next, we define Risk-Premium in two different conventions:

- **Absolute Risk-Premium** π_A :

$$\pi_A = \mathbb{E}[x] - x_{CE}$$

- **Relative Risk-Premium** π_R :

$$\pi_R = \frac{\pi_A}{\mathbb{E}[x]} = \frac{\mathbb{E}[x] - x_{CE}}{\mathbb{E}[x]} = 1 - \frac{x_{CE}}{\mathbb{E}[x]}$$

Now we develop mathematical formalism to derive formulas for Risk-Premia π_A and π_R in terms of the extent of Risk-Aversion and the extent of Risk itself. To lighten notation, we refer to $\mathbb{E}[x]$ as \bar{x} and Variance of x as σ_x^2 . We write $U(x)$ as the Taylor series expansion around \bar{x} and ignore terms beyond quadratic in the expansion, as follows:

$$U(x) \approx U(\bar{x}) + U'(\bar{x}) \cdot (x - \bar{x}) + \frac{1}{2} U''(\bar{x}) \cdot (x - \bar{x})^2$$

Taking the expectation of $U(x)$ in the above formula, we get:

$$\mathbb{E}[U(x)] \approx U(\bar{x}) + \frac{1}{2} \cdot U''(\bar{x}) \cdot \sigma_x^2$$

Next, we write the Taylor-series expansion for $U(x_{CE})$ around \bar{x} and ignore terms beyond linear in the expansion, as follows:

$$U(x_{CE}) \approx U(\bar{x}) + U'(\bar{x}) \cdot (x_{CE} - \bar{x})$$

Since $\mathbb{E}[U(x)] = U(x_{CE})$ (by definition of x_{CE}), the above two expressions are approximately the same. Hence,

$$U'(\bar{x}) \cdot (x_{CE} - \bar{x}) \approx \frac{1}{2} \cdot U''(\bar{x}) \cdot \sigma_x^2 \tag{1.1}$$

From Equation (1.1), Absolute Risk-Premium

$$\pi_A = \bar{x} - x_{CE} \approx -\frac{1}{2} \cdot \frac{U''(\bar{x})}{U'(\bar{x})} \cdot \sigma_x^2$$

We refer to the function:

$$A(x) = -\frac{U''(x)}{U'(x)}$$

as the *Absolute Risk-Aversion* function. Therefore,

$$\pi_A \approx \frac{1}{2} \cdot A(\bar{x}) \cdot \sigma_x^2$$

In multiplicative uncertainty settings, we focus on the variance $\sigma_{\frac{x}{\bar{x}}}^2$ of $\frac{x}{\bar{x}}$. So in multiplicative settings, we focus on the Relative Risk-Premium:

$$\pi_R = \frac{\pi_A}{\bar{x}} \approx -\frac{1}{2} \cdot \frac{U''(\bar{x}) \cdot \bar{x}}{U'(\bar{x})} \cdot \frac{\sigma_x^2}{\bar{x}^2} = -\frac{1}{2} \cdot \frac{U''(\bar{x}) \cdot \bar{x}}{U'(\bar{x})} \cdot \sigma_{\frac{x}{\bar{x}}}^2$$

We refer to the function

$$R(x) = -\frac{U''(x) \cdot x}{U'(x)}$$

as the *Relative Risk-Aversion* function. Therefore,

$$\pi_R \approx \frac{1}{2} \cdot R(\bar{x}) \cdot \sigma_{\frac{x}{\bar{x}}}^2$$

Now let's take stock of what we've learning here. We've shown that Risk-Premium is proportional to the product of:

- Extent of Risk-Aversion: either $A(\bar{x})$ or $R(\bar{x})$
- Extent of uncertainty of outcome (i.e., Risk): either σ_x^2 or $\sigma_{\frac{x}{\bar{x}}}^2$

We've expressed the extent of Risk-Aversion to be proportional to the negative ratio of:

- Concavity of the Utility function (at \bar{x}): $-U''(\bar{x})$
- Slope of the Utility function (at \bar{x}): $U'(\bar{x})$

So for typical optimization problems in financial applications, we maximize $\mathbb{E}[U(x)]$ (not $\mathbb{E}[x]$), which in turn amounts to maximization of $x_{CE} = \mathbb{E}[x] - \pi_A$. If we refer to $\mathbb{E}[x]$ as our "Expected Return on Investment" (or simply "Return" for short) and π_A as the "risk-adjustment" due to risk-aversion and uncertainty of outcomes, then x_{CE} can be conceptualized as "risk-adjusted-return." Thus, in financial applications, we seek to maximize risk-adjusted-return x_{CE} rather than just the return $\mathbb{E}[x]$. It pays to emphasize here that the idea of maximizing risk-adjusted-return is essentially the idea of maximizing expected utility, and that the utility function is a representation of an individual's risk-aversion.

Note that Linear Utility function $U(x) = a + bx$ implies *Risk-Neutrality* (i.e., when one doesn't demand any compensation for taking risk). Next, we look at typically-used Utility functions $U(\cdot)$ with:

- Constant Absolute Risk-Aversion (CARA)
- Constant Relative Risk-Aversion (CRRA)

1.5. Constant Absolute Risk-Aversion (CARA)

Consider the Utility function $U : \mathbb{R} \rightarrow \mathbb{R}$, parameterized by $a \in \mathbb{R}$, defined as:

$$U(x) = \begin{cases} \frac{1-e^{-ax}}{a} & \text{for } a \neq 0 \\ x & \text{for } a = 0 \end{cases}$$

Firstly, note that $U(x)$ is continuous with respect to a for all $x \in \mathbb{R}$ since:

$$\lim_{a \rightarrow 0} \frac{1 - e^{-ax}}{a} = x$$

Now let us analyze the function $U(\cdot)$ for any fixed a . We note that for all $a \in \mathbb{R}$:

- $U(0) = 0$
- $U'(x) = e^{-ax} > 0$ for all $x \in \mathbb{R}$
- $U''(x) = -a \cdot e^{-ax}$

This means $U(\cdot)$ is a monotonically increasing function passing through the origin, and its curvature has the opposite sign as that of a (note: no curvature when $a = 0$).

So now we can calculate the Absolute Risk-Aversion function:

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

So we see that the Absolute Risk-Aversion function is the constant value a . Consequently, we say that this Utility function corresponds to *Constant Absolute Risk-Aversion* (CARA). The parameter a is referred to as the Coefficient of CARA. The magnitude of positive a signifies the degree of risk-aversion. $a = 0$ is the case of being Risk-Neutral. Negative values of a mean one is "risk-seeking," i.e., one will pay to take risk (the opposite of risk-aversion) and the magnitude of negative a signifies the degree of risk-seeking.

If the random outcome $x \sim \mathcal{N}(\mu, \sigma^2)$, then using Equation (??) from Appendix ??, we get:

$$\mathbb{E}[U(x)] = \begin{cases} \frac{1 - e^{-a\mu + \frac{a^2\sigma^2}{2}}}{a} & \text{for } a \neq 0 \\ \mu & \text{for } a = 0 \end{cases}$$

$$x_{CE} = \mu - \frac{a\sigma^2}{2}$$

$$\text{Absolute Risk Premium } \pi_A = \mu - x_{CE} = \frac{a\sigma^2}{2}$$

For optimization problems where we need to choose across probability distributions where σ^2 is a function of μ , we seek the distribution that maximizes $x_{CE} = \mu - \frac{a\sigma^2}{2}$. This clearly illustrates the concept of "risk-adjusted-return" because μ serves as the "return" and the risk-adjustment $\frac{a\sigma^2}{2}$ is proportional to the product of risk-aversion a and risk (i.e., variance in outcomes) σ^2 .

1.6. A Portfolio Application of CARA

Let's say we are given \$1 to invest and hold for a horizon of 1 year. Let's say our portfolio investment choices are:

- A risky asset with Annual Return $\sim \mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}, \sigma \in \mathbb{R}^+$.
- A riskless asset with Annual Return $= r \in \mathbb{R}$.

Our task is to determine the allocation π (out of the given \$1) to invest in the risky asset (so, $1 - \pi$ is invested in the riskless asset) so as to maximize the Expected Utility of Consumption of Portfolio Wealth in 1 year. Note that we allow π to be unconstrained, i.e., π can be any real number from $-\infty$ to $+\infty$. So, if $\pi > 0$, we buy the risky asset and if $\pi < 0$, we "short-sell" the risky asset. Investing π in the risky asset means in 1 year, the risky asset's value will be a normal distribution $\mathcal{N}(\pi(1 + \mu), \pi^2\sigma^2)$. Likewise, if $1 - \pi > 0$, we lend $1 - \pi$ (and will be paid back $(1 - \pi)(1 + r)$ in 1 year), and if $1 - \pi < 0$, we borrow $1 - \pi$ (and need to pay back $(1 - \pi)(1 + r)$ in 1 year).

Portfolio Wealth W in 1 year is given by:

$$W \sim \mathcal{N}(1 + r + \pi(\mu - r), \pi^2\sigma^2)$$

We assume CARA Utility with $a \neq 0$, so:

$$U(W) = \frac{1 - e^{-aW}}{a}$$

We know that maximizing $\mathbb{E}[U(W)]$ is equivalent to maximizing the Certainty-Equivalent Value of Wealth W , which in this case (using the formula for x_{CE} in the section on CARA) is given by:

$$1 + r + \pi(\mu - r) - \frac{a\pi^2\sigma^2}{2}$$

This is a quadratic concave function of π for $a > 0$, and so, taking it's derivative with respect to π and setting it to 0 gives us the optimal investment fraction in the risky asset (π^*) as follows:

$$\pi^* = \frac{\mu - r}{a\sigma^2}$$

1.7. Constant Relative Risk-Aversion (CRRA)

Consider the Utility function $U : \mathbb{R}^+ \rightarrow \mathbb{R}$, parameterized by $\gamma \in \mathbb{R}$, defined as:

$$U(x) = \begin{cases} \frac{x^{1-\gamma}-1}{1-\gamma} & \text{for } \gamma \neq 1 \\ \log(x) & \text{for } \gamma = 1 \end{cases}$$

Firstly, note that $U(x)$ is continuous with respect to γ for all $x \in \mathbb{R}^+$ since:

$$\lim_{\gamma \rightarrow 1} \frac{x^{1-\gamma} - 1}{1 - \gamma} = \log(x)$$

Now let us analyze the function $U(\cdot)$ for any fixed γ . We note that for all $\gamma \in \mathbb{R}$:

- $U(1) = 0$
- $U'(x) = x^{-\gamma} > 0$ for all $x \in \mathbb{R}^+$
- $U''(x) = -\gamma \cdot x^{-1-\gamma}$

This means $U(\cdot)$ is a monotonically increasing function passing through $(1, 0)$, and it's curvature has the opposite sign as that of γ (note: no curvature when $\gamma = 0$).

So now we can calculate the Relative Risk-Aversion function:

$$R(x) = \frac{-U''(x) \cdot x}{U'(x)} = \gamma$$

So we see that the Relative Risk-Aversion function is the constant value γ . Consequently, we say that this Utility function corresponds to *Constant Relative Risk-Aversion (CRRA)*. The parameter γ is referred to as the Coefficient of CRRA. The magnitude of positive γ signifies the degree of risk-aversion. $\gamma = 0$ yields the Utility function $U(x) = x - 1$ and is the case of being Risk-Neutral. Negative values of γ mean one is "risk-seeking," i.e., one will pay to take risk (the opposite of risk-aversion) and the magnitude of negative γ signifies the degree of risk-seeking.

If the random outcome x is lognormal, with $\log(x) \sim \mathcal{N}(\mu, \sigma^2)$, then making a substitution $y = \log(x)$, expressing $\mathbb{E}[U(x)]$ as $\mathbb{E}[U(e^y)]$, and using Equation (??) in Appendix ??, we get:

$$\mathbb{E}[U(x)] = \begin{cases} \frac{e^{\mu(1-\gamma) + \frac{\sigma^2}{2}(1-\gamma)^2} - 1}{1-\gamma} & \text{for } \gamma \neq 1 \\ \mu & \text{for } \gamma = 1 \end{cases}$$

$$x_{CE} = e^{\mu + \frac{\sigma^2}{2}(1-\gamma)}$$

$$\text{Relative Risk Premium } \pi_R = 1 - \frac{x_{CE}}{\bar{x}} = 1 - e^{-\frac{\sigma^2 \gamma}{2}}$$

For optimization problems where we need to choose across probability distributions where σ^2 is a function of μ , we seek the distribution that maximizes $\log(x_{CE}) = \mu + \frac{\sigma^2}{2}(1-\gamma)$. Just like in the case of CARA, this clearly illustrates the concept of “risk-adjusted-return” because $\mu + \frac{\sigma^2}{2}$ serves as the “return” and the risk-adjustment $\frac{\gamma\sigma^2}{2}$ is proportional to the product of risk-aversion γ and risk (i.e., variance in outcomes) σ^2 .

1.8. A Portfolio Application of CRRA

This application of CRRA is a special case of [Merton’s Portfolio Problem](#) (Merton 1969) that we shall cover in its full generality in Chapter ???. This section requires us to have some basic familiarity with Stochastic Calculus (covered in Appendix ??), specifically Ito Processes and Ito’s Lemma. Here we consider the single-decision version of Merton’s Portfolio Problem where our portfolio investment choices are:

- A risky asset, evolving in continuous time, with value denoted S_t at time t , whose movements are defined by the Ito process:

$$dS_t = \mu \cdot S_t \cdot dt + \sigma \cdot S_t \cdot dz_t$$

where $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$ are given constants, and z_t is 1-dimensional standard Brownian Motion.

- A riskless asset, growing continuously in time, with value denoted R_t at time t , whose growth is defined by the ordinary differential equation:

$$dR_t = r \cdot R_t \cdot dt$$

where $r \in \mathbb{R}$ is a given constant.

We are given \$1 to invest over a period of 1 year. We are asked to maintain a constant fraction of investment of wealth (denoted $\pi \in \mathbb{R}$) in the risky asset at each time t (with $1 - \pi$ as the fraction of investment in the riskless asset at each time t). Note that to maintain a constant fraction of investment in the risky asset, we need to continuously rebalance the portfolio of the risky asset and riskless asset. Our task is to determine the constant π that maximizes the Expected Utility of Consumption of Wealth at the end of 1 year. We allow π to be unconstrained, i.e., π can take any value from $-\infty$ to $+\infty$. Positive π means we have a “long” position in the risky asset and negative π means we have a “short” position in the risky asset. Likewise, positive $1 - \pi$ means we are lending money at the riskless interest rate of r and negative $1 - \pi$ means we are borrowing money at the riskless interest rate of r .

We denote the Wealth at time t as W_t . Without loss of generality, assume $W_0 = 1$. Since W_t is the portfolio wealth at time t , the value of the investment in the risky asset at time t would need to be $\pi \cdot W_t$ and the value of the investment in the riskless asset at time t would need to be $(1 - \pi) \cdot W_t$. Therefore, the change in the value of the risky asset investment from time t to time $t + dt$ is:

$$\mu \cdot \pi \cdot W_t \cdot dt + \sigma \cdot \pi \cdot W_t \cdot dz_t$$

Likewise, the change in the value of the riskless asset investment from time t to time $t + dt$ is:

$$r \cdot (1 - \pi) \cdot W_t \cdot dt$$

Therefore, the infinitesimal change in portfolio wealth dW_t from time t to time $t + dt$ is given by:

$$dW_t = (r + \pi(\mu - r)) \cdot W_t \cdot dt + \pi \cdot \sigma \cdot W_t \cdot dz_t$$

Note that this is an Ito process defining the stochastic evolution of portfolio wealth. Applying Ito's Lemma (see Appendix ??) on $\log W_t$ gives us:

$$\begin{aligned} d(\log W_t) &= ((r + \pi(\mu - r)) \cdot W_t \cdot \frac{1}{W_t} - \frac{\pi^2 \cdot \sigma^2 \cdot W_t^2}{2} \cdot \frac{1}{W_t^2}) \cdot dt + \pi \cdot \sigma \cdot W_t \cdot \frac{1}{W_t} \cdot dz_t \\ &= (r + \pi(\mu - r) - \frac{\pi^2 \sigma^2}{2}) \cdot dt + \pi \cdot \sigma \cdot dz_t \end{aligned}$$

Therefore,

$$\log W_t = \int_0^t (r + \pi(\mu - r) - \frac{\pi^2 \sigma^2}{2}) \cdot du + \int_0^t \pi \cdot \sigma \cdot dz_u$$

Using the martingale property and Ito Isometry for the Ito integral $\int_0^t \pi \cdot \sigma \cdot dz_u$ (see Appendix ??), we get:

$$\log W_1 \sim \mathcal{N}(r + \pi(\mu - r) - \frac{\pi^2 \sigma^2}{2}, \pi^2 \sigma^2)$$

We assume CRRA Utility with $\gamma \neq 0$, so:

$$U(W_1) = \begin{cases} \frac{W_1^{1-\gamma} - 1}{1-\gamma} & \text{for } \gamma \neq 1 \\ \log(W_1) & \text{for } \gamma = 1 \end{cases}$$

We know that maximizing $\mathbb{E}[U(W_1)]$ is equivalent to maximizing the Certainty-Equivalent Value of W_1 , hence also equivalent to maximizing the log of the Certainty-Equivalent Value of W_1 , which in this case (using the formula for x_{CE} from the section on CRRA) is given by:

$$\begin{aligned} r + \pi(\mu - r) - \frac{\pi^2 \sigma^2}{2} + \frac{\pi^2 \sigma^2 (1 - \gamma)}{2} \\ = r + \pi(\mu - r) - \frac{\pi^2 \sigma^2 \gamma}{2} \end{aligned}$$

This is a quadratic concave function of π for $\gamma > 0$, and so, taking it's derivative with respect to π and setting it to 0 gives us the optimal investment fraction in the risky asset (π^*) as follows:

$$\pi^* = \frac{\mu - r}{\gamma \sigma^2}$$

1.9. Key Takeaways from this Chapter

- An individual's financial risk-aversion is represented by the concave nature of the individual's Utility as a function of financial outcomes.
- Risk-Premium (compensation an individual seeks for taking financial risk) is roughly proportional to the individual's financial risk-aversion and the measure of uncertainty in financial outcomes.
- Risk-Adjusted-Return in finance should be thought of as the Certainty-Equivalent-Value, whose Utility is the Expected Utility across uncertain (risky) financial outcomes.

Bibliography

- Barberà, Salvador, Christian Seidl, and Peter J. Hammond, eds. 1998. *Handbook of Utility Theory: Handbook of Utility Theory / Barberà, Salvador*. - Boston, Mass. [U.a.] : Kluwer, 1998- Vol. 1. Dordrecht [u.a.]: Kluwer. http://gso.gbv.de/DB=2.1/CMD?ACT=SRCHA&SRT=YOP&IKT=1016&TRM=ppn+266386229&sourceid=fwb_bibsonomy.
- Merton, Robert C. 1969. "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case." *The Review of Economics and Statistics* 51 (3): 247–57. <https://doi.org/10.2307/1926560>.