5. Notes on Risky Choice

Much of the material is covered in Varian Chapter 11. The Appendix below expands on some technical points.

1. Risk vs uncertainty.

   The opportunities considered so far ("bundles" of consumption goods) are riskless in the sense that you get to consume exactly what you choose. But some choices, especially in finance, are not like that. Your choices affect what you can consume, but you can't know exactly what you will get.

   In interesting situations you don't even know all the possible consequences of your choices, or the probabilities of some of the consequences; early 20th century economist Frank Knight referred to this as **uncertainty**. We will focus on the more tractable situations of what Knight called **risk**, where you know all possible consequences and their probabilities of each choice.

   Beyond the scope of this course lies research in the middle ground between Knightian uncertainty and risk. This includes Leonard Savage's work in the mid-20th century on subjective probabilities, recent behavioral work on ambiguity, and recent speculative forays on known versus unknown unknowns.

2. Risk as Lottery choice.

   For the rest of this course we will focus on cases where probabilities can be found for each possible outcome. The next unit shows how to use Bayes Theorem to compute probabilities from new information and previous beliefs. For now, we will assume that the probability $p_i \in [0, 1]$ is known for each possible outcome m_i.

   It is convenient to summarize each possible outcome as an increase (or possibly decrease) in purchasing power, relying on the standard theory of the previous unit that describes which bundles a person would choose given a total budget m and the corresponding level of satisfaction u. So here we will assume that each possible outcome m_i

is a real number, representing either total purchasing power (wealth) or, more often, an increase or decrease in purchasing power.

Thus we will regard choice as among a set of risky opportunities, each of which is described by possible monetary outcomes $m_i \in R$ with associated probabilities p_i. In some applications we take the range of outcomes to be an interval, in which case probabilities are given by a density function. For now, however, we assume that there are only a finite number of possible outcomes $i = 1, ..., k$ from any particular opportunity, so $\sum_{i=1}^{k} p_i = 1$. Economists refer to such well-defined risky opportunities as **lotteries**.

3. Modeling risky choice: Bernoulli rediscovered.

   How should a person choose among risky opportunities, i.e., among lotteries? The standard theory goes back to Daniel Bernoulli (1738); in the 1950s it was dusted off and made mainstream by researchers including John von Neumann, Leonard Savage, Ken Arrow, Milton Friedman and Harry Markowitz.

   The basic idea is that the decision maker has a utility function defined over all possible monetary outcomes, and that she rationally will choose the opportunity that maximizes utility on average. To explain, it is helpful to review some statistical and economic ideas.

   - Let L denote some lottery with outcomes $m_1, ..., m_k$ and associated probabilities $p_1, ..., p_k$. Then its **expected value** is $EL = \sum_{i=1}^{k} p_i m_i$, and its **variance** is $Var[L] = E(m - EL)^2 = \sum_{i=1}^{k} p_i (m_i - EL)^2$.
   - The expected value (sometimes denoted μ_L instead of EL) should be interpreted as the average monetary outcome, weighted by probability. Variance (sometimes denoted at σ_L^2 instead of $Var[L]$) is a measure of dispersion of monetary outcomes. Variance (or its square root, σ_L, called the **standard deviation**) is commonly used to quantify the riskiness of a lottery.
   - Suppose that the decision maker has the utility function $u(m)$, where u is a

differentiable increasing function, so $u' > 0$. Then the **expected utility** of lottery L is $E_L u = \sum_{i=1}^{k} p_i u(m_i)$.

- Suppose that the decision maker has to choose among a set of lotteries. According to **expected utility theory**, she will choose so as to maximize expected utility.

Here is a very simple example. The decision maker has utility function $u(m) = m^{0.6}$ defined for $m \geq 0$. She must choose either lottery R or lottery S. Lottery R pays \$10 if a fair coin comes up heads and \$0 otherwise. Thus $ER = 10 \cdot 0.5 + 0 \cdot 0.5 = \$5.00$ and $Var[R] = E(m - ER)^2 = (10-5)^2 \cdot 0.5 + (0-5)^2 \cdot 0.5 = 25$, so $\sigma_R = \sqrt{25} = \$5.00$. Lottery S pays \$4.00 for sure. Hence $ES = 4 \cdot 1.0 + 0 \cdot 0 = \$4.00$ and $Var[R] = E(m - ER)^2 = (4-4)^2 \cdot 1.0 + (0-4)^2 \cdot 0.0 = 0$, so $\sigma_S = \sqrt{0} = \$0.00$.

The decision maker's expected utility for lottery R is

$$E_R u = \sum_{i=1}^{k} p_i u(m_i) = p_1 u(m_1) + p_2 u(m_2) = .5 \cdot 10^{0.6} + 5 \cdot 0^{0.6} \approx .5 \cdot 3.98 = 1.99,$$

while $E_S u = 1.0 \cdot 4^{0.6} \approx 2.30$. Since $2.30 > 1.99$, expected utility theory predicts that she will pick S and not R. On the other hand (check this yourself) if another decision maker had utility function $w(m) = m^{0.9}$ and faced the same opportunities, then he would make the opposite choice.

4. Cardinal utility.

Notice that the last utility function w can be obtained from the previous utility function by means of a monotone increasing transformation: $w(m) = h(u(m)$ where $h(x) = x^{1.5}$. The theory used in previous section said that preferences (over bundles) are ordinal and thus not changed by applying such a transformation. What gives?

For several decades, economists criticized expected utility theory because it is not ordinal, but they eventually got over it. When it comes to evaluating lotteries, it is not enough to say that, for example, \$10 is better than \$4 which is better than \$0. Ordinal is not enough; you really have to know *how much* better. On the other hand, it doesn't affect comparisons of expected utility if you add or subtract a fixed amount from every monetary outcome, or if you change the scale (say from dollars to cents, or to a foreign currency at fixed exchange rates).

The technical jargon is that risk preferences are **cardinal**, not ordinal, and that the utility functions we use here are unique only up to a positive affine transformation. It is straightforward to check that if lottery L1 has higher expected utility than L2 using utility function u, then it still has higher expected utility using utility function $v(m) = \alpha u(m) + b$ as long as $\alpha > 0$.

5. Useful utility functions.

   (a) The function family $f(m|r) = m^{1-r}/(1-r)$ is called **CRRA with parameter** r. We just saw two special cases, for $r = .4$ and $r = .9$. (Actually, we dropped the constant factor $1/(1-r)$, but that doesn't matter, as explained in the previous paragraph). This family is often used in macroeconomics, where the parameter r is chosen so as to align the model with data.

   Using L'Hospital's rule, you can show that in the limiting case $r \to 1$, the function f takes the form $f(m|1) = \ln(m)$, the function that Daniel Bernoulli originally proposed in 1738 !

   (b) The function family $u(m|a) = 1 - e^{-am}$ is called **CARA with parameter** $a > 0$. It is sometimes used in applied work, where the parameter a is fitted to the data. A higher value of a is interpreted as greater risk aversion, or more cautious preferences, as we will see next.

6. Measuring Risk Aversion.

   Given a twice continuously differentiable utility function u, the *coefficient of absolute risk aversion* at monetary outcome $m \in (-\infty, \infty)$ is

$$A(m) = \frac{-u''(m)}{u'(m)}. \tag{1}$$

   It is straightforward to verify that $A(m) = a$ in the CARA function family. That is why it has its name: CARA is an acronym for constant absolute risk aversion.

The *coefficient of relative risk aversion* at $m > 0$ is

$$R(m) = \frac{-u''(m)}{u'(m)} m = mA(m). \tag{2}$$

It will not surprise you to hear (but check this anyway) that $R(m)$ is constant for all functions in the CRRA family, including $\ln m$. Can you now decipher the CRRA acronym?

You should also check that the functions $A(m)$ and $R(m)$ are unaffected by positive affine transformations, that is, they are the same for $v(m) = \alpha u(m) + b$ as they are for u if $\alpha > 0$. So these are valid cardinal measures of risk preferences.

7. Intuition about A and R.

   Suppose that $w(m) = \alpha m + b$ is linear and upward sloping, $\alpha > 0$. Then w represents the same cardinal preferences as $u(m) = m$. Of course, expected utility for u is exactly the same thing as expected value, so the same is true for w. Thus according to expected utility theory, a person with increasing linear utility will always choose the lottery with highest expected value, irrespective of variance. Such a person is said to be **risk-neutral**.

   A risk-neutral person has $u''(m) = 0$, and so by equations (1) and (2), such a person has $A(m) = R(m) = 0$.

   Higher values of A and R indicate greater aversion to risk. Why? Look at equations (1) and (2) again. They are sign-adjusted measures of concavity, appropriately normalized, and greater concavity means less willingness to accept risk. To spell this out,

   - recall that $u'' < 0$ for a concave function, so the numerators are $-u'' > 0$.
   - The denominators u' normalize u'' so that positive affine transformations have no effect, as you checked with u and v in the previous item.
   - Greater normalized curvature implies a larger drop in utility for a risky lottery relative to a non-risky lottery with the same expected value.

The last point is illustrated in Figure 1. There would be no gap between $u(EL)$ and $E_L u$ if the utility function were linear, and a very large gap if the utility function were tightly curved in the neighborhood of EL. The next item spells out the connection between risk and utility.

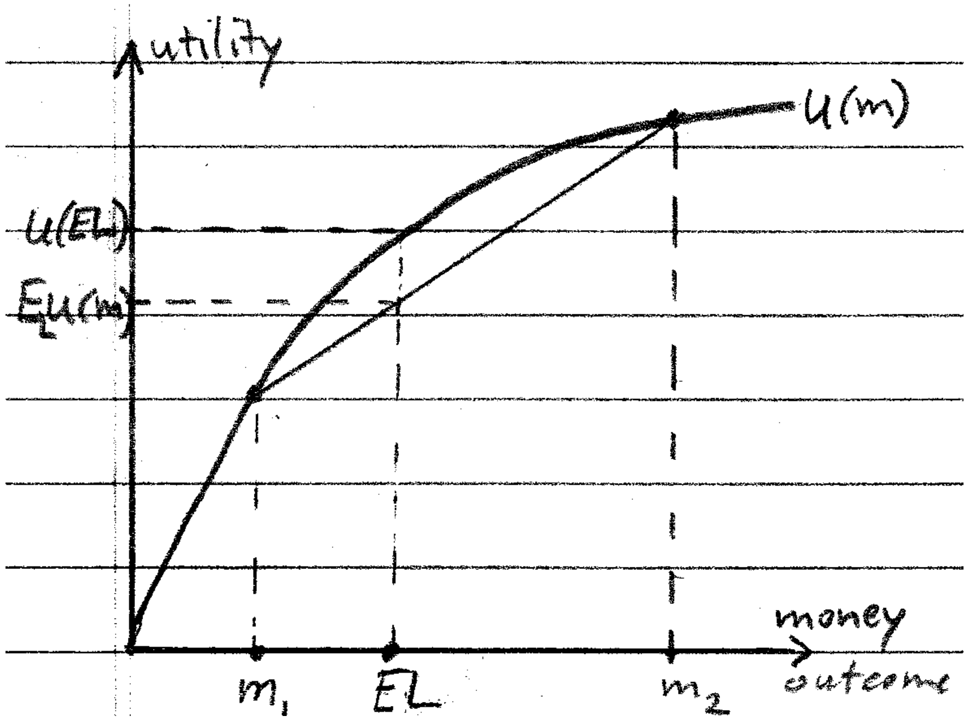

Figure 1: Expected Utility. Here the lottery prizes are $m_1 = 20$ and $m_2 = 90$ with probabilities $p_1 = .3$ and $p_2 = .7$; the expected value is $EL = 69$. The expected utility $E_L u = p_1 u(m_1) + p_2 u(m_2)$ is the height of the point above EL on the line segment connecting the points $(m_1, u(m_1))$ and $(m_2, u(m_2))$ on the utility curve. The gap between $u(EL)$ on the u-curve and the expected utility $E_L u$ on the line segment is larger the greater the degree of concavity (curvature) in the utility function u.

8. Mean and Variance.

By Taylor's Theorem, any smooth (continuously differentiable 3 times) utility function u can be expanded at any point z in its domain as a quadratic function plus remainder:

$$u(z+h) = u(z) + u'(z)h + \frac{1}{2}u''(z)h^2 + R^3(z,h), \tag{3}$$

where $R^3(z,h) = \frac{1}{6}u'''(y)h^3$ for some point y between z and $z+h$.

In equation (3), set $z = EL$ and $m = z + h$, and take the expected value of both sides. The second (linear) term disappears because $Eh = E(m - EL) = EL - EL = 0$. Since $Eh^2 = Var[L]$ the equation becomes

$$E_L u(m) = u(EL) + \frac{1}{2}u''(EL)Var[L] + R^3. \tag{4}$$

That is, expected utility of the lottery is equal to the utility of the mean outcome, plus a term proportional to the variance of the lottery and to the second derivative of u evaluated at the mean of the lottery, plus a remainder term.

- The second term in equation (4) is key. It says that variance reduces expected utility to the extent that u is concave, as measured by (unnormalized) $A(m)$. Otherwise put, a person with higher A will be more averse to variance than another rational person with A closer to zero. If $A(m) = 0$ for all m, then the Bernoulli function is linear and that person is risk neutral.

- We can ignore the remainder term if either

  (a) h is small because all likely outcomes m are near EL, or

  (b) $u'''(EL)$ is small because u'' is almost constant in the neighborhood of EL.

  In other words, the approximation using the first two terms is reliable if either the risks are small, or the utility function is nearly quadratic. A separate argument shows that the approximation is exact if the probability distribution is Normal (aka Gaussian).

- Using a suitable affine transformation and ignoring the remainder term, expected utility takes the form $EL - cVar[L]$, or in other notation $\mu_L - c\sigma_L^2$, where c is proportional to the coefficient of absolute risk aversion.

- A lot of the finance literature assumes directly that utility takes this mean-variance form. In many important cases it is a good approximation, but it can be misleading for highly risky non-Normal lotteries. See homework problem 3.2.

9. Justifications for EUT.

   There are two reasons to think that EUT should work, i.e., that a person would have some particular utility function $u(m)$ and, when faced with a set of risky opportunities, would choose one with highest expected utility.

   (a) EUT says that people choose what, on average, will make them happiest. This seems reasonable, and was Bernoulli's original justification, back in 1738.

   (b) If you write down a general preference relation over lotteries (instead of ordinary bundles), the same basic properties we used in the last unit ensure that it can be represented by a utility function U defined on the space of lotteries. This will not necessarily take the form $U(L) = E_L u$ for some utility function u defined on real numbers (monetary outcomes). But a beautiful theorem (developed in the 1950s by the economists mentioned earlier) shows that it does take that form if a few additional, very reasonable, properties are assumed.

       Thus a second reason for EUT is that it is the logical consequence of some very sensible axioms (like continuity, transitivity, and other consistency requirements) about how people make risky choices.

   This Expected Utility Theorem is explained in the Appendix below. (The function u is called a Bernoulli function there to avoid confusing it with the utility function U.) See also Varian pp. 173-176, and many other references.

10. Some caveats

    - Unless you are trained to do so, it seems unnatural to actually calculate expected utilities when making a risky choice decision.

    - EUT should be understood to say that, if a person is rational and consistent in the face of risk, her choices can be nicely summarized by finding a utility function such that her actual choices maximize the expectation of that function over available opportunities.

- Unfortunately, lots of empirical research, in the lab and in the field, has not uncovered utility functions that enable EUT to make good predictions of actual risky choices by most people.
- Behavioral economists have used Prospect Theory (Kahneman and Tversky, 1979) to account for many discrepancies. PT generalizes EUT by (a) categorizing m's as either gains or losses relative to some reference point, (b) allowing the utility function u to be concave in gains, convex in losses, and kinked at the reference point, and (c) allowing for distorted probabilities, especially overweighting the probability of rare large gains or losses.
- A recent book, *Risky Curves* by D. Friedman et al. (Routledge, 2013), argues that PT has so many free parameters that, although it can account well for many data sets after the fact, it shows little or no improvement relative to EUT in predicting what people will do in new settings. The book argues that, as far as prediction is concerned, sophisticated versions of risk-neutral expected utility (taking into account indirect effects) may offer the best predictions currently available.

Bottom line. In your instructor's opinion, EUT is good normative theory (i.e., in saying what you *should* do), and is an essential building block of modern finance. But you should be cautious in using it to predict what ordinary people will do.

Appendix: Some formalities.

Basic definitions. A *lottery* $L = (M, P)$ is a finite list of monetary outcomes $M = \{m_1, m_2, ..., m_k\} \subset \Re$ together with a corresponding list of probabilities $P = \{p_1, p_2, ..., p_k\}$, where $p_i \geq 0$ and $\sum_{i=1}^{k} p_i = 1$. The symbol $\Re$ denotes the real numbers $(-\infty, \infty)$. The *space of all lotteries* is denoted $\mathcal{L}$.

The *expected value* of lottery $L = (M, P)$ is $EL = \sum_{i=1}^{k} p_i m_i$.

A utility function over monetary outcomes (henceforth called a *Bernoulli function*) is a strictly increasing function $u : \Re \to \Re$. A *utility function over lotteries* is a function $U : \mathcal{L} \to \Re$.

Given a Bernoulli function u, the *expected utility* of lottery $L = (M, P)$ is $E_L u = \sum_{i=1}^{k} p_i u(m_i)$.

Preferences $\succeq$ over any set refer to a complete and transitive binary relation. A utility function U *represents* preferences $\succeq$ if $x \succeq y \iff U(x) \geq U(y)$ for all x, y in that set, where the symbol " $\iff$ " means "if and only if." In particular, a utility function $U : \mathcal{L} \to \Re$ represents preferences $\succeq$ over $\mathcal{L}$ if

$$L \succeq L' \iff U(L) \geq U(L')$$

for all $L, L' \in \mathcal{L}$.

Expected utility theorem. Preferences $\succeq$ over $\mathcal{L}$ have the *expected utility property* if they can be represented by a utility function U that is the expected value of some Bernoulli function u. That is, there is some Bernoulli function u, such that for all $L, L' \in \mathcal{L}$, we have $L \equiv (M, P) \succeq (M', P') \equiv L' \iff$

$$U(L) \equiv E_L u \equiv \sum_{i=1}^{k} p_i u(m_i) \geq \sum_{i=1}^{k} p_i' u(m_i') \equiv E_{L'} u \equiv U(L'). \tag{5}$$

It might seem that preferences with the expected utility property are quite special, and indeed they are. For example, their indifference surfaces are parallel and flat. Thus preferences over for lotteries with only $k = 3$ monetary outcomes have indifference curves on $\mathcal{L}$

(represented as the probability simplex, here a triangle) that are all straight lines with the same slope.

The expected utility theorem (EUT) is therefore surprising. It states that preferences over lotteries that satisfy a seemingly mild set of conditions will automatically satisfy the expected utility property, and thus be representable via a Bernoulli function.

Over the decades since the original results of Von Neumann and Morgenstern, many different sets of conditions have been shown to be sufficient. Here we mention the set used in a leading textbook, Mas-Colell et al. [2010]. It consists of four axioms that an individual's preferences $\succeq$ on $\mathcal{L}$ should satisfy:

1. Rationality: Preferences $\succeq$ are complete and transitive on $\mathcal{L}$.

2. Continuity: The precise mathematical expressions are rather indirect (they state that certain subsets of real numbers are closed sets), but they capture the intuitive idea that U doesn't take jumps on the space of lotteries. This axiom rules out lexicographic preferences.

3. Reduction of Compound Lotteries: Compound lotteries have outcomes that are themselves lotteries in $\mathcal{L}$. By taking the expected value, one obtains the *reduced* lottery, a simple lottery in $\mathcal{L}$. The axiom states that the person is indifferent between any compound lottery and the corresponding reduced lottery.

4. Independence: Let $L, L', L'' \in \mathcal{L}$ and $\alpha \in (0, 1)$. Suppose that $L \succeq L'$. Then $\alpha L + (1-\alpha)L'' \succeq \alpha L' + (1-\alpha)L''$. "In other words, if we mix two lotteries with a third one, then the preference ordering of the resulting two mixtures does not depend upon (is independent of) the particular third lottery used." (Mas-Colell et al. p. 171).

Theorem 1 (EUT)**.** *Let preferences $\succeq$ on $\mathcal{L}$ satisfy axioms 1-4 above. Then $\succeq$ has the expected utility property, i.e., there is a Bernoulli function u such that (5) holds.*

As Mas-Colell et al. point out, all four axioms seem innocuous. Someone who cares only about the ultimate monetary payoffs and whose calculations are not affected by indirect ways of stating the probabilities will satisfy the third axiom. For example, such a person would be indifferent between the compound lottery "get 0 with probability 0.5 and with probability 0.5 play the lottery that pays 10 with independent probability 0.5 and 0 otherwise," and the reduced lottery "get 0 with probability 0.75 and 0 with probability 0.25." The fourth axiom enforces a degree of consistency by requiring that preference rankings over lotteries are not changed by nesting each of those lotteries within a generic compound lottery. The first two axioms are even less controversial or problematic.[1]

For a proof of the EUT, see Mas-Colell et al. and the references cited therein. Here is a sketch of how the function u can be constructed for given preferences. Denote by m_+ and m_- the maximum and minimum monetary outcomes in the lottery. Set $u(m_+) = 1$ and $u(m_-) = 0$. Consider any other monetary outcome m, and the set of lotteries $\{([m_+, m_-], [p, (1-p)]) : p \in [0, 1]\}$. For $p = 1$ the lottery is preferred to m and for $p = 0$ the outcome m is preferred to the lottery. Using the continuity axiom, one can show that for some intermediate p^* the person is indifferent between m and that lottery. Set $u(m) = p^*$. Then use the other axioms to verify that the Bernoulli function so constructed indeed represents the given preferences.

Notice that this construction of the Bernoulli function suggests an empirical procedure: vary the probabilities on best and worst outcomes to try to find a person's point of indifference. There are many variations on this theme; one of the currently more popular is the Multiple Price List scheme introduced by Holt and Laury (*American economic review*, 2002).

The rest of the material in these notes will not be covered in Econ 200, but students taking a course in Finance may (or may not!) find it helpful.

[1] On the other hand, the EUT's conclusion is quite strong, and not consistent with some actual choice data. One response is to accommodate some of the anomalous data by weakening the axioms, usually the third or fourth. For an extensive and skeptical review of these matters, see my recent book *Risky Curves* (Routledge, 2014) coauthored with M. Isaac, D. James and S. Sunder.

0.1 Basic Definitions

The *return* k on an asset over a given period is defined in terms of the cash flow x received during that period and the beginning and end prices P_0 and P_1 as follows:

$$k = \frac{P_1 - P_0 + x}{P_0}. \tag{6}$$

Suppose that the return in situation (or scenario) s is k_s, and that each possible situation $s = 1, ..., S$ has probability $p_s > 0$, where $\sum_s p_s = 1$. Then the *expected return* is

$$Ek = \sum_{s=1}^{S} p_s k_s. \tag{7}$$

An important special case is with equally-weighted historical data. Each observation receives equal probability weight $p_s = 1/S$.

The *variance* is

$$\textsc{Var}\ k = \sigma_k^2 = E(k - Ek)^2 = \sum_{s=1}^{S} p_s (k_s - Ek)^2. \tag{8}$$

The square root of the variance, $\sigma_k = \sqrt{\textsc{Var}\ k}$ is called the *standard deviation* of the return, or the *volatility* of the asset. It is the most popular measure of risk.

The *covariance* between returns k and h on two different assets is

$$\textsc{Cov}(k, h) = \sigma_{kh} = E(k - Ek)(h - Eh) = \sum_{s=1}^{S} p_s (k_s - Ek)(h_s - Eh). \tag{9}$$

The *correlation* ρ_{kh} is covariance normalized by the volatilities. Thus $\rho_{kh} = \textsc{Cov}(k, h)/(\sigma_k \sigma_h)$. Standard theorems in mathematics demonstrate that $-1 \leq \rho_{kh} \leq 1$.

Consider a portfolio $x = (x_k, x_h)$ with initial dollar value x_k in asset k and initial value x_h in asset h. The numbers x_k, x_h are called *positions* or *exposures*. Usually we have $x_k > 0$, which is called a *long* position, but a *short* position $x_k < 0$ sometimes occurs, e.g., if you borrow an asset. The expected return on portfolio x with total initial investment $x_T = x_k + x_h$ is $(x_k/x_T)Ek + (x_h/x_T)Eh$. It is the obvious weighted average of the returns on the two assets.

0.2 Diversification

Portfolio variance is more complicated and interesting. First note that the definitions imply that:

- variance $\text{Var}\, ak = a^2 \text{Var}\, k$ increases in the square of the position a
- volatility $\sigma_{ak} = |a|\sigma_k$ increases linearly in the absolute position
- the portfolio with unit positions in two assets has variance

$$\begin{aligned} \text{Var}\,(k+h) &= E[(k+h) - E(k+h)]^2 = E(k-Ek)^2 + 2E(k-Ek)(h-Eh) + E(h-Eh)^2 \\ &= \text{Var}\, k + 2\text{Cov}(k,h) + \text{Var}\, h. \end{aligned} \tag{10}$$

From the first and third bullet items above, it follows that the variance on the portfolio x above is $x_k^2 \text{Var}\, k + 2x_k x_h \text{Cov}(k,h) + x_h^2 \text{Var}\, h$.

This piece of mathematics has a very important implication. It shows that risk can be reduced substantially by diversification. To see this clearly, suppose that the two assets have the same variance, say $\text{Var}\, k = \text{Var}\, h = 100$, and consider various portfolios where the positions add to 10, say $x_k = a$ and $x_h = 10 - a$. Clearly if all 10 is put in either asset and the other is left out, then the portfolio volatility is $|10|\sigma_k = 10 * 10 = 100$. As discussed in class (using diagrams in (σ_k, Ek) space), there are three special cases of interest:

1. if $\rho_{kh} = 0$, then the portfolio volatility is $\sqrt{a^2\sigma_k^2 + (10-a)^2\sigma_h^2} < 100$. For an equal-weight portfolio $a = 5$, for example, the portfolio volatility is only $\sqrt{25 * 100 + 25 * 100} = 50\sqrt{2} \approx 70.7$, almost a 30% reduction in volatility or riskiness.

2. if $\rho_{kh} = -1$, then risk can be eliminated entirely. Now $\text{Cov}(k,h) = \rho_{kh}\sigma_k\sigma_h = -1 * 10 * 10 = -100$. For $a = 5$, the portfolio volatility is $\sqrt{5^2\sigma_k^2 + 2 * 5 * 5\text{Cov}(k,h) + 5^2\sigma_k^2} = \sqrt{2500 - 5000 + 2500} = 0$. The idea is that when

$\rho = -1$, any fluctuation in the first asset is exactly offset by the fluctuation in the other asset, and risk is eliminated.

3. if $\rho_{kh} = 1$, then there is no diversification effect and all portfolios with the same total position have the same risk. The mathematical reason is that now $\text{Cov}(k, h) = \sigma_k \sigma_h$. In the example, the portfolio volatility is $\sqrt{a^2\sigma_k^2 + 2 * a * (10 - a)\sigma_k\sigma_h + (10 - a)^2\sigma_k^2}$ $= 10\sqrt{a^2 + 2 * a * (10 - a) + (10 - a)^2} = 10 * \sqrt{100} = 100$ for any a.

The take-home point here is that there is a reduction in risk when a given initial investment is spread among several assets whose returns are not perfectly correlated. For positive positions, the diversification effect is stronger when the correlation is smaller (or more negative). The effect strengthens as the fraction of the investment in one asset increases from zero, and eventually weakens as the fraction approaches one.

Exercises.

1. Show that if the volatilities of two assets are not equal, you can still find a riskless portfolio when $\rho = -1$, with positions related to the relative volatilities of the two assets.

2. Also show that risk can be eliminated when $\rho = 1$ by taking a short position in one asset and a long position in the other.

0.3 Marginal Risk in a Portfolio

The rest of these notes concern portfolios involving N risky assets with returns k_i for $i = 1, ..., N$ and covariance matrix $C = ((\sigma_{ij}))$. (Note that the diagonal elements $\sigma_{ii} = \sigma_i^2$ are the variances of the risky asset, and the off-diagonal elements are covariances as defined earlier.) The positions are denoted x_i and the total investment is $x_T = \sum_{i=1}^{N} x_i$. The portfolio shares are $a_i = (x_i/x_T)$. The reasoning in the previous section shows that the expected return on portfolio x is

$$E[x] = a \cdot Ek = \sum_{i=1}^{N} a_i Ek_i, \tag{11}$$

the weighted average expected return on the N assets, with weights given by the portfolio shares. The earlier reasoning also shows that the variance of the portfolio value is

$$\text{Var}[x] = x \cdot Cx = \sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j \sigma_{ij} = \sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j \sigma_i \sigma_j \rho_{ij}, \tag{12}$$

where $\rho_{ij} = \text{Cov}(k_i, k_j)/(\sigma_i \sigma_j)$ is the correlation between the returns on assets i and j. The portfolio volatility is denoted $\sigma_x = \sqrt{\text{Var}[x]}$.

The question is: Holding constant the overall investment, how does the portfolio volatility change as the the position in one particular asset i increases? Or more briefly, what is the marginal risk of asset i in portfolio x?

The question is key to portfolio analysis, and its answer requires a little linear algebra and vector calculus. Let $x = (x_1, ..., x_N)$ be the original portfolio and let $y = \alpha x_T e^i + (1-\alpha)x$ be the shifted portfolio, where $\alpha \geq 0$ is the proportional amount of the shift in direction $e^i = (0, ..., 0, 1, 0, ..., 0)$, the unit portfolio with only asset i. The question is how fast portfolio volatility σ_y changes as α increases from zero. The question and its answer are captured in:

Lemma 1. For portfolio y as just defined, we have

$$\left.\frac{d\sigma_y}{d\alpha}\right|_{\alpha=0} = \frac{-\sigma_x^2 + \sigma_{ix}}{\sigma_x}. \tag{13}$$

That is, portfolio volatility increases at a rate proportional to the covariance $\sigma_{ix} = x_T e^i \cdot Cx$ of asset i with the rest of the portfolio. The other term $-\sigma_x^2/\sigma_x = -\sigma_x$ in (13) simply reflects the reduced size of the rest of the portfolio when the asset i share increases.

The following proof can be skipped by readers who do not enjoy such things.

Proof. From equation (12) and the definition of y we have

$$\text{Var}[y] = y \cdot Cy = (\alpha x_T e^i + (1-\alpha)x) \cdot C(\alpha x_T e^i + (1-\alpha)x) \tag{14}$$

$$= (1-\alpha)^2 x \cdot Cx + \alpha^2 x_T^2 e^i \cdot Ce^i + 2\alpha x_T (1-\alpha) e^i \cdot Cx, \tag{15}$$

where the last expression uses the fact that $e^i \cdot Cx = x \cdot Ce^i$ since the covariance matrix is symmetric. Hence

$$\left.\frac{d\mathrm{Var}[y]}{d\alpha}\right|_{\alpha=0} = -2(1-\alpha)x \cdot Cx + 2\alpha x_T^2 e^i \cdot Cx + 2(1-2\alpha)x_T e^i \cdot Cx \,|_{\alpha=0} \tag{16}$$

$$= -2x \cdot Cx + 2x_T e^i \cdot Cx = -2\sigma_x^2 + 2\sigma_{ix}. \tag{17}$$

Marginal risk now can be calculated using the chain rule of ordinary calculus:

$$\left.\frac{d\sigma_y}{d\alpha}\right|_{\alpha=0} = \left.\frac{d\sqrt{\mathrm{Var}[y]}}{d\alpha}\right|_{\alpha=0} = \frac{1}{2}[\mathrm{Var}[y]]^{-1/2} \left.\frac{d\mathrm{Var}[y]}{d\alpha}\right|_{\alpha=0} = \frac{-2\sigma_x^2 + 2\sigma_{ix}}{2\sigma_x}, \tag{18}$$

which immediately simplifies to the desired expression. $\diamond$

0.4 The Capital Asset Pricing Model (CAPM)

CAPM is no longer unchallenged orthodoxy but it still is the central model in modern financial theory. Its main result is that the expected return of each risky asset i is the risk-free return k_F, plus a risk premium which is the market price of risk RP_M times the quantity of risk β_i for that asset. That is, we have the

Security Market Line Theorem. If conditions B1, B2, M1 and M2 below hold, then the return on each asset i satisfies

$$Ek_i = k_F + \beta_i \mathrm{RP}_M \tag{19}$$

where $\mathrm{RP}_M = E[M] - k_F$ is the expected return on the Market portfolio—the portfolio consisting of all assets traded in financial markets—in excess of the risk-free return, and $\beta_i = \sigma_{iM}/\sigma_M^2 = \rho_{iM}\sigma_i/\sigma_M$ is the normalized covariance of asset i with the market portfolio.

The conclusion follows from various sets of assumptions. One short, convenient and transparent set of behavioral (B) and market (M) assumptions is as follows.

- (B1) Investors are rational and care only about portfolio expected return (+) and volatility (-).

- (B2) They have homogeneous beliefs about asset returns Ek and covarariances C.
- (M1) A financial market lasts just one period (from $t = 0$ to $t = 1$); it is competitive and frictionless (i.e., no market power and zero transactions costs) and is in equilibrium.
- (M2) All N risky assets are traded on the financial market plus one risk-free asset with return $k_F > 0$. The portfolio M consisting of all assets has expected return $E[M] > k_F$.

Here is a sketch of the proof of the SML Theorem; see a standard text for diagrams, etc. First plot in (σ, Ek)-space the volatilities and expected return for each risky asset in isolation. Using equations (3, 11, 12), note that the portfolios consisting of varying shares of two or more risky assets trace out parabolic arcs that fill a convex region of that space. Efficient risky portfolios (those with the largest expected return for given volatility, or smallest volatility for given expected return) lie along the Northwestern frontier of this region.

The Capital Market line (CML) has one endpoint at the risk-free point F with coordinates $\sigma = 0$ and $Ek = k_F > 0$, and passes through a point $T = (\sigma_T, E[T])$ of tangency to the Northwestern frontier. That is, the line satisfies the equation

$$E[x] = k_F + [(E[T] - k_F)/\sigma_T]\sigma_x \tag{20}$$

Since the market is frictionless, each investor can borrow or lend as much as he or she wants at the risk-free rate. Taking this into account, each efficient portfolio must lie on the CML, and by (B1) each investor will choose an efficient portfolio. Thus we have the

Markowitz Separation Theorem. Each investor's I's portfolio is of the form

$$w_I a_I F + w_I(1 - a_I)T, \tag{21}$$

that is, she lends some fraction a_I of her wealth w_I at the risk-free rate k_F (or borrows if $a_I < 0$) and invests the rest in the tangency portfolio.

The striking implication is that investors' risk preferences do not affect the mix of risky assets—all investors choose the same combination T. It's just that more risk averse investors put more of their wealth into the riskless asset, and the least risk averse investors borrow at the riskless rate to buy more of the T portfolio, i.e., they use leverage. This striking implication, of course, is not exactly true in the real world, but some finance economists argue that the vast majority of investors (at least in terms of wealth) choose portfolios that are near T in (σ, Ek)-space.

The next step in the argument (due to Sharpe and Lintner in the 1960s) is that financial market equilibrium implies that the tangency portfolio T must be the same as the market portfolio M. To spell it out, the market demand for risky assets is the N-vector $w\tau$, where $w = \sum_I w_I(1-a_I)$ is the total wealth invested in risky assets, and $\tau = (\tau_1, ..., \tau_N)$ is the unit vector of assets that produces the point T in (σ, Ek)-space. The supply of risky assets is the N-vector $(s_1, ..., s_N)$. But in equilibrium, supply = demand, i.e., $s_i = w\tau_i$ for each asset i. Thus portfolios M and T have identical shares of the risky assets and therefore represent the same point in (σ, Ek)-space.

The SML (19) now follows from direct calculations. Let $y = \alpha e^i + (1-\alpha)M$ be a small shift in the market portfolio towards asset i along the efficient frontier. The slope of the frontier at M in (σ, Ek)-space, according to the implicit function theorem, is given by a ratio of derivatives, the numerator being $\frac{dE[y]}{d\alpha}\Big|_{\alpha=0}$ and the denominator being $\frac{d\sigma_y}{d\alpha}\Big|_{\alpha=0}$. By (11) the numerator is $\frac{d}{d\alpha}[\alpha Ek_i + (1-\alpha)E[M]]\big|_{\alpha=0} = Ek_i - E[M]$. The denominator is given by Lemma 1 of the previous section as $\frac{-\sigma_M^2+\sigma_{iM}}{\sigma_M}$. Since the CML is tangent to the frontier at M, the slope of the frontier just computed must be equal to the slope of the CML given by the bracketed term in (20), i.e.,

$$\frac{Ek_i - E[M]}{\frac{-\sigma_M^2+\sigma_{iM}}{\sigma_M}} = \frac{(E[M]-k_F)}{\sigma_M}. \tag{22}$$

Cross-multiplying we get

$$Ek_i - E[M] = (E[M]-k_F)\frac{-\sigma_M^2+\sigma_{iM}}{\sigma_M^2} = (E[M]-k_F)\frac{\sigma_{iM}}{\sigma_M^2} - E[M] + k_F. \tag{23}$$

Cancelling the $-E[M]$ terms on both sides of the equation we have exactly the expressions given in the Security Market Line Theorem.