A Primer on Portfolio Theory

Part I: Some Basics

The bottom line with portfolio construction is learning how to deal with uncertainty. To begin, let’s start with some definitions:

- A random variable is a number associated with an outcome that is uncertain. For example, suppose that you roll a fair dice. The result of that roll is a random variable. Every random variable can be described by a probability density function. This is just a list of the possible outcomes with each outcome’s probability of occurring. The probability density for rolling a fair dice would be as follows (note that the probabilities must always sum to one – something has to happen!):

<table>
<thead>
<tr>
<th>X</th>
<th>P(X)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/6</td>
</tr>
<tr>
<td>2</td>
<td>1/6</td>
</tr>
<tr>
<td>3</td>
<td>1/6</td>
</tr>
<tr>
<td>4</td>
<td>1/6</td>
</tr>
<tr>
<td>5</td>
<td>1/6</td>
</tr>
<tr>
<td>6</td>
<td>1/6</td>
</tr>
</tbody>
</table>

- For many applications, it is not necessary to know the exact probability density for a random variable. All that is needed are some descriptive statistics. The two most common statistics are expected value and variance. The expected value of a random variable, X, is the weighted-average of all possible values where the weights are the probabilities that each value will occur.

\[
\bar{X} = E(X) = \sum p_i X_i
\]

For example, the expected value of the dice roll would be as follows:

\[
E(X) = (1/6) \times 1 + (1/6) \times 2 + (1/6) \times 3 + (1/6) \times 4 + (1/6) \times 5 + (1/6) \times 6 = 21/6 = 3.5
\]

The expected value is meant to capture the central tendency of a random variable (i.e. the average). The variance of a random variable, X, is the weighted-average of squared deviations from the expected value where, again, the weights are the probabilities of each event.

\[
\sigma^2 = Var(X) = \sum p_i (X_i - E(X))^2
\]
For the dice roll, the variance would be as follows:

\[ \text{Var}(X) = \frac{1}{6} (1 - 3.5)^2 + \frac{1}{6} (2 - 3.5)^2 + \frac{1}{6} (3 - 3.5)^2 + \frac{1}{6} (4 - 3.5)^2 + \frac{1}{6} (5 - 3.5)^2 + \frac{1}{6} (6 - 3.5)^2 = \frac{17.5}{6} = 2.91 \]

The variance is meant to describe the degree to which a variable fluctuates around its mean. Sometimes standard deviation is used rather than the variance. The standard deviation is simply the square root of variance.

\[ \sigma = \text{Std.Dev}(X) = \sqrt{\text{Var}(X)} \]

So, for example, the standard deviation of the dice roll is 1.71.

- The last statistic of interest involves the relationship between two random variables and is called the covariance. The covariance of two random variables \( X \) and \( Y \) is calculated by computing the expected value of the product of \( X \) and \( Y \) and then subtracting the product of the expected values of \( X \) and \( Y \)

\[ \sigma_{xy} = \text{Cov}(X,Y) = E(XY) - E(X) \cdot E(Y) \]

If there is a large probability that \( X \) and \( Y \) will take on large or small values at the same time then the covariance between \( X \) and \( Y \) will be positive. If there is a large probability that \( X \) will be large when \( Y \) is small and vice versa, then the covariance will be negative. Sometimes, instead of covariance, correlation is used. The correlation is simply the covariance divided by the product of standard deviations:

\[ \rho_{xy} = \frac{\text{Cov}(X,Y)}{\sigma_x \sigma_y} \]

**Part II: Some useful formulas**

There are some properties of expected value, variance, and covariance that will be very useful for portfolio construction.

1) If \( X \) is a random variable and ‘\( k \)’ and ‘\( l \)’ are constant, then:

\[ E(kX) = kE(X) \]
\[ \text{Var}(kX) = k^2 \text{Var}(X) \]
\[ \text{Cov}(kX, lY) = kl\text{Cov}(X,Y) \]

2) If \( X \) and \( Y \) are random variables, then the sum, \( X+Y \), is also a random variable with
Given (1) and (2) above, we can calculate the expected value and variance of any linear combination of random variables. For example, suppose we have the following two variables X and Y. If we define a new variable as $Z = kX + lY$ where $k$ and $l$ are constants, then the new variable $Z$ has the following expected value and variance:

$$E(Z) = kE(X) + lE(Y)$$
$$Var(Z) = k^2Var(X) + l^2Var(Y) + 2klCov(X, Y)$$

### Part III: The power of diversification

Diversification involves choosing a portfolio of several stocks instead of holding a single stock. The idea here is that each company has associated with it some idiosyncratic risk. That is, each individual company might respond slightly differently to various events (i.e., when it rains, umbrella company stocks go up while picnic basket companies stocks go down). By diversifying, you will be able to lower the variance of your overall portfolio return. However, there are limits to how much variance you can diversify away. Suppose that you have a portfolio composed of two stocks: $w_1$ is the percentage of your overall investment in stock 1, $w_2$ is the percentage of your overall investment in stock 2. From the formulas above, the variance of your overall portfolio will be as follows:

$$\sigma_p^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2Cov(1, 2)$$

Your ability to diversify depends on the covariance between stocks 1 and 2. When the covariance between stocks 1 and 2 is negative, we are dealing with idiosyncratic (company specific) risk. This can be diversified away. Note, however, that as the covariance between stocks 1 and 2 gets large and positive, the portfolio variance gets bigger. Positive covariance (stocks that move in the same direction) is a sign of systemic risk. For example, all stocks typically fall if oil prices rise. Systemic risk cannot be diversified away.

Let's be a bit more rigorous about the power of diversification. Suppose that you have a portfolio of ‘n’ stocks. $w(i.)$ refers to the percentage of your overall portfolio invested in stock i. From the above formulas, we can calculate the variance of your portfolio as follows:

$$\sigma_p^2 = \sum_{i=1}^{n} w_i^2\sigma_i^2 + \sum_{j=1}^{n} \sum_{i=1}^{n} w_iw_jCov(i, j)$$
Now, suppose that you follow a naïve investment strategy of dividing your portfolio equally among each stock. That is, \( w(i) = \frac{1}{n} \). Then, the above formula becomes

\[
\sigma_p^2 = \sum_{i=1}^{n} \left( \frac{1}{n} \right) \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} \left( \frac{1}{n^2} \right) \text{Cov}(i, j)
\]

Now, suppose that all stocks have a common variance and that all security pairs have a common covariance. Then, the above formula simplifies to the following:

\[
\sigma_p^2 = \left( \frac{1}{n} \right) \sigma^2 + \left( \frac{n-1}{n} \right) \text{Cov}
\]

The following chart shows the resulting portfolio standard deviation as more and more securities are added for three cases: positive correlation (systemic risk), zero correlation, and negative correlation (idiosyncratic risk).

<table>
<thead>
<tr>
<th>Number of Stocks (n)</th>
<th>Proportion (1/n)</th>
<th>Standard Dev. (corr = .4)</th>
<th>Standard Dev. (corr = 0)</th>
<th>Standard Dev. (corr = -.4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>.5</td>
<td>41.83</td>
<td>35.36</td>
<td>35.35</td>
</tr>
<tr>
<td>5</td>
<td>.2</td>
<td>36.06</td>
<td>22.36</td>
<td>22.35</td>
</tr>
<tr>
<td>6</td>
<td>.1667</td>
<td>35.36</td>
<td>20.41</td>
<td>20.40</td>
</tr>
<tr>
<td>10</td>
<td>.1</td>
<td>33.91</td>
<td>15.81</td>
<td>15.80</td>
</tr>
<tr>
<td>11</td>
<td>.0909</td>
<td>33.71</td>
<td>15.08</td>
<td>15.06</td>
</tr>
<tr>
<td>20</td>
<td>.05</td>
<td>32.79</td>
<td>11.18</td>
<td>11.16</td>
</tr>
<tr>
<td>21</td>
<td>.0476</td>
<td>32.73</td>
<td>10.91</td>
<td>10.89</td>
</tr>
<tr>
<td>100</td>
<td>.01</td>
<td>31.86</td>
<td>5.00</td>
<td>4.96</td>
</tr>
<tr>
<td>101</td>
<td>.0099</td>
<td>31.86</td>
<td>4.98</td>
<td>4.93</td>
</tr>
</tbody>
</table>

**Part IV: The Markowitz Portfolio Selection Model**

In the above section, we used a very naïve investment strategy (ie, simply dividing our portfolio equally between various stocks). The Markowitz method is a more sophisticated investment strategy. Here, we take the properties of the stocks as given (expected value, variance, and covariance), and choose the weights in our portfolio based on these statistics. Specifically, assume the following:

1) There are two risky assets (stocks) available. Each has an associated expected return and variance:

\[
E(r_a) = \bar{r}_a \\
E(r_b) = \bar{r}_b \\
\text{Var}(r_a) = \sigma_a^2 \\
\text{Var}(r_b) = \sigma_b^2
\]
2) The stocks have a covariance given by

\[ \text{Cov}(r_a, r_b) = \sigma_{ab} \]

3) There is a risk free asset available

\[ E(r_f) = r_f \]
\[ \text{Var}(r_f) = 0 \]

4) The investor is choosing the proportions of the portfolio invested in stock \( a \), stock \( b \) and the risk free asset in order to maximize the following:

\[ \frac{E(r_p) - r_f}{\sigma_p} \]

In other words, the investor is choosing weights to maximize his/her reward-to-risk ratio.

This problem is done in two steps. First, suppose that the risky portion of your portfolio is divided into \( w(a) \) percent in stock \( a \) and \( w(b) \) in stock \( b \). We know that the expected return and variance associated with this division is as follows:

\[ E(r_p) = w_a r_a + w_b r_b \]
\[ \text{Var}(r_p) = w_a^2 \sigma_a^2 + w_b^2 \sigma_b^2 + 2 w_a w_b \text{Cov}(r_a, r_b) \]

Suppose that we plotted the variance associated with various weights in stock \( a \). We would something like the following:
Point A would be the minimum variance portfolio. Mathematically, we could solve for the weights associated with point A.

\[
\begin{align*}
    w_a &= \frac{\sigma_b^2 - \text{Cov}(r_a, r_b)}{\sigma_a^2 + \sigma_b^2 - 2\text{Cov}(r_a, r_b)} \\
    w_b &= 1 - w_a
\end{align*}
\]

Assume that the variance and return of the minimum variance portfolio is given by the following:

\[
\begin{align*}
    E(\text{return}) &= r \\
    \text{var(\text{return})} &= \sigma^2
\end{align*}
\]

While these weights are useful, they are not our ultimate goal. We are interested in the tradeoff between risk (variance) and return. Suppose that you plot the expected return versus the portfolio standard deviation associated with the various portfolio weights. You would get something like the diagram below:

The minimum variance portfolio is still given by point A. Clearly, any weights associated with points below point A are not optimal. We could find other weights with the same standard deviation, but with higher returns. Any point above A is a possible optimum (with a positive risk/return tradeoff).

Suppose that we chose to allocate a fraction ‘y’ of our total portfolio in risky assets (where the weights on the two assets are chosen to minimize their combined risk) and the remaining (1-y) in the risk free asset. We can calculate the expected return and standard deviation of our overall portfolio.
First, note that the ratio of risk to return is independent of ‘y’. Secondly, note that graphically, the ratio of these is simply the slope of a line connecting point A with the point \((0, r(f))\) on the vertical axis.

Can we do better than this? Sure we can! Suppose that, rather than choosing the minimum variance portfolio, we selected the weights on our risky assets such that we were at point B on the above graph. Clearly, point B has a better risk/reward tradeoff (as seen by the steeper slope). In fact, it appears that point B generates the steepest possible line between \(r(f)\) and a point on the curve. Therefore, point B represents the optimal portfolio selection. Specifically, point B is the weights for the two risky assets that maximize the reward/risk ratio. The actual weights are as follows:

\[
w_a = \frac{\left( r_a - r_f \right) \sigma_a^2 - \left( r_b - r_f \right) \sigma_b^2 + \left( r_a - r_b \right) \sigma_{a,b}^2}{\left( r_a - r_f \right) \sigma_a^2 + \left( r_b - r_f \right) \sigma_b^2 - \left( r_a + r_b - 2r_f \right) \sigma_{a,b}^2}
\]

\[w_b = 1 - w_a\]

Finally, once the portfolio associated with point B has been chosen, (with its associated return and standard deviation), the overall portfolio depends on the amount allocated towards the risky assets \((y)\). This decision will be based purely on preferences towards risk. (Different y’s will simply be different points on the line connecting A and B).