1

Definitions of risk and return

1.1 Introduction

Mathematics can be applied to the practice of finance in a number of ways. These include trying to use mathematics to predict asset price movements (*statistical arbitrage*), measuring and controlling risk in trading books (*risk management,*) pricing options and other contingent claims by assessing hedging strategies (*derivatives pricing,*) and the use of mathematics to maximise the risk–return trade-off when investing in the markets: *portfolio theory*. It is portfolio theory which we will address in this volume. This subject is sometimes called *modern portfolio theory* or *MPT*.

At first glance the objective of maximising the risk–return trade-off when investing in the markets appears straightforward and intuitive; that is, any rational investor will want to maximise the anticipated return on his or her investment whilst minimising the risk of unexpected loss. However, in order to apply the rigour of mathematics to this activity, we need first to carefully define these terms. What is risk? What is return? How do we decide the trade-off between them? There are multiple ways of doing this, and we will examine the more widely-used ones.

Throughout we will make two fundamental assumptions. The first is that individual assets are correctly priced. This means that "stock picking" is pointless and, accordingly, our efforts will focus on how to compare portfolios with each other.

Second, to ensure consistency, we will generally work across a fixed timeframe, for example, one year. We should think of ourselves as a funds manager whose performance is assessed on a yearly basis. The funds manager will be given a statement by his/her client or the board stating the required risk–return trade-off and then it becomes his or her job to achieve it.

1.2 Measuring return

The return on an asset is the percentage change in its value over a given time period. A negative return is possible (and the financial markets upheavals of 2008 and beyond have seen many negative returns). Notably, change can occur in multiple fashions. For example, in the case of a stock, its market price can vary both up and down due to company performance and general market conditions. Second, the stock may pay dividends which will always be considered as part of the return. Dividends may be paid either in cash, or as *scrip* dividends in the form of additional shares. For a scrip dividend the number of shares held by the investor increases but no cash changes hands; value increases and the contribution to return is positive. By way of contrast, certain exotic financial instruments as well as real-world assets require the holder to pay money in order to retain his or her rights: this results in negative cash-flows which must also be considered as part of the return.

Definition 1.1 The *return* on a portfolio is the percentage change in its value taking into account all cash in-flows and out-flows.

In financial mathematics, we are generally interested in the future rather than the past, so the return will normally be uncertain. It is therefore *expected* return that is important rather than *actual* return. We therefore assume that the return, R, follows some probability distribution with density f. And the *expected return* is

$$\mathbb{E}(R) = \overline{R} = \int R f(R) dR.$$

In simple examples, R will be a discrete random variable, and f is then a sum of point densities (delta functions) and R follows some probability distribution taking values R_i with probability p_i.

The *expected return* is then

$$\mathbb{E}(R) = \overline{R} = \sum_{i=1}^{n} p_i R_i.$$

For example, if R has probabilities of taking values as follows

p_i	R_i
$\frac{1}{3}$	5%
$\frac{1}{6}$	6%
$\frac{1}{2}$	7%

,

then the expected percentage return is

$$\frac{1}{3}5+\frac{1}{6}6+\frac{1}{2}7.$$

The expectation operator is linear; that is

$$\mathbb{E}(aX+bY)=a\mathbb{E}(X)+b\mathbb{E}(Y),$$

and so for a portfolio P consisting of assets A_i, with return R_i, in proportions X_i, we have

$$\mathbb{E}(R_P)=\overline{R}_P=\sum_{i=1}^{n}X_i\overline{R}_i.$$

We can write this as

$$\overline{R}_P=\langle X,\overline{R}\rangle, \tag{1.1}$$

with $X=(X_1,\ldots,X_n)$ and $\overline{R}=(\overline{R}_1,\ldots,\overline{R}_n)$.

Here $\langle x,y\rangle$ denotes the inner product of two vectors: $\sum_{i=1}^{n} x_i y_i$.

1.3 Portfolio constraints

Generally, when given the task of building an investment portfolio, we will be entrusted with a fixed sum of money to divide between a total of n possible assets, $A_1,\ldots A_n$. We will solely be interested in what fraction of our money to put into each asset and hence we can take the sum of money to be 1. If we put X_i into each asset our constraint becomes

$$\sum_{i=1}^{n}X_i=1. \tag{1.2}$$

What further constraints do we put on X_i? Often, there is a requirement that all holdings be non-negative. In other words, we are not allowed to *short-sell* assets. Effective short-selling is possible in the markets under certain constraints, although in the wake of the 2008 global financial crisis, various regulators intervened to ban the short-selling of financial stocks in attempts to control speculation and spiraling confidence loss. The long-term efficacy of this strategy has been a topic of some debate and in any case, derivatives can generally be used to obtain the same effects. For a portfolio prohibiting short-sales, we will have the extra constraint

$$X_i\geq 0.$$

It is also not unusual for funds to place restrictions on the fraction of wealth that can be put into a single asset so often there will be a single asset constraint of the form

$$X_i \leq \varepsilon_i,$$

for some given ε_i.

Other commonly-encountered portfolio constraints include restrictions on the fraction of wealth to be invested in particular geographies, industries, credit ratings or asset classes. These may be set to reflect preferences of the board or client as the case may be and will vary according to the investment strategy of a particular fund or portfolio.

Note that if our sole objective is to maximise expected return, the portfolio selection problem is easy to solve. We simply put as much money as possible into the highest returning asset. In other words, we find the asset, A_j, that maximises $\overline{R}_j$ and invest all the money in that. If there is a constraint on how much money can be placed in each asset, then we put as much as we can in A_j and then as much of what is left as we can in the next best and so on.

The reason that there is some work (and mathematics) to this subject is that generally there is a requirement to control risk to some stated level of "risk appetite" as well as to maximise returns. To control risk we first need to define it and to guide our thinking, we start by looking at some very simple examples.

Example 1.2 Suppose we have to choose between two assets. Asset A pays \$ 1,000,000 with 25% probability and pays 0 with 75% probability. Asset B however pays \$ 250,000 with 100% probability.

Which would an investor prefer?

Both assets have the same mean of \$ 250,000. However, B guarantees the investor will receive the mean whereas A involves a great deal of risk.

Generally then, B would be preferred as it involves no risk. ◇

Example 1.3 Suppose that now we have to choose between two assets as follows. Asset A pays \$ 1,000,000 with 25% probability and pays 0 with 75% probability. Asset B pays \$ 260,000 with 100% probability.

Which would an investor prefer?

Asset B has higher mean and lower risk. You would have to be very risk-loving to prefer A. Note, however, that if you play roulette or a lottery, then A is the sort of investment you are making. Of course, owning a casino or running a lottery is a different matter and is highly recommended.

Incidentally, the mathematician's way of playing the lottery is to pick some numbers and then not buy a ticket. When the numbers do not come up, he has won a dollar. ◇

Whilst the above two examples are straightforward, the next raises some more interesting points.

Example 1.4 Asset A pays \$ 1,000,000 with 25% probability and pays 0 with 75% probability. Asset B however pays \$ 240,000 with 100% probability.

Which would an investor prefer?

Asset B has lower mean and lower risk. Most investors would prefer B on the grounds that the extra risk is not worth the extra \$10,000 to be gained on average. However, some might go for A since if one had the opportunity to do many such independent investments then the risk would average out (through diversification) and A would be preferable. This is in essence what banks do when they make a very large number of small loans to different customers and is often referred to as taking a "portfolio view." ◇

Example 1.5 Now suppose we have two assets A and B. A coin is tossed and A pays 1 on heads and zero otherwise. Asset B pays 1 on tails and zero otherwise. The two assets are based on the same coin toss. How much are A and B worth?

The mean pay-off for each asset is 0.5, yet we would expect the value to be lower because of risk aversion. We would also expect the two assets to trade at the same price. However, if we consider the portfolio of A and B together then it will always be worth 1 (the assets are complementary). We therefore conclude that the individual assets are worth 0.5 despite risk aversion.

This example illustrates the fact that a risk premium is generally not available for risk that is diversifiable or hedgeable. Whilst we will generally be unable to remove all risk, we will be able to remove some via portfolio diversification. ◇

1.4 Defining risk with variance

There are many ways to define and control risk. The first and simplest way is to use variance. The variance of a random variable is defined via

$$\mathrm{Var}(R) = \mathbb{E}((R - \overline{R})^2) = \mathbb{E}(R^2) - \mathbb{E}(R)^2.$$

The standard deviation is a related measure of risk. It is defined by

$$\sigma_R = (\mathrm{Var}\, R)^{\frac{1}{2}}.$$

It therefore contains the same information as the variance.

Standard deviation is harder to work with computationally because of the square root, but has the virtue that it has the same scale as the expectation. That is, we have

$$\mathbb{E}(\lambda R) = \lambda \mathbb{E}(R),$$
$$\mathrm{Var}(\lambda R) = \lambda^2 \mathrm{Var}(R),$$
$$\sigma_{\lambda R} = |\lambda| \sigma_R.$$

Note the important modulus sign, $||$, in the final equation: standard deviation is never negative.

We will be interested in the variance of portfolio returns given the variances of individual asset's returns. If we have assets with returns $R_1, \ldots, R_n$, held in amounts $X_1, \ldots, X_n$ then we can compute the variance of the portfolio.

We proceed by direct computation. We seek the value of

$$\mathrm{Var}(R_P) = \mathrm{Var}\left(\sum_{i=1}^{n} X_i R_i\right).$$

We compute

$$\begin{aligned}
\mathrm{Var}\left(\sum_{i=1}^{n} X_i R_i\right) &= \mathrm{Var}\left(\sum_{i=1}^{n} X_i (R_i - \mathbb{E}(R_i))\right), \\
&= \mathbb{E}\left(\left(\sum_{j=1}^{n} X_j (R_j - E(R_j))\right)^2\right), \\
&= \mathbb{E}\left(\sum_{i=1}^{n} X_i (R_i - \mathbb{E}(R_i)). \sum_{j=1}^{n} X_j (R_j - \mathbb{E}(R_j))\right), \\
&= \sum_{i,j=1}^{n} X_i X_j \mathbb{E}((R_i - \mathbb{E}(R_i))(R_j - \mathbb{E}(R_j))).
\end{aligned}$$

We define the covariance of R_i and R_j via

$$\mathrm{Cov}(R_i, R_j) = \mathbb{E}((R_i - \overline{R}_i)(R_j - \overline{R}_j)). \tag{1.3}$$

(Recall $\overline{R}_i = \mathbb{E}(R_i)$.)

So

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i R_i\right) = \sum_{i,j=1}^{n} X_i \mathrm{Cov}(R_i, R_j) X_j. \tag{1.4}$$

If we let C be a matrix with entries

$$C_{ij} = \mathrm{Cov}(R_i, R_j),$$

we can rewrite the variance as

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i R_i\right) = x^t C x. \tag{1.5}$$

We call C the *covariance matrix* of the returns. It is clearly symmetric: that is $C_{ij} = C_{ji}$ for all i and j. In addition, the diagonal entries are simply the variances of the individual assets. Covariances of asset pairs may be negative. For example, if whenever one asset goes up, another tends to go down, they will be negatively correlated.

Note that to interpret (1.5) correctly, we regard x as a vector which is a matrix with one column and n rows. The matrix C is of size $n \times n$. The transpose of x is written x^{T}, and has one row and n columns, so

$$x^{\mathrm{T}} = (X_1, X_2, \ldots, X_n).$$

We are multiplying a $1 \times n$ matrix by a $n \times n$ matrix, and then by a $n \times 1$ matrix to get a 1×1 matrix, i.e. a number. Since this number is a variance, it is always greater than or equal to zero.

Definition 1.6 If C is a symmetric matrix and

$$x^{\mathrm{T}} C x \geq 0,$$

for all x, then C is said to be positive semi-definite. It is said to be positive definite if $x^{\mathrm{T}} C x > 0$, for $x \neq 0$.

So all covariance matrices are positive semi-definite, and furthermore it can be shown that any positive semi-definite matrix is the covariance matrix of some collection of random variables. (See [8].)

Note that the variance of an asset is the covariance of an asset with itself. It follows from (1.4) that the variance of a portfolio will be the sum of the variances of the individual assets if and only if the covariances between the pairs of the different asset pairs are all zero. That is, if and only if the assets are uncorrelated. We then have

$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i R_i\right) = \sum_{i=1}^{n} X_i^2 \operatorname{Var} R_i.$$

Note that we can write the covariance of two asset returns R and S as

$$\operatorname{Cov}(R, S) = \sigma_R \sigma_S \rho_{RS},$$

where ρ_{RS} is referred to as the *correlation coefficient* and is defined in such a way as to make this statement true. Assets will have zero correlation if and only if they have zero covariance.

One condition that will lead to zero correlation is the much stronger condition of independence. If two random variables, C, D, are independent then

$$\mathbb{E}(CD) = \mathbb{E}(C)\mathbb{E}(D),$$

which in turn implies that

$$\operatorname{Cov}(C,D) = \mathbb{E}(C-\overline{C})\mathbb{E}(D-\overline{D}) = 0.$$

When considering independent assets we might ask what happens if we take a large number of independent assets and invest the same fraction of our wealth in each? That is, suppose we take n assets and put $1/n$ into each asset. We have

$$\operatorname{Var}\left(\sum_{i=1}^{n}\frac{1}{n}R_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var}(R_i).$$

If we assume that $\operatorname{Var}(R_i) \leq C$ for some C for all i then we have

$$\operatorname{Var}\left(\sum_{i=1}^{n}\frac{1}{n}R_i\right) \leq \frac{C}{n},$$

as n goes to infinity the variance will go to zero. This says that given a great enough number of independent assets, we can achieve an arbitrarily small amount of portfolio risk. What happens then if we allow covariance to be non-zero? In this case, we get

$$\begin{aligned}\operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^{n}R_i\right) &= \frac{1}{n^2}\sum_{i=1}^{n}\operatorname{Var}(R_i) + \frac{2}{n^2}\sum_{i<j}\operatorname{Cov}(R_i,R_j)\\ &= \frac{1}{n}\overline{\operatorname{Var}(R_i)} + \frac{n-1}{n}\overline{\operatorname{Cov}(R_i,R_j)}.\end{aligned}$$

Here we have used the fact that

$$1+2+3+\cdots n-1 = \frac{n(n-1)}{2},$$

so the number of elements in the sum $\sum_{i<j}$ is $n(n-1)/2$. Letting n tend to infinity, the variance converges to

$$\overline{\operatorname{Cov}(R_i,R_j)}.$$

Thus by taking equal proportions of a large number of assets, we obtain a portfolio whose variance is the average covariance of the assets in the pool. This tells us a very important fact that the background covariance in a pool of assets will have an effect on how much risk we can diversify away. This suggests that we want to invest in as large a class of assets as possible.

1.5 Other risk measures

Variance can be criticised for penalising upside volatility as well as downside volatility. We generally only care about our possibility of loss, not our possibility of a windfall gain. With this in mind, we can define the *semi-variance* of a variable X via

$$\mathbb{E}((X-\mathbb{E}(X))^2 I_{X<\mu}),$$

where $I_{X<\mu}$ equals 1 for $X < \mu$ and 0 otherwise. Markowitz therefore devoted a chapter of his book, [11], to the problem of portfolio selection using semi-variance rather than variance. Here, however, we will tend to focus on cases where X is reasonably symmetric. It follows then that the semi-variance will not give much beyond the variance and so we will not study it further.

Another measure of risk, highly popular with financial institutions and regulators alike, is Value-At-Risk (VAR). The idea here is to define a maximum loss to be tolerated at a given level of probability over a given time horizon. We will discuss this measure in some detail, both in terms of its uses and shortcomings in Chapter 12.

1.6 Review

By the end of this chapter, the reader should be able to answer the following theoretical questions.

1. What is the objective of modern portfolio theory?
2. How is return defined in MPT?
3. How is expected return defined?
4. How do we maximise return if there are no risk constraints?
5. Derive the formula for the variance of returns of a portfolio.
6. What is a covariance matrix?
7. What special properties does a covariance matrix have?
8. Derive the formula for the variance of return of a large pool of correlated assets.
9. Define semi-variance.

1.7 Problems

Question 1.1 An asset has the following distribution of returns. Compute the mean return, standard deviation of returns, and semi-variance of returns.

Probability	Return
0.1	-3
0.1	-2
0.2	-1
0.2	0
0.2	1
0.1	2
0.1	10

Question 1.2 Suppose x and y are vectors. Let

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

For some fixed y, let $f(x) = \langle x, y \rangle$. Compute $\frac{\partial f}{\partial x_i}$.

Question 1.3 Let A be a square matrix with n rows. Let A^{T} denotes its transpose. Let u and v be vectors with n entries. Show

$$\langle Au, v \rangle = \langle u, A^{\mathrm{T}} v \rangle = u^{\mathrm{T}} A^{\mathrm{T}} v.$$

Question 1.4 Let A be a real symmetric square matrix (so $A = A^{\mathrm{T}}$), and let

$$g(x) = x^{\mathrm{T}} A x.$$

Compute

$$\frac{\partial g}{\partial x_k}.$$

Question 1.5 Assets A, B and C have expected returns of 8, 10, and 12. Their standard deviations of returns are 12, 10, and 8. The pairwise correlations are all 0.5. Find the variances and expected returns of the following portfolios:

- equal weights of all assets;
- 0.5 units of A, 0.25 units of B, and 0.25 units of C.

Question 1.6 Repeat the previous question with standard deviations 10, 10, and 10.

Question 1.7 Assume that the average standard deviation of return for an individual security is 7 and that the average correlation is 0.2. Estimate the standard deviation of returns of portfolios composed of 10 and 100 securities.