Xinfu Chen

Mathematical Finance II

Department of mathematics,
University of Pittsburgh
Pittsburgh, PA 15260, USA

MATHEMATICAL FINANCE II

Course Outline

This course is an introduction to modern mathematical finance. Topics include

1. single period portfolio optimization based on the mean-variance analysis, capital asset pricing model, factor models and arbitrage pricing theory.
2. pricing and hedging derivative securities based on a fundamental state model, the well-received Cox-Ross-Rubinstein's binary lattice model, and the celebrated Black-Scholes continuum model;
3. discrete-time and continuous-time optimal portfolio growth theory, in particular the universal log-optimal pricing formula;
4. necessary mathematical tools for finance, such as theories of measure, probability, statistics, and stochastic process.

Prerequisites

Calculus, Knowledge on Excel Spreadsheet, or Matlab, or Mathematica, or Maple.

Textbooks

David G. Luenberger, INVESTMENT SCIENCE, Oxford University Press, 1998.

Xinfu Chen, LECTURE NOTES, available online www.math.pitt.edu/~xfc.

Recommended References

Steven Roman, INTRODUCTION TO THE MATHEMATICS OF FINANCE, Springer, 2004.

John C. Hull, OPTIONS, FUTURES AND OTHER DERIVATIVES, Fourth Edition, Prentice-Hall, 2000.

Martin Baxter and Andrew Rennie, FINANCIAL CALCULUS, Cambridge University Press, 1996.

P. Wilmott, S. Howison & J. Dewynne, THE MATHEMATICS OF FINANCIAL DERVIATIVES, CUP, 1999.

Stanley R. Pliska, INRODUCTION TO MATHEMATICAL FINANCE, Blackwell, 1999.

Grading Scheme

Homework 40 % Take Home Midterms 40% Final 40%

Contents

Chapter 1

Mean-Variance Portfolio Theory

Typically, when making an investment, the initial outlay of capital is known, but the amount to be returned is uncertain, and one makes efforts to minimize the uncertainties. Such situation is studied in this part of text. We shall restrict attention to the case of a single investment period: money is invested at the initial time and payoff is attained at the end of the period. The uncertainty is treated by **mean-variance analysis**, developed by Nobel prize winner Markowitz. This method leads to convenient mathematical expressions and procedures, and forms the basis for the more important **capital asset pricing model**.

1.1 Assets and Portfolios

An **asset** is *an investment instrument that can be bought and sold.* Its **return** is *the percentage of value increased from time bought to time sold.* By **return rate** it means return per unit time.

A **portfolio** is *a collection of shares of assets. The proportions in value of assets in a portfolio* are called the **weights**. A **portfolio's return** is *the percentage of value increased from time bought to time sold.*

In this chapter, we study one period investment and take the period as unit time, so return and return rate are interchangeable.

Example 1.1. (1) With \$10,000 cash Jesse bought 100 shares Stone Inc. stock at \$100 per share at the beginning of a period. She hold the stock for one period and sold the stock at \$105.00 per share, ending up with \$10,500 cash.

Assume that during the period, the stock did not pay any dividend and there was no transaction cost. Then, she made a profit of \$10,500-\$10,000=\$500 from her \$10,000 investment. Thus, the return is

$$\frac{\text{payment} - \text{investment}}{\text{investment}} = \frac{\$10,500 - \$10,000}{\$10,000} = 5\%.$$

The return rate is 5% per period.

(2) Similarly, suppose John spent $\$5,000$ bought 100 shares of Rock Inc. stock at \$50 per share at the beginning the same period as Jesse and sold all his stocks at \$55 per share at the end of period, with no dividend received during the period. Then John ended up with \$5500 cash, making a profit of \$500 with a \$5000 capital investment. The return of his investment is \$500/\$5000=10%.

(3) Consider a hypothetical investment. Suppose John is a trusted friend of Jesse and promised to take care of Jesse's investment. So at the beginning of the period, John received $10,000 cash from Jesse who instructed John to make investment on her behalf on a one period investment on Stone's stock. By this, Jesse means John has to give her the cash price of 100 shares of Stone Inc. stock at the end of the period.

John has his own cash $5000 at the beginning of the period. With Jesse's $10,000, he now has $15,000 cash. Instead of buying 100 share of Stone' Inc stock on Jesse's behalf, John bought 300 shares of Rock Inc. stock at $50 per share. By doing so, John means that he will go to the market buy the stock for Jesse whenever she wants them.

(a) Suppose at the end of period, the Rock Inc. stock unit share price is $55 and Stone Inc. stock price is $105. After selling all his Rock Inc. stock holding, John obtains $300 * \$55 = \$16500$ cash.

    Now Jesse asks John to pay her the payment of her investment, totalling $100*$105=$10,500. after the payment, John now has $16500-$10500= $6000 cash left.

    In her investment Jesse made a profit of $500 with $10,000, as she would have done herself.

    On the other hand, John made a profit of $6000-$5000=$1000 out of $5000 investment. Hence his return is $1000/$5000=20%.

(b) Suppose at the end of period, Stone Inc's stock price is $110 and Rock Inc. stock price is $53.

    Then after cashing in the 300 shares of Rock Inc stock, John has $15900 cash. But he has the obligation to pay Jesse 100*$110=$11000 the payment of her investment. Upon doing so, John ends up with $4900 cash.

    In this investment, Jesse made a profit of $1000 with an $10,000 investment, so the return is 10%.

    However, John lost $100 with an $5000 investment. Hence, his return is -2%.

In this example, John's action on Jesse's request is known as short selling: He takes in the cash and owes certain shares of the named stock. Typically, one shorts with dealers instead of with friends, and dealers charge a certain amount of extra fees. In this course, we shall assume that not only there is no transaction cost, but also there is no extra charge on short selling.

Example 1.2. The following table illustrates a typical example of a portfolio:

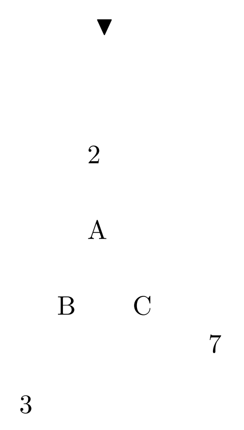
assets (security)	Number of shares	unit price	cost	portfolio weight	return (rate)	total return	weighted return	new portfolio weight
Rock Inc.	200	$20	$4,000	0.40	10%	$4400	0.4*10%	4400/11150
Jazz Inc.	300	$30	$9,000	0.90	10 %	$9900	0.90*10%	9900/11150
Stone Inc.	-100	$30	-$3,000	-0.30	5%	-$3150	-0.30*5%	-3150/11150
Portfolio Total			$10,000	1.00		$11150	11.5 %	1

In this example, the assets (also called securities) in consideration are stocks of three companies. Initially, the investor has a total of $10,000 cash available. By short selling, e.g. borrowing 100 shares of Stone Inc.'s stock, selling it to generate cash for other stocks, and then returning the borrowed stock at the end of period, the investor is lucky enough to make an 11.5% return. Here, we assume that selling and buying are symmetric, no extra charges are accounted. Of course, if the investor made a wrong judgement by short selling 100 shares of Jazz Inc's stock to generating cash buying 300 shares of Stone Inc.'s stock, the final return would be $0.4 * 10\% - 0.3 * 10\% + 0.9 * 5\% = 5.5\%$; that is, the final wealth would be \$10,550.

Note that the weight changes at the end of period. For a multiple period investment, one may consider adjusting the weights from time to time.

Example 1.3. Investments have risks. This is the same as gambling. Here we illustrate such an aspect by using an investment wheel.

You are able to place a bet on any of the three sectors, named A, B and C respectively. In fact, you may invest different amounts on each of sectors independently. The numbers in sectors denote the winnings (multiplicative factor to your bet) for that sector after the wheel is spun. For example, if the wheel stops with the pointer at the top sector A after a spin, you will receive \$2 for every \$1 you invested on that sector (which means a net profit of \$1); all bets on other sectors are lost.



An Investment Wheel

Let's use $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ to denote the investment plan by place \$1.00 bet on sectors A, B and C respectively.

Denote $\Omega = \{A, B, C\}$ the space all possible events and by $\text{Prob}(x)$, the probability that event $x \in \Omega$ occurs. We have

$$\text{Prob}(A) = \frac{1}{2}, \qquad \text{Prob}(B) = \frac{1}{3}, \qquad \text{Prob}(C) = \frac{1}{6}.$$

The return R of an investment depends on the actual event that occurs. Mathematically R is a random variable, i.e. a measurable function from Ω to $\mathbb{R}$. Denote by $R_{\mathcal{A}}, R_{\mathcal{B}}, R_{\mathcal{C}}$ the returns of the investment plan $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, respectively. Then they are functions from Ω to $\mathbb{R}$ valued as follows:

$$R_{\mathcal{A}}(A) = 100\%, \quad R_{\mathcal{A}}(B) = -100\%, \quad R_{\mathcal{A}}(C) = -100\%;$$
$$R_{\mathcal{B}}(A) = -100\%, \quad R_{\mathcal{B}}(B) = 200\%, \quad R_{\mathcal{B}}(C) = -100\%,$$
$$R_{\mathcal{C}}(A) = -100\%, \quad R_{\mathcal{C}}(B) = -100\%, \quad R_{\mathcal{C}}(C) = 600\%;$$

For random variables, the most often used quantities are mean (expectation $\mathbf{E}[\cdot]$), variance, and covariance. For random variables ξ and η on a finite probability space Ω,

$$\begin{aligned}
\text{mean of } \xi &= \mathbf{E}[\xi] := \sum_{x\in\Omega} \xi(x)\text{Prob}(x), \\
\text{variance of } \xi &= \mathbf{Var}[\xi] := \mathbf{E}\Big[(\xi - \mathbf{E}[\xi])^2\Big] = \mathbf{E}[\xi^2] - \mathbf{E}[\xi]^2, \\
\text{standard derivation of } \xi &= \sqrt{\mathbf{Var}[\xi]} \\
\text{covariance between } \xi \text{ and } \eta &= \mathbf{Cov}[\xi,\eta] := \mathbf{E}[(\xi - \mathbf{E}[\xi])(\eta - \mathbf{E}[\eta])] = \mathbf{E}[\xi\eta] - \mathbf{E}[\xi]\mathbf{E}[\eta], \\
\text{corretion between } \xi \text{ and } \eta &= \mathbf{cor}[\xi,\eta] := \frac{\mathbf{Cov}[\xi,\eta]}{\sqrt{\mathbf{Var}[\xi]\mathbf{Var}[\xi]}} \in [-1,1].
\end{aligned}$$

Hence, the mean return of investment $\mathcal{A}$ is

$$\mu_{\mathcal{A}} := \mathbf{E}[R_{\mathcal{A}}] = \sum_{x\in\Omega} R_{\mathcal{A}}(x)\text{Prob}(x) = 100\% * \tfrac{1}{2} - 100\% * \tfrac{1}{3} - 100\% * \tfrac{1}{6} = 0\%.$$

The variance of the return $R_{\mathcal{A}}$ is

$$\sigma_{\mathcal{A}\mathcal{A}} = \mathbf{Var}[R_{\mathcal{A}}] := \mathbf{E}\Big[(R_{\mathcal{A}} - \mathbf{E}[R_{\mathcal{A}}])^2\Big] = \mathbf{E}[R_{\mathcal{A}}^2] - \mathbf{E}[R_{\mathcal{A}}]^2 = 1.$$

The standard deviation of $R_{\mathcal{A}}$ is

$$\sigma_{\mathcal{A}} = \sqrt{\sigma_{\mathcal{A}\mathcal{A}}} = \sqrt{\mathbf{Var}[R_{\mathcal{A}}]} = 1 = 100\%.$$

The covariance between $R_{\mathcal{A}}$ and $R_{\mathcal{B}}$ is

$$\sigma_{\mathcal{A}\mathcal{B}} = \mathbf{E}\Big[(R_{\mathcal{A}} - \mathbf{E}(R_{\mathcal{A}})(R_{\mathcal{B}} - \mathbf{E}[R_{\mathcal{B}}]\Big] = \sum_{x\in\Omega} R_{\mathcal{A}}(x)R_{\mathcal{B}}(x)\text{Prob}(x) - \mathbf{E}[R_{\mathcal{A}}]\mathbf{E}[R_{\mathcal{B}}] = -1.$$

The correlation between R_1 and R_2 is

$$\rho_{\mathcal{A}\mathcal{B}} = \frac{\sigma_{\mathcal{A}\mathcal{B}}}{\sigma_{\mathcal{A}}\sigma_{\mathcal{B}}} = -\frac{1}{\sqrt{2}} = -0.707.$$

Similarly, we can calculate other statistical quantities. The result is summarized in the following tables.

Investment Plan	Return Under A	B	C	Mean return	Covariance σ_{ij} $\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$	Correlation ρ_{ij} $\mathcal{A}$	$\mathcal{B}$	$\mathcal{C}$
$\mathcal{A}$	100%	-100%	-100%	0%	1	-1	-1.17	1.00	-0.707	-0.447
$\mathcal{B}$	-100%	200%	-100%	0 %	-1	2	-1.17	-0.707	1	-0.316
$\mathcal{C}$	-100%	-100%	600%	17%	-1.17	-1.177	6.81	-0.447	-0.316	1
Probability	1/2	1/3	1/6							

We now consider a market system consisting of m assets, named $\mathsf{a}_1, \cdots, \mathsf{a}_m$. Denote by R_i the return of asset a_i. Then

$$1 + R_i = \frac{\text{value of unit asset } \mathsf{a}_i \text{ at time sold}}{\text{initial value of unit asset } \mathsf{a}_i}.$$

The basic assumption here is that R_i is a random variable, with mean μ_i and variance σ_i^2:

$$\mu_i = \mathbf{E}[R_i], \qquad \sigma_i = \sqrt{\mathbf{Var}[R_i]}.$$

We call μ_i the **expected return** and in the current context σ_i (or σ_i^2) the **risk** of the asset a_i. Also we denote the covariance and correlation between the returns of the asset a_i and a_j by

$$\sigma_{ij} := \mathbf{Cov}(R_i, R_j) := \mathbf{E}((R_i - \mu_i)(R_j - \mu_j)), \quad \sigma_{ii} = \sigma_i^2, \quad \rho_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}.$$

We now consider a portfolio that consists of a collection of the above assets. Since the sizes of units of these assets are quite different, we shall not pay any attention on the particular numbers of units, rather, we are concerned about the percentage of each asset value in the portfolio.

Suppose the total value of the portfolio is V_0 and the value in asset a_i is V_i, $i = 1, \cdots, m$. Then the weight of the asset a_i in the portfolio is

$$w_i = \frac{\text{value in asset } \mathsf{a}_i}{\text{total value of portfolio}} = \frac{V_i}{V_0}.$$

We denote the portfolio's weight by the row vector

$$\mathbf{w} = (w_1, w_2, \cdots, w_m).$$

Then

$$\sum_{i=1}^m w_i = \sum_{i=1}^m \frac{V_i}{V_0} = \frac{\sum_{i=1}^m V_i}{V_0} = 1.$$

In general, the weight is a function of time, since the returns of different assets are different. In this chapter, we shall consider only two times, the time when the portfolio is bought and the time when it is sold.

Denote by R the portfolio's return:

$$R := \frac{\text{portfolio value at time sold} - \text{initial portfolio value}}{\text{initial portfolio value}} \quad .$$

A simple arithmetic gives the relation among portfolio return, asserts return and weight:

$$R = \sum_{i=1}^m w_i R_i. \tag{1.1}$$

The **expected return** μ and **risk** σ ($\sigma \geqslant 0$) of the portfolio can be calculated by

$$\mu = \mathbf{E}[R] = \mathbf{E}\Big[\sum_{i=1}^m w_i R_i\Big] = \sum_{i=1}^m w_i \mu_i, \tag{1.2}$$

$$\sigma^2 = \mathbf{Var}[R] = \mathbf{Var}\Big[\sum_{i=1}^m w_i R_i\Big] = \sum_{i=1}^m \sum_{j=1}^m \sigma_{ij} w_i w_j. \tag{1.3}$$

We shall assume that $\mathbf{u} := (\mu_1, \cdots, \mu_m)$ and $\mathbf{C} := (\sigma_{ij})_{m\times m}$ are known; that is, they can be calculated from historical data. Thus, the problem here is to choose appropriate weights $\mathbf{w} = (w_1, \cdots, w_m)$

which satisfies the constraint $\sum_{i=1}^{m} w_i = 1$. By varying the weight, one obtains different portfolios of different **risk-return** balances. There are people who are willing to take high risk expecting high returns, whereas there are also people who want security thus are willing to accept moderate returns with small risks.

Mathematically, we are going to find optimal weights that minimizes risk with given expected return or maximizes the expect return with given risk. These two problems are dual to each other.

Since μ is a linear and σ is a quadratic function of the weights, as one shall see, the problem can be solved explicitly. For the convenience of presentation, we shall assume that the market is **fair** in the sense that

any weight $\mathbf{w} \in \{(w_1, \cdots, w_m) \in \mathbb{R}^m \mid \sum_{i=1}^{m} w_i = 1\}$ *is attainable.*

Suppose the total investment is V_0. When a weight w_i is positive, it means to buy (long) asset a_i certain units worth $V_0 w_i$. When $w_i < 0$, it mean selling (short) the asset certain units to generate $V_0|w_i|$ cash that can be used to buy other assets. By doing that, one owes certain shares of assets a_i which has to be paid back, with the same amount of units, at time the portfolio is sold[1]

Example 1.4. Consider the three investment plans, $\mathcal{A}, \mathcal{B}, \mathcal{C}$, in Example 1.3. With a total capital $V_0 = \$50$, consider the following investment:

Put \$10 on sector A, \$10 on sector B, and \$30 on sector C.

Denote by $V_T(x)$ the value of the portfolio at the end of investment under event $x \in \Omega = \{A, B, C\}$. Then

$$V_T(A) = \$20, \qquad V_T(B) = \$30, \qquad V_T(C) = \$210.$$

Hence, denote by $R(x)$ the return under event $x \in \Omega = \{A, B, C\}$. It is easy to see

$$R(A) = \frac{\$20}{\$50} - 1 = -60\%, \qquad R(B) = \frac{\$30}{\$50} - 1 = -40\%, \qquad R(C) = \frac{\$210}{\$50} - 1 = 320\%.$$

The mean return is

$$\mu = \mathbf{E}[R] = \sum_{x \in \Omega} R(x) \ \mathrm{Prob}(x) = -60\% * \tfrac{1}{2} - 40\% * \tfrac{1}{3} + 320\% + \tfrac{1}{6} = 10\%.$$

The risk is

$$\sigma = \sqrt{\mathbf{Var}[R]} = \sqrt{\sum_{x \in \Omega} (R(x) - \mu)^2 \mathrm{Prob}(x)} = 139\%.$$

Portfolios with only a few assets may be subject to a high degree of risk, represented by a relatively large variance. As a general rule, the variance of the return of a portfolio can be reduced by including additional assets in the portfolio, a process referred to as **diversification**. This process reflects the maxim:

Don't put all yours eggs in one basket.

[1]When a stock pays dividend, typically one has the choice of receiving cash or a percentage of share of stock equivalent to the cash. In such scenario, number of units of to be returned from shorting will be larger than the number that one initially shorts. Similarly, it is very common that stock splits; namely, one share becomes two share; in such case, one of course has to pay double number of units.

Example 1.5. Consider the following simple yet illustrative situation. Suppose there are m assets each of which has return $\hat{\mu}$ and variance $\hat{\sigma}^2$. Suppose also that all these assets are mutually uncorrelated. One then construct a portfolio by investing equally into these assets, namely, taking $w_i = 1/m$ for all $i = 1, \cdots, m$. The overall expected rate of return is still $\hat{\mu}$. Nevertheless, the overall risk becomes

$$\mathbf{Var}[R] = \sum_{i=1}^{m}\sum_{j=1}^{m} \frac{1}{m}\frac{1}{m}\hat{\sigma}^2\delta_{ij} = \frac{\hat{\sigma}^2}{m},$$

which decays rapidly as m increases. The situation is different when returns of the available assets are correlated; see exercise 1.4.

Example 1.6. Consider a portfolio of two assets, $\mathsf{a}_1, \mathsf{a}_2$, with the following statistical parameters:

$$\mu_1 = 5\%, \quad \mu_2 = 10\%, \quad \sigma_1 = 10\%, \quad \sigma_2 = 40\%, \quad \rho_{12} = -0.5 \ .$$

The weight of an arbitrary portfolio can be denoted as $\mathbf{w} = (\theta, 1-\theta)$. Denote the return of such portfolio by $R(\theta)$. We have

$$\mu(\theta) := \mathbf{E}(R(\theta) = \theta\mu_1 + (1-\theta)\mu_2 = 0.1 - 0.05\theta.$$

Hence to have a portfolio of wanted expected return μ, we need only take θ such that $\mu = 0.1 - 0.05\theta$, i.e.

$$\theta = (0.1 - \mu)/0.05 = 2 - 20\mu.$$

Also, the variance of this portfolio is

$$\sigma(\theta)^2 := \mathbf{Var}(R(\theta)) = \sigma_1^2\theta^2 + 2\rho_{12}\sigma_1\sigma_2\theta(1-\theta) + \sigma_2^2(1-\theta)^2 = 0.16 - 0.36\theta + 0.21\theta^2.$$

To see a direct relation between the expected return $\mu = \mu(\theta)$ and the risk $\sigma = \sigma(\theta)$, we substitute $\theta = 2 - 20\mu$ in the above expresion, obtaining

$$\sigma = \sqrt{0.16 - 0.36\theta + 0.21\theta^2}\Big|_{\theta=0.1-20\mu} = \sqrt{0.076^2 + 84(\mu - 5.7\%)^2}.$$

The relation between μ and σ is depicted in Figure 1.1. Among all the portfolios, the one that has the minimum risk is

$$\theta = 0.86, \qquad \mu = 5.7\%, \qquad \sigma = 7.6\%.$$

Clear, such a mutual fund, with 86% capital on the first asset a_1 and 14% capital on the second asset a_2 is much better than a_1 alone, both in the expected return and in the risk.

Also, consider the portfolio $\mathbf{w} = (-1, 2)$; i.e. $\theta = -1$. Then on finds that the return and risk are

$$\mu = 15\%, \qquad \sigma = 85\%.$$

Here the large expected return $\mu = 15\%$ is obtained under the large risk $\sigma = 85\%$.

Exercise 1.1. *(a) Derive and illustrate with Examples 1.2 and Example 1.4 the formulas* (1.1)–(1.3).

(b) In a portfolio, the number of shares of each asset is assumed to be constant in the time period of our consideration. As the price of unit share changes, so is the relative proportion of values of each asset in the portfolio.

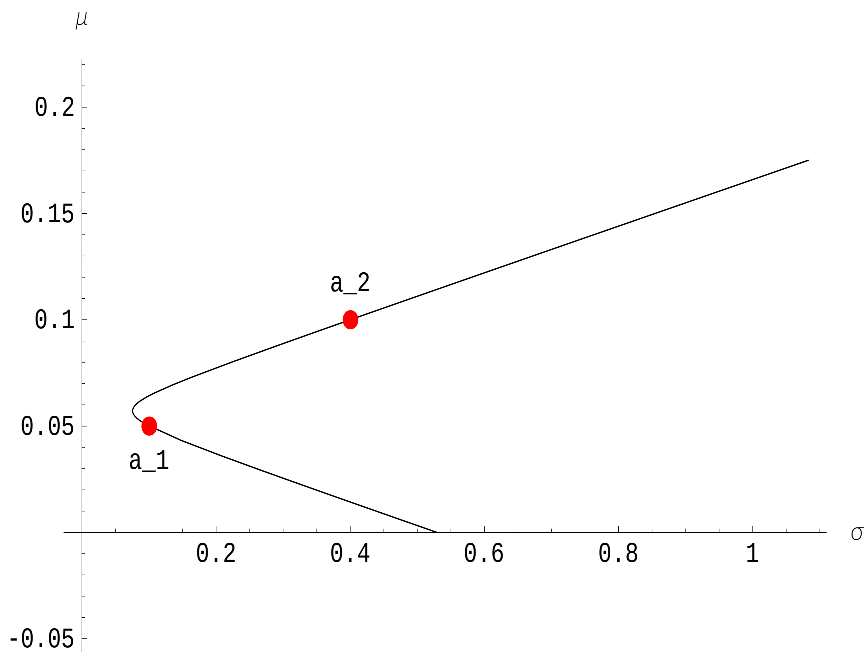


Figure 1.1: The risk-return curve. The two dots correspond to the asset a_1 and a_2 respectively.

(i) Suppose in Example 1.2, one holds the same number of shares of stock for the second period. Find the weights at the beginning of the investment for the second period.

(ii) Denote by $\mathbf{w}(t) = (w_1(t), \cdots, w_m(t))$ the weight of portfolio at time t. Show that at the end of first period, the new weight becomes

$$w_i(1-0) = \frac{(1+R_i)w_i(0)}{1+R} \qquad \forall\, i = 1, \cdots, m.$$

Exercise 1.2. *Consider the investment opportunities $\mathcal{A}, \mathcal{B}, \mathcal{C}$ in Example 1.3. Consider the following mutual funds:*

(F1) \$10 in $\mathcal{A}$, \$10 in $\mathcal{B}$, and \$10 in $\mathcal{C}$;

(F2) \$21 in $\mathcal{A}$, \$14 in $\mathcal{B}$, \$6 in $\mathcal{C}$;

(F3) x in $\mathcal{A}$, y in $\mathcal{B}$, $1-x-y$ in $\mathcal{C}$.

Find the mean return μ and risk σ of each mutual funds.

Exercise 1.3. *Calculate the statistics as in Example 1.3 for an investment wheel where multiplicative for A, B, and C are $3, 3, 7$ respectively, keeping the same probability of the occurrence of A,B and C.*

Exercise 1.4. *Suppose $\mathbf{Cov}(R_i, R_j) = 0.3\hat{\sigma}^2$ for all $i \neq j$ and $\mathbf{Var}(R_i) = \hat{\sigma}^2$. Calculate the risk of the following portfolios:*

(i) $w_i = 1/m$ for all $i = 1, \cdots, m$;

(ii) $w = 3/m$ for all $i = 1, \cdots, m/2$ and $w_i = -1/m$ for $i = m/2+1, \cdots, m$. (Assume m is even)

Exercise 1.5. *Consider a portfolio of two assets. Write $\mathbf{w} = (\theta, 1-\theta)$, $\mu = \mu(\theta), \sigma = \sigma(\theta)$.*

(a) For each of the cases when the correlation coefficient ρ_{12} *is* $1, 1/2, 0, -1/2, -1$*, plot the curve* $\sigma(\theta)$*. Also plot the curve* σ *against* μ*, taking (i)* $\mu_1 = 0.1, \mu_2 = 0.2, \sigma_1 = 0.2, \sigma_2 = 0.3$*, (ii)* $\mu_1 = 0.1, \mu_2 = 0.2, \sigma_1 = 0.3, \sigma_2 = 0.2$*.*

(b) Find a portfolio that has the minimum risk possible.

(c) Find a portfolio that has the minimum risk possible, where short selling is forbidden.

Exercise 1.6. *Suppose short selling is unlimited and consider a system of two assets with* $\mu_1 > \mu_2$ *and* $\sigma_1 = \sigma_2 > 0$*. Show that one can make money out of nothing if and only if* $\rho_{12} = 1$*.*

Exercise 1.7. *For a system of two and three assets respectively, find the portfolios that have minimum risk under condition (i) shorting selling is allowed (ii) short selling is forbidden. Assume the covariance matrix* (σ_{ij}) *is known.*

1.2 The Markowitz Portfolio Theory

It is reasonable to assume that not all asset returns are the same. Since if all the returns are the same, the expected return of the portfolio does not change with the weights. As a consequence, the problem becomes the study of risks alone; see exercise 1.7.

The covariance matrix $\mathbf{C} = (\sigma_{ij})_{m\times m}$ is symmetric and semi-positive-definite. For simplicity, we assume that it is invertible so it is positive definite.

A portfolio is called **efficient** if *its risk is no larger than any other portfolio of the same expected return.* The Markowitz portfolio theory is to find all efficient portfolios. Mathematically, the problem can be formulated as follows:

Efficient Portfolio Problem: Given $\mu \in \mathbb{R}$, find a portfolio $\mathbf{w} = (w_1, \cdots, w_m) \in \mathbb{R}^m$ that

$$\text{minimizes} \quad \mathbf{Var}[R] = \sum_{i=1}^{m}\sum_{j=1}^{m} w_i w_j \sigma_{ij} \qquad \text{subject to} \quad \sum_{i=1}^{m} w_i = 1, \quad \mathbf{E}[R] = \sum_{i=1}^{m} \mu_i w_i = \mu.$$

The solution. This problem can be solved by using the Lagrange multipliers. Thus, we consider the unconditional critical points of the functional

$$L(\lambda_1, \lambda_2, \mathbf{w}) = \sum_{i=1}^{m}\sum_{j=1}^{m} w_i w_j \sigma_{ij} + \lambda_1\Big(1 - \sum_{i=1}^{m} w_i\Big) + \lambda_2\Big(\mu - \sum_{i=1}^{m} w_i\mu_i\Big), \quad (\lambda_1, \lambda_2, \mathbf{w}) \in \mathbb{R}^{m+2}.$$

The system of equations for critical points of L is

$$\frac{\partial L}{\partial \lambda_1} = 0, \quad \frac{\partial L}{\partial \lambda_2} = 0, \quad \frac{\partial L}{\partial w_k} = 0 \;\forall\, k = 1, \cdots, m.$$

The first two equations give the constraint conditions whereas the remaining equations are

$$0 = \frac{\partial L}{\partial w_k} = \sum_{i=1}^{m} (\sigma_{ik} + \sigma_{ki}) w_i - \lambda_1 - \lambda_1 \mu_k, \qquad k = 1, \cdots, m.$$

These equations can be written in the matrix form as $2\mathbf{wC} + \lambda\mathbf{1} + \lambda_2\mathbf{u} = \mathbf{0}$ where $\mathbf{0} = 0\mathbf{1}$ and

$$\mathbf{1} = (1, \cdots, 1)_{1\times m}, \qquad \mathbf{u} = (\mu_1, \cdots, \mu_m), \qquad \mathbf{C} = (\sigma_{ij})_{m\times m}.$$

Here we identify a row vector with a row matrix. Using $(\cdot,\cdot)$ for $\mathbb{R}^m$ dot product, we then have

$$\mathbf{w} = \frac{\lambda_1 \mathbf{1}\mathbf{C}^{-1}}{2} + \frac{\lambda_2 \mathbf{u}\mathbf{C}^{-1}}{2} = \theta \mathbf{w}_1 + (1-\theta)\mathbf{w}_2,$$

where $\theta = \lambda_1(\mathbf{1}, \mathbf{1}\mathbf{C}^{-1})/2$ and $\mathbf{w}_1, \mathbf{w}_2$ are weights of two portfolios given by

$$\mathbf{w}_1 := \frac{\mathbf{1}\mathbf{C}^{-1}}{(\mathbf{1}, \mathbf{1}\mathbf{C}^{-1})}, \qquad \mathbf{w}_2 := \frac{\mathbf{u}\mathbf{C}^{-1}}{(\mathbf{1}, \mathbf{u}\mathbf{C}^{-1})}.$$

Here the proportion $(1-\theta)$ is obtained by using $(\mathbf{1}, \mathbf{w}) = 1$. Note that the portfolio $\mathbf{w}$ consists of θ portion of portfolio $\mathbf{w}_1$ and $(1-\theta)$ portion of portfolio $\mathbf{w}_2$.

Hence, substituting the expression for $\mathbf{w}$ into the constraint $(\mathbf{w}, \mathbf{u}) = \mu$ we obtain the value of θ. After substituting it back into the expression for $\mathbf{w}$ we then find the solution to the efficient portfolio problem to be

$$\mathbf{w} = \mathbf{e}_1 + \mu \mathbf{e}_2,$$

where

$$\mathbf{e}_1 := \frac{(\mathbf{w}_2, \mathbf{u})\, \mathbf{w}_1 - (\mathbf{w}_1, \mathbf{u})\, \mathbf{w}_2}{(\mathbf{w}_2, \mathbf{u}) - (\mathbf{w}_1, \mathbf{u})}, \qquad \mathbf{e}_2 := \frac{\mathbf{w}_2 - \mathbf{w}_1}{(\mathbf{w}_2, \mathbf{u}) - (\mathbf{w}_1, \mathbf{u})}\ .$$

The Rigorous Analysis. While the method of Lagrange multiplier is powerful enough to provide needed solutions, it does not necessarily always provide the correct answer. Verification of the solution is often needed. Hence, here we provide a rigorous analysis, showing that the solution we obtained is indeed the unique solution to the conditional minimization problem. We use the same notation $\mathbf{1}, \mathbf{u}, \mathbf{C}, \mathbf{w}_1, \mathbf{w}_2, \mathbf{e}_1, \mathbf{e}_2$ as before.

Let $\mathbf{w} = (w_1, \cdots, w_m)$ be any weight satisfying $\sum_{i=1}^m w_i = 1$ and $\sum_{i=1}^m u_i w_i = \mu$, i.e. $(\mathbf{1}, \mathbf{w}) = 1, (\mathbf{u}, \mathbf{w}) = \mu$. Consider the vector

$$\mathbf{w}_\perp := \mathbf{w} - \mathbf{e}_1 - \mu \mathbf{e}_2.$$

We find that $(\mathbf{w}_\perp, \mathbf{1}) = (\mathbf{w}, \mathbf{1}) - (\mathbf{e}_1, \mathbf{1}) - (\mathbf{e}_2, \mathbf{1}) = 1 - 1 - 0 = 0$ and $(\mathbf{w}_\perp, \mathbf{u}) = (\mathbf{w}, \mathbf{u}) - (\mathbf{e}_1, \mathbf{u}) - \mu(\mathbf{e}_2, \mathbf{u}) = \mu - 0 - \mu = 0$. That is $\mathbf{w}_\perp \perp \mathbf{1}, \mathbf{w}_\perp \perp \mathbf{u}$. Write $\mathbf{w} = \mathbf{e}_1 + \mu \mathbf{e}_2 + \mathbf{w}_\perp$. Note that $\mathbf{C}$ is symmetric and both $\mathbf{e}_1\mathbf{C}$ and $\mathbf{e}_2\mathbf{C}$ are linear combinations of $\mathbf{1}$ and $\mathbf{u}$, we have $(\mathbf{e}_1\mathbf{C}, \mathbf{w}_\perp) = 0$ and $(\mathbf{e}_2\mathbf{C}, \mathbf{w}_\perp) = 0$. Hence,

$$\begin{aligned}
\sigma^2 &= (\mathbf{w}\mathbf{C}, \mathbf{w}) \\
&= (\mathbf{w}_\perp\mathbf{C}, \mathbf{w}_\perp) + (\mathbf{e}_1\mathbf{C}, \mathbf{e}_1) + 2\mu(\mathbf{e}_1\mathbf{C}, \mathbf{e}_2) + \mu^2(\mathbf{e}_2\mathbf{C}, \mathbf{e}_2) \\
&= (\mathbf{w}_\perp\mathbf{C}, \mathbf{w}_\perp) + (\mathbf{e}_2\mathbf{C}, \mathbf{e}_2)\Big(\mu + \frac{(\mathbf{e}_1\mathbf{C}, \mathbf{e}_2)}{(\mathbf{e}_2\mathbf{C}, \mathbf{e}_2)}\Big)^2 + (\mathbf{e}_1\mathbf{C}, \mathbf{e}_1) - \frac{(\mathbf{e}_1\mathbf{C}, \mathbf{e}_2)^2}{(\mathbf{e}_2\mathbf{C}, \mathbf{e}_2)} \\
&= (\mathbf{w}_\perp\mathbf{C}, \mathbf{w}_\perp) + \sigma_*^2 + \kappa^2(\mu - \mu_*)^2
\end{aligned}$$

where, by the definition of $\mathbf{e}_1, \mathbf{e}_2$,

$$\begin{aligned}
\kappa^2 &= (\mathbf{e}_2\mathbf{C}, \mathbf{e}_2) = \frac{(\mathbf{1}\mathbf{C}^{-1}, \mathbf{1})}{(\mathbf{1}\mathbf{C}^{-1}, \mathbf{1})(\mathbf{u}\mathbf{C}^{-1}, \mathbf{u}) - (\mathbf{1}\mathbf{C}^{-1}, \mathbf{u})^2}, \\
\sigma_*^2 &= (\mathbf{e}_1\mathbf{C}, \mathbf{e}_1) - \frac{(\mathbf{e}_1\mathbf{C}, \mathbf{e}_2)^2}{(\mathbf{e}_2\mathbf{C}, \mathbf{e}_2)} = \frac{1}{(\mathbf{1}\mathbf{C}^{-1}, \mathbf{1})}, \\
\mu_* &= -\frac{(\mathbf{e}_1\mathbf{C}, \mathbf{e}_2)}{(\mathbf{e}_2\mathbf{C}, \mathbf{e}_2)} = \frac{(\mathbf{u}\mathbf{C}^{-1}, \mathbf{1})}{(\mathbf{1}\mathbf{C}^{-1}, \mathbf{1})}.
\end{aligned}$$

We remark that $\kappa > 0$ and $\sigma_* > 0$ since $\mathbf{C}$ positive definite implies $\mathbf{C}^{-1}$ is also positive definite. Hence we have the following

Theorem 1.1 *Assume that not all asset's expected returns are equal and* $\mathbf{C} = (\sigma_{ij})$ *is positive definite.*

(i) For every weight $\mathbf{w}$ *with expected return* μ*, its risk* σ^2 *satisfies*

$$\sigma \geqslant \sqrt{\sigma_*^2 + \kappa^2(\mu - \mu_*)^2}.$$

(ii) The equality in the above inequality is attained at and only at **minimum risk weight line**

$$\mathbf{w} = \mathbf{e}_1 + \mu\mathbf{e}_2, \quad i.e. \quad \mathbf{w} = \theta\mathbf{w}_1 + (1-\theta)\mathbf{w}_2. \tag{1.4}$$

We note that on the σ-μ plane, the curve

$$\sigma^2 = \sigma_*^2 + \kappa^2(\mu - \mu_*)^2, \qquad \sigma > 0, \tag{1.5}$$

or

$$\mu = \mu_* \pm \frac{1}{\kappa}\sqrt{\sigma^2 - \sigma_*^2}, \quad \sigma > 0, \tag{1.6}$$

is a hyperbola with tip at (σ_*, μ_*); see Figure 1.2. The hyperbola is called the **Markowitz curve**. A portfolio is **efficient** if and only if its expected return and standard deviation is on the Markowitz curve. The unbounded region on the right-hand side of the hyperbola is called the **Markowitz bullet** or **attainable region**; the top half of the hyperbola is called the **Markowitz efficient frontier**.

(a) For any expected return μ, the attainable risk is an interval $[\sigma^2, \infty)$ where (μ, σ) is on the Markowitz curve. That is, fixing any expected return, the minimum risk is given by (1.5) with weight given by (1.4).

(b) The positive number σ_*^2 is the absolute minimum risk among all weights, i.e. there is no weight that can provide a risk smaller than σ^*. For any chosen risk $\sigma \geqslant \sigma_*$, the attainable expected return μ is an interval centered at μ_* with maximum on the Markowitz efficient frontier.

(c) Any attainable point is dominated by an attainable point on the Markowitz efficient frontier. Investors who seek to minimize risk for any expected return need only look on the Markowitz efficient frontier, that is, for efficient portfolios, whose weight are given by the minimum risk weight line, being a linear combinations of two special weights.

Theorem 1.2 **(Two-Fund Theorem)** *Two efficient funds (portfolios) can be established so that any efficient portfolio can be duplicated, in terms of mean and variance, as a combination of these two. That is, all investors seeking efficient portfolio need only invest in combinations of these funds.*

This result has dramatic implications. According to the two-fund theorem, two **mutual funds** (for example, portfolios with weights $\mathbf{w}_1$ and $\mathbf{w}_2$ respectively) could provide a complete investment service for everyone. There would be no need for everyone to purchase individual stocks separately; they could just purchase shares in the two mutual funds.

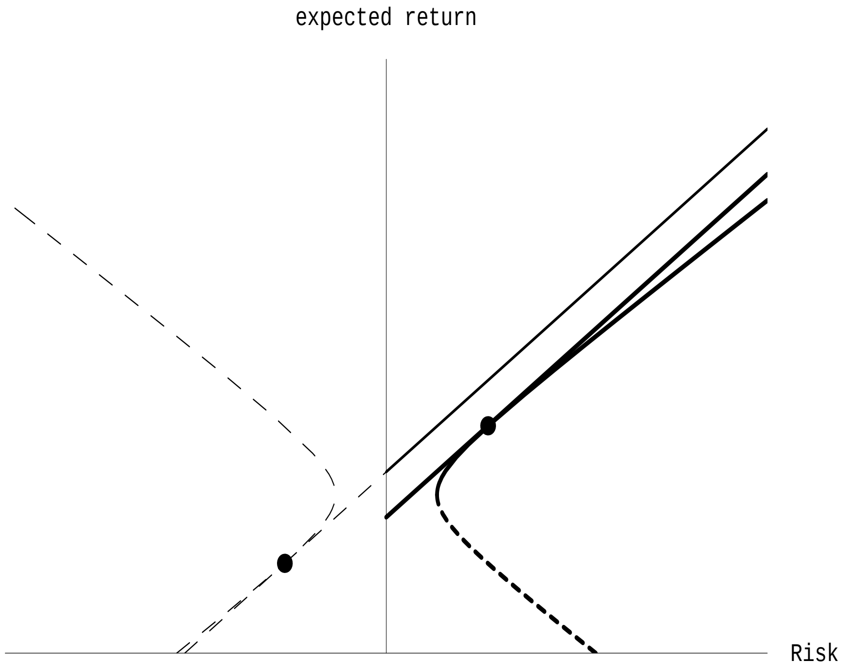


Figure 1.2: Thick hyperbola is the Markowitz frontier, where dashed thick curve is the remaining part of the Markowitz curve. The thick tangent line is the capital market line when risk-free rate μ_0 is less that μ_*. The thin tangent line is an analogous of the capital market line when $\mu_0 > \mu_*$.

Example 1.7. Consider portfolios of three assets with the following statistics:

		Cov(R_i, R_j)		
Assets	Mean Return	a$_1$	a$_2$	a$_3$
a$_1$	0.08	0.02	-0.01	-0.02
a$_2$	0.08	-0.01	0.04	0.01
a$_3$	0.12	-0.02	0.01	0.09

Then we have

$$\mathbf{u} = (0.08\ 0.08\ 0.12), \quad \mathbf{C} = \begin{pmatrix} 0.08 & 0.02 & -0.01 \\ -0.01 & 0.04 & 0.01 \\ -0.02 & 0.01 & 0.09 \end{pmatrix}, \quad \mathbf{C}^{-1} = \begin{pmatrix} 71.4 & 14.3 & 14.3 \\ 14.3 & 28.6 & 0 \\ 14.3 & 0 & 14.3 \end{pmatrix}.$$

Consequently, denoting by $*$ the matrix transpose, we obtain

$$\mathbf{w}_1 = \frac{\mathbf{1C}^{-1}}{\mathbf{1C}^{-1}\mathbf{1}^*} = (0.583\ 0.250\ 0.167), \quad \mathbf{w}_2 = \frac{\mathbf{uC}^{-1}}{\mathbf{uC}^{-1}\mathbf{1}^*} = (0.577\ 0.231\ 0.192).$$

This is the weights of a particular pair of two funds in the two fund theorem.

If one takes θ portion of mutual fund of weight $\mathbf{w}_1$ and $1-\theta$ portion of mutual fund with weight $\mathbf{w}_2$, then its return is

$$\begin{aligned} \mu &= \theta \mathbf{w}_1 \mathbf{u}^* + (1-\theta)\mathbf{w}_2 \mathbf{u}^* = 0.0877 - 0.00106\theta, \\ \sigma^2 &= \theta^2 \mathbf{w}_1 \mathbf{C} \mathbf{w}_1^* + 2\theta(1-\theta)\mathbf{w}_1 \mathbf{C} \mathbf{w}_2^* + (1-\theta)^2 \mathbf{w}_2 \mathbf{C} \mathbf{w}_2^* \\ &= 0.00590252 - 0.000138\theta + 0.0000690\theta^2 = 0.49875 - 11.375\mu + 65.625\mu^2 \\ &= 0.076^2 + 65.62(\mu - 0.087)^2. \end{aligned}$$

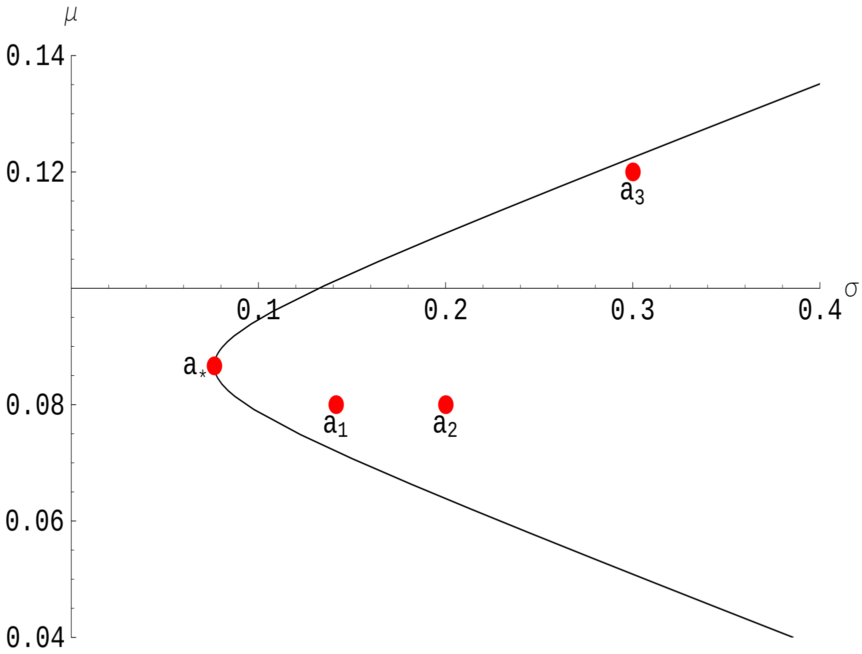


Figure 1.3: The risk-return curve and the risk-return of original assets.

Hence, the Markowitz curve is given by

$$\sigma = \sqrt{0.076^2 + 65.62(\mu - 0.087)^2}.$$

In Figure 1.3, we plot the Markowitz curve, the location of the (risk, expected return) of the three assets, and the location of the minimum risk asset. Note that none of the assets are efficient, since their return and risk are not on the efficient curve.

Thus, the minimum risk of all portfolio is $\sigma_* = 0.076$, attained at $\mu_* = 0.083$. The weight is $\theta_* = 1$, i.e. $\mathbf{w}_* = \mathbf{w}_1$. Hence,

$$\mathbf{w}_* = \mathbf{w}_1 = (0.583,\ 0.250,\ 0.167).$$

In the above example, the first asset a_1 has expect return 8% with risk $\sigma_1 = \sqrt{0.2}$, whereas the second asset a_1 has expect return 8% with risk $\sigma_1 = \sqrt{0.04}$. Just comparing these two assets, one can say that a_1 is more preferable than a_2. However, asset a_2 is not excluded from efficient portfolios.

Exercise 1.8. *When all μ are the same. Find all the attainable region on the return-risk plane.*

Exercise 1.9. *Suppose $\mathbf{C} = (\sigma_{ij})_{m\times m}$ is degenerate, i.e. there exists a non-zero vector $\mathbf{w} = (w_1, \cdots, w_m)$ such that $\mathbf{wC} = \mathbf{0}$. Show that the random variable $R := \sum_{i=1}^m w_i R_i$ is risk-free, i.e., $\mathbf{Var}[R] = 0$ so R is a constant function. Find necessary (and sufficient) conditions for the exclusion of possibility of making money without risk and without any initial vestment.*

Exercise 1.10. *Assume that $\mathbf{C}$ is positive definite. Show that there is a unique portfolio that has the minimum risk. In addition, the weight of this portfolio is given by $\mathbf{w}_1$.*

Exercise 1.11. *Analyze in detail when only two assets are considered. Assume the expected return satisfies $\mu_1 < \mu_2$ and $\sigma_1 > 0, \sigma_2 > 0$. Consider first the case $(\sigma_{ij})_{2\times 2}$ is positive definite and then the case when it degenerate.*

Mark on the Markowitz curve the segments where short selling is not needed.

Exercise 1.12. *Consider a system of three assets, with parameters given as follows:*

	Mean Return	$\mathbf{Cov}(R_i, R_j)$		
Assets		a_1	a_2	a_3
a_1	*0.1*	*0.04*	*-0.006*	*0.016*
a_2	*0.2*	*-0.006*	*0.09*	*0.024*
a_3	*0.3*	*0.016*	*0.024*	*0.14*

1. *Find two examples of two funds that satisfy the two fund theorem.*
2. *Plot the Markowitz curve. Also Mark the risk-return of the three assets.*
3. *Suppose the maximum risk is set at* 0.10*, find the maximum expected return and the corresponding weight.*
4. *Suppose one wants an expect return of* 100%*. How to achieve that?*
5. *Suppose the market is not complete in the sense that one cannot short assets valued more than the portfolio's total worth; (i.e. the sum of all negative* w_i *is no smaller than* -1*.) Find the maximum expected return, regardless how high the risk may be, but still want the risk as small as possible.*
6. *Is it true that in a incomplete market as in (5), the minimum risk-maximum expect return curve always lies on the Markowitz efficient frontier? Either prove of disprove your conclusion.*

Exercise 1.13. *Using the Lagrange multiplier method solving the following problem: Given* $\sigma > 0$,

$$\text{maximize} \quad \mathbf{E}[R] = \sum_{i=1}^{m} w_i R_i \qquad \text{subject to} \quad \sum_{i=1}^{m} w_i = 1, \quad \mathbf{Var}[R] = \sum_{i=1}^{m}\sum_{j=1}^{m} w_i w_j \sigma_{ij} = \sigma^2.$$

Exercise 1.14. *For the three investment plans* $\mathcal{A}, \mathcal{B}, \mathcal{C}$ *in Example 1.3, find one example of two mutual fund that provide all needed efficient portfolios. Also plot the Markowitz curve, as well as the locations of risk-return of the investment plan* $\mathcal{A}, \mathcal{B}, \mathcal{C}$.

1.3 Capital Asset Pricing Model

Now we take a look at the Capital Asset Pricing Model, developed by the Nobel Prize winner William Sharpe and also independently by John Lintner and J. Mossin, thus called SLM CAPM model. The major factor that turns the Markowitz portfolio theory into a capital market theory is the inclusion of a risk–free asset in the model.

A **risk-free asset** is *an asset that gives a fixed return without variability.*

Example 1.8. Suppose today Mellon bank offers the following annul interest rates:

1. Checking account: 2%
2. One year deposit: $4\frac{1}{2}$%;

3. 5 year deposit: 6%
4. 10 year deposit: $5\,\frac{3}{4}$ %.

Assume that each of the investment is guaranteed by federal insurance. Then each of the investment can be regarded as a risk-free investment.

Different from investing on stock for which the return are uncertain at time of investment, the investment on riskless asset has a known return.

As we shall see, the inclusion of a risk-free asset can improve the risk-return balance by investing in a portfolio partially in risky assets and partially in a risk-free asset.

Let us denote by $\mu_0 = R_0$ (almost sure) the return of the underlying risk-free asset, denoted by a_0. Here almost sure means

$$\mathbf{Var}(R_0) = \int_\Omega \Big(R_0(x) - \mu_0\Big)^2 \text{Prob}(dx) = 0.$$

Altogether we have $m+1$ assets $\mathsf{a}_0, \mathsf{a}_1, \cdots, \mathsf{a}_m$ to choose. We use weight $\hat{\mathbf{w}} = (\hat{w}_0, \hat{w}_1, \cdots, \hat{w}_m)$, where $\sum_{i=0}^m \hat{w}_i = 1$, for a generic portfolio. In order to use the Markowitz theory, we can decompose the weight as

$$\hat{\mathbf{w}} = (1-\theta, \theta w_1, \theta w_2, \cdots, \theta w_m), \qquad \sum_{i=1}^m w_i = 1, \quad \theta \in \mathbb{R}.$$

Here θ is the portion of risky assets and $1-\theta$ the portion of the risk-free asset; among risky assets, the relative weight is $\mathbf{w} = (w_1, \cdots, w_m)$. The portfolio return is the random variable

$$\hat{R} = (1-\theta)R_0 + \sum_{i=1}^m \theta w_i R_i = (1-\theta)\mu_0 + \theta R, \quad R = \sum_{i=1}^m w_i R_i.$$

Here $R_0 = \mu_0$ (a.s) is a constant, so that $\mathbf{Cov}(R_0, R_i) = 0$ for all $i = 0, \cdots, m$.

With the inclusion of a risk-free asset, the portfolio with weight $\hat{\mathbf{w}} = (1-\theta, \theta\mathbf{w})$ has the expected return $\hat{\mu}$ and risk $\hat{\sigma}^2$ given by

$$\begin{aligned} \hat{\mu} &= \mathbf{E}(\hat{R}) = (1-\theta)\mu_0 + \theta\mu, \qquad \mu = \mathbf{E}(R) = (\mathbf{w}, \mathbf{u}), \\ \hat{\sigma}^2 &= \mathbf{Var}(\hat{R}) = \theta^2\sigma^2, \qquad \sigma^2 := \mathbf{Var}(R) = (\mathbf{wC}, \mathbf{w}). \end{aligned}$$

Here μ and σ are expected return and risk for the portfolio without risk-free asset. It then follows that the risk-return relation can be expressed in the parametric form, with θ as a free parameter,

$$\begin{cases} \hat{\mu} = \mu_0 + \theta(\mu - \mu_0), \\ \hat{\sigma} = |\theta\sigma| \end{cases} \qquad \theta \in \mathbb{R}. \tag{1.7}$$

Eliminating θ and keeping in mind that only $\hat{\mu} \geqslant \mu_0$ are of our interest, we then obtain the relation

$$\hat{\mu} - \mu_0 = \frac{|\mu - \mu_0|}{\sigma}\,\hat{\sigma}.$$

Here σ and μ, being functions of the relative weight $\mathbf{w}$ on risky assets, can be regarded as parameters which have to be in the attainable region, also called the Markowitz bullet.

Now we see that to obtain the maximum expected return, we need only find the maximum of the slope $|\mu - \mu_0|/\sigma$. As (σ, μ) is in the Markowitz bullet, we see that the maximum can only be attained at the Markowitz curve, i.e., when $\sigma = \sqrt{\sigma_*^2 + \kappa^2(\mu - \mu^*)^2}$. Therefor, the maximum expected return is obtained on the line

$$\hat{\mu} - \mu_0 = \lambda_M \hat{\sigma}, \qquad \lambda_M := \max_{\mu \in \mathbb{R}} \frac{|\mu - \mu_0|}{\sqrt{\sigma_*^2 + \kappa^2(\mu - \mu_*)^2}}. \tag{1.8}$$

This line is called the **Capital Market Line**. It is easy to see that the line is tangent to the Markowitz curve; see Figure 1.2. There are two cases.

(i) $\mu_0 < \mu_*$. In this case, the capital market line is tangent to the Markowitz efficient frontier; see the thick line in Figure 1.2. One can show that the maximum of λ_M is obtained at

$$\mu = \mu_M := \mu_* + \frac{\sigma_*^2}{\kappa^2(\mu_* - \mu_0)}. \tag{1.9}$$

Substituting this μ into the minimum-risk weight formula $\mathbf{w}_M = \mathbf{e}_1 + \mu \mathbf{e}_2$ in the previous section, we then obtain the **market portfolio**

$$\mathbf{w}_M := \frac{(\mathbf{u} - \mu_0 \mathbf{1})\mathbf{C}^{-1}}{((\mathbf{u} - \mu_0)\mathbf{C}^{-1}, \mathbf{1})}. \tag{1.10}$$

(ii) $\mu_0 > \mu_*$. [2] In this case, the capital market line is the extension of the line passing $(0, \mu_0)$ and tangent to the reflection of Markowitz curve about the μ axis; see the thin line in Figure 1.2. One can show that λ_M is obtained at μ given by (1.9), which gives the same relative weight (1.10).

We can now summarize our calculation as follows:

Theorem 1.3 *Consider a market system consisting of a risk-free asset a_0 of return rate μ_0 and risky asserts $a_1, \cdots, a_m$ of expect return $\mu_1, \cdots, \mu_m$ and covariance matrix $\mathbf{C}$.*
For any given risk $\hat{\sigma}$, the maximum expected return $\hat{\mu}$ among all possible portfolios is given by the capital market line equation (1.8). In addition, the relative weight on risky assets are given by (1.10).
In a complete market, any expected return of minimum risk can be attained at a unique portfolio.

Note that the relative weight $\mathbf{w}_M$ in (1.10) on risky assets does not depend on any particular choice of efficient portfolio. This observation is indeed the key to the CAPM.

Theorem 1.4 **(The One-Fund Theorem)** *There is a single fund F of risky assets such that any efficient portfolio can be constructed as a combination of the fund F and the risk-free asset.*

We now explain in more detail on what we have.

(i) If $\mu_0 < \mu_*$, the capital market line (the thick half line in Figure 1.2) is the unique line that passes $(0, \mu_0)$ and is tangent to the Markowitz efficient frontier. By adjusting the portion between the risk-free asset and the risky assets in the portfolio, that is by adjusting the parameter θ which is the total portion of all risk asserts, any risk-return balance on the capital market line can be achieved. To get a point

[2] This case does not have much meaning in finance and therefore its discussion is omitted in most textbooks. Since by investing in risky-asset, one expects larger expected return, and hence, it is meaningful only when $\mu_0 > \mu_*$.

to the right of the market portfolio (the intersection of the line and the frontier curve) requires selling the risk free asset short (since $\theta > 1$ and $1 - \theta < 0$) and using the money to buy more of the market portfolio.

(ii) If $\mu_0 > \mu^*$, the capital market line is above the Markowitz efficient frontier. Nevertheless, to achieve this, one needs (since $\theta < 0$ and $1 - \theta > 1$) to sell the risky assets short and use the money to buy more of the risk-free asset. In reality, the situation $\mu_0 > \mu^*$ does not happen.

(iii) In any situation, to achieve an optimal risk-return balance (i.e. the capital market line), the relative weight of the risky asset has to be the unique weight $\mathbf{w}_M$ given by (1.10).

(iv) The equation $\hat{\mu} = \mu_0 + \lambda_M \hat{\sigma}$ for the capital market line proclaims that the quantity $S_M \hat{\sigma}$, called the **risk premium**, is the additional return beyond the risk-free return μ_0 that one may expect for assuming the risk $\hat{\sigma}$. Of course, it is the presence of risk that the investor may not actually see this additional return. Hence, λ_M is also called the **market price of risk**.

We now state the suggestion provided by the CAPM model to any investor, no matter which kind of risk he/she is willing to take to maximize the return:

In order to maximize the expect return for a given level of risk, what should invest is an **efficient portfolio** *consisting of the risk-free asset and the risky assets with relative weight given by (1.10), where the relative proportion between risk-free asset and risky assets is determined by the level of acceptable risk.*

Example 1.9. Consider the three assets in Example 1.7. Assume the risk-free return is 7%. Then the Market Portfolio has weight

$$\mathbf{w}_M := \frac{(\mathbf{u} - 0.07\mathbf{1})\mathbf{C}^{-1}}{(\mathbf{u} - 0.07\mathbf{1})\mathbf{C}^{-1}\mathbf{1}^*} = (0.55, 0.15, 0.30).$$

The return μ_M and risk σ_M of this market portfolio are respectively

$$\mu_M = \mathbf{w}_M \mathbf{u}^* = 0.092, \qquad \sigma_M = \sqrt{\mathbf{w}_M \mathbf{C} \mathbf{w}_M^*} = 0.0877.$$

Also, the market price of risk is

$$\lambda_M = \frac{|\mu_M - \mu_0|}{\sigma_M} = 0.25.$$

The Capital Market line is the line with the equation

$$\hat{\mu} = 0.07 + 0.25\hat{\sigma}.$$

See Figure 1.4

Finally, we can calculate, for a generic risk-free rate $\mu_0 \in (0, \mu_*)$, the weight of the market portfolio

$$\begin{aligned} \mathbf{w}_M = \frac{(\mathbf{u} - \mu_0\mathbf{1})\mathbf{C}^{-1}}{(\mathbf{u} - \mu_0\mathbf{1})\mathbf{C}^{-1}\mathbf{1}^*} &= (w_1(\mu_0), w_2(\mu_0), w_2(\mu_0)) \\ &= \Big(\frac{0.05 - 0.583\mu_0}{0.0867 - \mu_0}, \quad \frac{0.02 - 0.25\mu_0}{0.0867 - \mu_0}, \quad \frac{0.0167 - 0.0167\mu_0}{0.0867 - \mu_0}\Big). \end{aligned}$$

The three functions are plotted in Figure 1.5, in the unit of percentage.

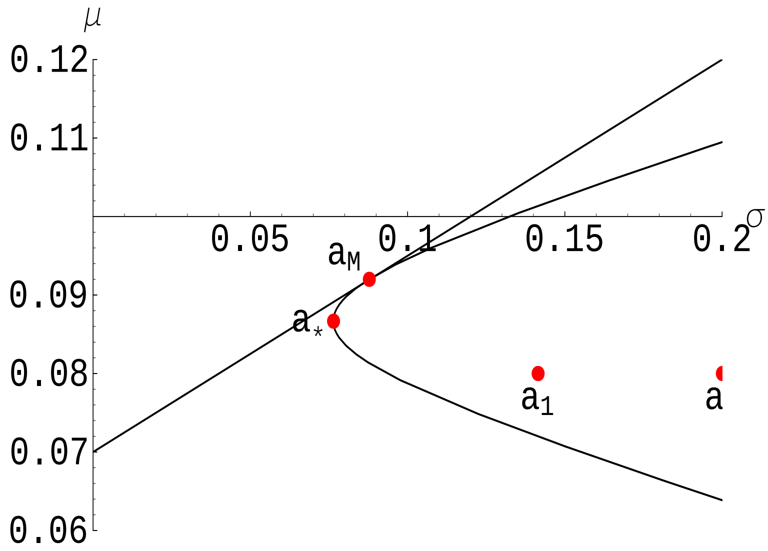


Figure 1.4: The Markowitz Curve and Capital Market Line.

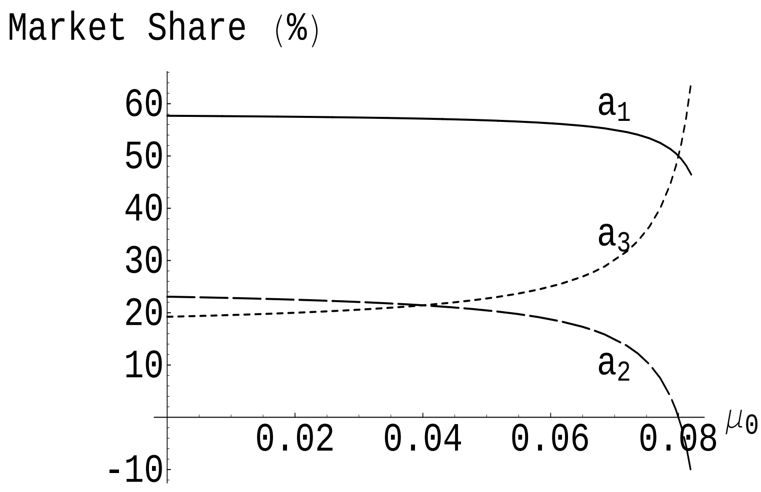


Figure 1.5: Percentage of Market Shares $\mathbf{w}_M = (w_1(\mu_0), w_2(\mu_0), w_3(\mu_0))$ as function of risk-free rate μ_0.

1.3.1 Derivation of the Market Portfolio

Here we derive the formula for the market portfolio. The mathematical problem is following minimization problem: Given $\mu_0, \mu \in \mathbb{R}, \mathbf{u} \in \mathbb{R}^m$, $\mathbf{C} \in \mathbb{R}^{m\times m}$, find $(\theta, \mathbf{w}) \in \mathbb{R}^{1+m}$ that

$$\text{minimize} \quad \theta^2 \mathbf{w} C \mathbf{w}^* \qquad \text{subject to } (1-\theta)\theta\mu_0 + \theta\mathbf{w}\mathbf{u}^* = \mu, \quad \mathbf{w}\mathbf{1}^* = 1. \tag{1.11}$$

We consider the Lagrangian

$$L(\theta, \mathbf{w}, \lambda_1, \lambda_2) := \tfrac{1}{2}\theta^2 \mathbf{w}\mathbf{C}\mathbf{w}^* - \lambda_1\{(1-\theta)\mu_0 + \theta\mathbf{w}\mathbf{u}^*\} - \lambda_2\{\mathbf{w}\mathbf{1}^* - 1\}.$$

If we have a minimizer $(\theta, \mathbf{w})$, then for some Lagrange multiplier λ_1, λ_2, $(\theta, \mathbf{w}, \lambda_1, \lambda_2)$ is a critical point of L, i.e.

$$\frac{\partial L}{\partial \theta} = 0, \quad \frac{\partial L}{\partial w_i} = 0 \ (i = 1, \cdots, m), \quad \frac{\partial L}{\partial \lambda_1} = 0, \quad \frac{\partial L}{\partial \lambda_2} = 0.$$

This leads to the following system of equations

$$\begin{cases} \theta\mathbf{w}\mathbf{C}\mathbf{w}^* = \lambda_1(\mathbf{w}\mathbf{u}^* - \mu_0), \\ \theta^2\mathbf{w}\mathbf{C} = \lambda_1\theta\mathbf{u} + \lambda_2\mathbf{1}, \\ \mu = (1-\theta)\mu_0 + \theta\mathbf{w}\mathbf{u}^* \\ \mathbf{w}\mathbf{1}^* = 1 \end{cases} \tag{1.12}$$

Multiply on right the second equation by $\mathbf{w}^*$ and subtract the resulting equation from the first equation multiplied by θ we obtain

$$0 = \lambda_1(\mathbf{w}\mathbf{u}^* - \mu_0)\theta - \lambda_1\theta\mathbf{u}\mathbf{w}^* - \lambda_2\mathbf{1}\mathbf{w}^* = \lambda_1\theta\mu_0 - \lambda_2$$

since $\mathbf{w}\mathbf{u}^* = \mathbf{u}\mathbf{w}^*$ and $\mathbf{1}\mathbf{w}^* = \mathbf{w}\mathbf{1}^* = 1$. Hence, $\lambda_2 = -\lambda_1\theta\mu_0$. Consequently, multiplying the second equation by $(\theta^2\mathbf{C})^{-1}$ from the right we obtain

$$\mathbf{w} = (\lambda_1\theta\mathbf{u} + \lambda_2\mathbf{1})\mathbf{C}^{-1}\theta^{-2} = \lambda_1\theta^{-1}(\mathbf{u} - \mu_0\mathbf{1})\mathbf{C}^{-1} = \frac{(\mathbf{u} - \mu_0\mathbf{1})\mathbf{C}^{-1}}{(\mathbf{u} - \mu_0\mathbf{1})\mathbf{C}^{-1}\mathbf{1}^*}$$

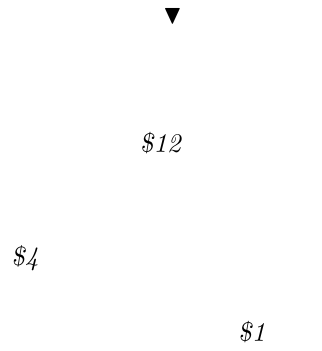
where the last equation os obtained by using $\mathbf{w}\mathbf{1}^* = 1$ which implies $\theta/\lambda_1 = (\mathbf{u} - \mu_0\mathbf{1})\mathbf{C}^{-1}\mathbf{1}^*$.

Exercise 1.15. *Assume that* $\mathbf{u} \neq \mu_0\mathbf{1}$ *and that* $\mathbf{C}$ *is positive definite. Show that for each* $\mu \in \mathbb{R}$*, the minimization problem* (1.11) *admits at least one solution. Consequently, the calculation using the method of Lagrange multiplies shows that the solution is unique.*

Also show that λ_M *in (1.8) is attained at* μ *in* (1.9). *Also, from* (1.9), *derive* (1.10). *Finally, derive a formula for* λ_M.

Exercise 1.16. *Explain what would happen if* $\mu_0 > \mu_*$. *Also explain that in reality it is unlikely that* $\mu_0 > \mu_*$.

Exercise 1.17. *Consider a betting wheel divided into 3 sectors with payoffs \$1,\$4 and \$12 and chances* $0.7, 0.20$ *and* 0.10 *respectively. The game is to place a chip on one of the segment and win the designated amount if the segment appears after a spin and win nothing otherwise.*



A Betting Wheel

Suppose you have 1000 chips, each chip cost $0.80.

(1) How can one place the chips so that the amount to win is independent of the outcome? What is the risk-free rate of the return for the wheel?

(2) Consider the investment plans: (A) put chip on $1 awards segment, (B) put chip on $4 awards segment, and (C) put chip on $12 awards segment. Find the expect return and risk of each investment. Also calculate the correlation matrix.

(3) Find the efficient frontier. [Assume that there is no shorting.]

Exercise 1.18. *Consider a market system consists of three assets with parameters given in Exercise* 1.12.

(a) Assume the risk-free rate is 0.2. Plot the Markowitz curve and the Capital Market line.

(b) Assume the risk-free rate is 0.1. Plot the Markowitz curve and the Capital Market line.

(c) Let μ_0 *be a free parameter. Write the weight of the market portfolio as* $\mathbf{w}_M = (w_1(\mu_0), w_2(\mu_0), w_3(\mu_0))$. *Plot the three curves* $w_1(\mu_0), w_2(\mu_0), w_3(\mu_0)$.

1.4 The Market Portfolio and Risk Analysis

According the CAPM, any rational investor will invest in the market according to efficient portfolios that consist of a $1-\theta$ portion of risk-free asset and the remaining θ portion of risky assets, where θ is a parameter chosen according to the individuals willingness to take the risk to enhance the expected return according to $\hat{\mu} = \mu_0 + \lambda_M \hat{\sigma}$. The most amazing conclusion is that in the portion of the risky assets, the relative weight of the distribution of investment among $a_1, \cdots, a_m$ is given by (1.10). This weight is universal in the sense that it is independent of any individual investor.

That everybody invest according to the CAPM theory has profound consequences.

(a) The market has to contain all assets. Since if an asset a_i is not in the portfolio (e.g. the associated component $\mathbf{w}_M^i$ in the weight in $\mathbf{w}_M = (w_M^1, \cdots, w_M^m)$ is zero), then no one will want to purchase (suggested by the CAMP model) so the asset will wither and die, thus out of market.

(b) If everyone purchases the same mutual fund of risky assets, then the total of this fund must match the **capitalization weights**, being the proportions of each individual asset's total capital value to the total market capital value.

How could capitalization weight equal to the relative weight of risky assets in the efficient portfolio? The answer is based on an **equilibrium** argument. If, based on (1.10), there is a large demand of one particular asset thereby causing short supply, its price will arise, thereby decreasing its rate of return. Similarly, assets under light demand has to decrease its price thereby increase the return. The price change affect the estimates of the assets return directly and also the weights (1.10) in the efficient portfolio. This process continues until demand, base on the market portfolio calculated from the CAPM, exactly matches supply; that is, it continues until there is equilibrium. Under equilibrium, the *percentage of market share of each asset is exactly the weight of the asset in the market portfolio.*

Though this argument has a degree of plausibility and weakness, for the time we shall be content with it. Thus, we assume that the capitalization weight equals the minimum risk weight given in (1.10) and call the corresponding portfolio on risky assets the market portfolio. More precisely,

the **market portfolio** is *the portfolio on risky assert with weight given by* (1.10).

(c) Under a **market equilibrium**, the market portfolio has no unsystematic risk—this risk has been completely diversified out. Here unsystematic risk refers to those risks that affects only individual or localized group of assets. Thus, all risk associated with the market portfolio is systematic risk, i.e., the risk that affects all assets, such as a risk-free rate change, war, terrorism, etc.

To see why we have (c), we shall play around the equations derived from the CAPM.

In the previous sections, we calculated the market portfolio according to risk-free return rate, the risky assets' expect return and their covariance matrix. Now we want to see how the market portfolio affects individual risky asset's system risk.

For convenience, we use a row vector

$$\mathbf{R} = (R_1, \cdots, R_m)$$

to denote all the random variables representing the returns of all risky assets (at end time). As the market portfolio has weight $\mathbf{w}_M$, its return is the random variable

$$R_M = \sum_{i=1}^{m} w_{Mi} R_i = (\mathbf{R}, \mathbf{w}_M).$$

Hence, the market portfolio's expected return μ_M and risk σ_M^2 can be calculated by

$$\mu_M = \mathbf{E}(R_M) = (\mathbf{u}, \mathbf{w}_M), \qquad \sigma_M^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} w_{Mi}\mathbf{Cov}(R_i, R_j) w_{Mj} = (\mathbf{w}_M\mathbf{C}, \mathbf{w}_M).$$

Now here comes the key to our calculation. The CAPM says that $\mathbf{w}_M$ has to be that in (1.10), regardless of the risk that each investor is willing to take. The expression in (1.10) can be written as

$$\mathbf{w}_M = \frac{1}{d}\Big\{\mathbf{u}\mathbf{C}^{-1} - \mu_0\mathbf{1}\mathbf{C}^{-1}\Big\}, \qquad d := (\mathbf{u} - \mu_0\mathbf{1})\mathbf{C}^{-1}\mathbf{1}^*.$$

This implies that $\mathbf{u} = \mu_0\mathbf{1} + d\mathbf{w}_M\mathbf{C}$. This has two consequences, writing $\mathbf{C} = \mathbf{Cov}(\mathbf{R}^t, \mathbf{R})$ for simplicity,

(i) $\quad \mu_M = (\mathbf{u}, \mathbf{w}_M) = (\mu_0\mathbf{1} + d\mathbf{w}_M C, \mathbf{w}_M) = \mu_0(\mathbf{1}, \mathbf{w}_M) + d\,(\mathbf{w}_M\mathbf{C}, \mathbf{w}_M) = \mu_0 + d\,\sigma_M^2,$

(ii) $\quad \mathbf{u} = \mu_0\mathbf{1} + d\,\mathbf{w}_M\mathbf{Cov}(\mathbf{R}^t, \mathbf{R})) = \mu_0\mathbf{1} + d\,\mathbf{Cov}(\mathbf{R}_M, \mathbf{R}),$

i.e. $\quad \mu_k = \mu_0 + d\,\mathbf{Cov}(R_M, R_k) \quad \forall\, k = 1, \cdots, m.$

Now consider the important constant

$$\beta_k := \frac{\mathbf{Cov}(R_k, R_M)}{\sigma_M^2}.$$

(1) From the formulas we just derived, $\beta_k = \frac{(\mu_k-\mu_0)/d}{(\mu_M-\mu_0)/d}$. Thus,

$$\beta_k = \frac{\mu_k - \mu_0}{\mu_M - \mu_0}, \quad \mu_k - \mu_0 = \beta_k(\mu_M - \mu_0), \qquad \mu_k = \mu_0 + \beta_k(\mu_M - \mu_0). \tag{1.13}$$

This line $\mu = \mu_0 + \beta(\mu_M - \mu_0)$ on the β-μ plane is called **security market line** (SML for short). Thus, β_k is the ratio of the **risk premium** $\mu_k - \mu_0$ of the asset a_k and the risk premium $\mu_M - \mu_0$ of the marker portfolio; that is, the risk premium of the asset a_k **magnifies** the risk premium $\mu_M - \mu_0$ of the market portfolio by β_k times. The last equation shows that the expected return of an asset is equal to the return of the risk-free asset plus the **risk premium** $\beta_k(\mu_M - \mu_0)$ of the asset.

It is worthy to point that there is a β book [19] that gives estimates on company's β values. Of course, the book has to be updated from time to time.

(2) Let's see what β_k really is. Decompose R_k as

$$R_k = \beta_k R_M + \varepsilon_k$$

Then

$$\mathbf{Cov}(\varepsilon_k, R_M) \;=\; \mathbf{Cov}(R_k - \beta_k R_M, R_M) = \mathbf{Cov}(R_k, R_M) - \beta_k \sigma_M^2 = 0.$$

Thus, β_k *is the slope of the* **best linear predicator** *for the linear regression of R_k with respect to R_M:*

$$\mathbf{Var}(R_k - \beta_k R_M) = \min_{\beta\in\mathbb{R}} \mathbf{Var}(R_k - \beta R_M). \tag{1.14}$$

(3) Now it is easy to calculate

$$\sigma_k^2 \;=\; \mathbf{Cov}(R_k, R_k) = \beta_k^2 \mathbf{Cov}(R_M, R_M) + \mathbf{Cov}(\varepsilon_k, \varepsilon_k) = \beta_k^2 \sigma_M^2 + \mathbf{Var}(\varepsilon_k).$$

This equation indicates that the risk σ_k^2 of the asset a_k can be decomposed into two parts: $\beta_k^2\sigma_M^2$, called the **systematic risk**, and $\mathbf{Var}(\varepsilon_k)$, called the **unique risk** or **unsystematic risk** of the particular asset; the former depends only on the whole market system whereas the latter depends only on the individual asset (recall $\mathbf{Cov}(R_M, \varepsilon_k) = 0$).

(4) Once we know the meaning of β_k, we can understand better the security market line (1.13).

An asset's expect return $\mu_k = \beta_k(\mu_M - \mu_0) + \mu_0$ depends only on the asset's system risk $\beta_k^2\sigma_M^2$ and does not depend on its unique risk $\mathbf{Var}(\varepsilon_k)$*.*

(5) That $\mathbf{Cov}(R_M, \varepsilon_k) = 0$ for all $k = 1, \cdots, n$ states the following:

The market portfolio has no unsystematic risk, i.e., its expected return does not depend on each individual's unique risk $\mathbf{Var}(\varepsilon_k)$*. All risk associated with the market portfolio is systematic risk.*

Finally, if an efficient portfolio consists of β portion of market portfolio and $1-\beta$ portion of risk-free asset, then its expected return μ and risk σ are given by $\mu = \mu_0 + \beta(\mu_M - \mu_0)$ and $\sigma^2 = \beta^2\sigma_M^2$. From here, we see that *any efficient portfolio does not contain any non-system risk.*

We can also explain the consistency of our conclusion with a **market equilibrium** theory.

(1) The market portfolio has risk σ_M^2 and expected return μ_M. The portion $\mu_M - \mu_0 > 0$ is the "bonus" expected from taking the risk σ_M^2. The expected return μ_M is considered by public as reasonable under risk σ_M^2.

(2) For a particular asset a_k with $\beta_k < 1$, its systematic risk $\beta_k^2\sigma_M^2$ is smaller than the risk of the market portfolio, so its expected return μ_k is smaller than the expect market portfolio return μ_M since $\frac{\mu_k-\mu_0}{\mu_M-\mu_0} = \beta_k < 1$. This is reasonable under the following principal in **market equilibrium**:

(a) if an asset has risk smaller and expected return larger than that of the market portfolio, then more people will buy it and hence raising its price and lowing its expect return.

(3) For a particular asset with $\beta_k > 1$, its systematic risk $\beta_k^2\sigma_M^2$ is larger than the risk σ_M^2 of the market portfolio, so its expect return μ_k is large than that of the marker portfolio since $\frac{\mu_k-\mu_0}{\mu_M-\mu_0} = \beta_k > 1$. This make sense—the more the systematic risk in an asset the higher should be its expected return under another principal of the **market equilibrium**:

(b) if an asset is return less than the market feels is reasonable with respect to the asset's perceived risk, then no one will buy that asset and its price will decline thus increasing the asset's return.

We formalize the discussion in to the following:

Theorem 1.5 *Let R_M be the market portfolio's return with expected return $\mu_M = \mathbf{E}(R_M)$ and risk $\sigma_M = \sqrt{\mathbf{Var}(R_M)}$, under risk-free rate μ_0. Then for each individual asset a_k in the system with return R_k, expected return $\mu_k = \mathbf{E}(R_k)$ and risk $\sigma_k = \sqrt{\mathbf{Var}(R_k)}$, there is a constant, denoted by β_k (that is attained by the driving force of the market equilibrium dynamics) such that*

$$\mu_k - \mu_0 = \beta_k(\mu_M - \mu_0), \qquad \sigma_k^2 = \beta_k^2\sigma_M^2 + \mathbf{Var}(\varepsilon_k),$$
$$\mathbf{Cov}(\varepsilon_k, R_M) = 0, \qquad R_k = \beta_k R_M + \varepsilon_k.$$

In particular, any efficient portfolio consists of a certain β portion of market portfolio and $1-\beta$ portion of risk-free asset and has expected return μ and risk σ given by

$$\mu = \mu_0 + \beta(\mu_M - \mu_0), \qquad \sigma = |\beta|\,\sigma_M.$$

Any efficient portfolio does not contain any non-system risk.

Finally, we introduce two important indexes used in finance community:

$$\textbf{Jensen Index} \qquad J_k = \mu_k - \mu_0 - \beta_k(\mu_M - \mu_0),$$
$$\textbf{Sharp index} \qquad \lambda_k = \frac{\mu_k - \mu_0}{\sigma_k}$$

Theoretically, $J_k = 0$. The real data J_k thus measures approximately how much the performance of an asset has deviated from the theoretical value of zero. A positive value of J_k presumably implies that the fund did better than the CAPM prediction (but of course we recognize that approximations are quite often introduced by insufficient amount of data to estimate the important quantities).

The Shape index measures the efficiency of risk premium of an asset. A lower value of the index implies that the fund is probably insufficient. We note that for the market portfolio and any efficient portfolio, their sharp index is $\lambda_M = (\mu_M - \mu_0)/\sigma_M$, which is the slope of the capital market like, or the market price of risk; cf. (1.9).

Example 1.10. Consider the system of three risky assets as in Example 1.7. Assume that the riskless return is $\mu_0 = 7\%$. Then the weight $\mathbf{w}_M$, return μ_M, and risk σ_M of the market portfolio is

$$\mathbf{w}_M = \frac{(\mathbf{u} - \mu_0\mathbf{1})\mathbf{C}^{-1}}{(\mathbf{u} - \mu_0\mathbf{1})\mathbf{C}^{-1}\mathbf{1}^*} = (0.55, 0.15, 0.3), \quad \mu_M = \mathbf{w}_M\mathbf{u}^* = 0.092, \quad \sigma_M = \sqrt{\mathbf{w}_M\mathbf{C}\mathbf{w}_M^*} = 0.0877.$$

The beta values of the three assets are given by

$$(\beta_1,\ \beta_2,\ \beta_3) = \frac{\mathbf{w}_M\mathbf{C}}{\sigma_M^2} = (0.4545, 0.454545, 2.27).$$

One can check that

$$\mathbf{u} - \mu_0\mathbf{1} = (0.01, 0.01, 0.05), \quad (\mu_M - \mu_0)(\beta_1, \beta_2, \beta_2) = (0.01, 0.01, 0.05).$$

Hence, $\mu_i - \mu_0 = \beta_i(\mu_M - \mu_0)$ for $i = 1, 2, 3$. That is, the Jensen index of each asset is zero.

The Sharp indexes of all assets are

$$\begin{aligned}(\lambda_1, \lambda_2, \lambda_3, \lambda_M) &= \Big(\frac{\mu_1 - \mu_0}{\sigma_1}, \frac{\mu_2 - \mu_0}{\sigma_2}, \frac{\mu_3 - \mu_0}{\sigma_3}, \frac{\mu_M - \mu_0}{\sigma_M}\Big) \\ &= (0.07, 0.05, 0.167, 0.25).\end{aligned}$$

Clearly, the market portfolio has the largest Sharp index, or market price of risk.

The Pricing Formula

The CAPM is a **pricing model**. We now see why.

First of all, the market is driven by demand according to which asset's share price changes. Take an extreme example. Suppose the weight of a stock is negative in the market portfolio; then, according to the CAPM theory, everybody will short sell it. This will drive its price down and consequently increases its return. When price is down to certain level (equivalently the return is increased high enough), the new calculated market portfolio's weight will be positive.

Thus, under the assumption that the market is at equilibrium, we can use certain index fund as a reasonably accurate approximation of the market portfolio to calculate the "true" value of each individual asset in the system thereby determining if it is over priced (due to large demand) or underpriced (due to low demand).

Suppose an asset is purchased at price P and later sold at price Q. The return is $R = (Q - P)/P$. Here P is known and Q is a random variable. Under certain assumption, we may reasonably believe that Q is independent of the price P. The price is in certain way artificial (driven by demand). The fair price of an asset should be judged by its revelation value Q at end of the period. The CAPM uses exactly the information on Q to find its fair price P, at least theoretically.

This is in certain way analogous to an auction process during which a property to be auctioned does not change any bit whereas its price may change significantly.

Putting $R = (Q - P)/P$ in the CAPM formula, we have

$$\mu = \mathbf{E}(R) = \frac{\mathbf{E}(Q)}{P} - 1 = \mu_0 + \beta(\mu_M - \mu_0).$$

Solving P gives **price formula of the CAPM**

$$P = \frac{\mathbf{E}(Q)}{1 + \mu_0 + \beta(\mu_M - \mu_0)}$$

where β is the beta of the asset.

The pricing formula is indeed a linear formula: It depends linearly on Q. To see why, we notice that

$$\beta = \frac{\mathbf{Cov}(R, R_M)}{\sigma_M^2} = \frac{\mathbf{Cov}(Q/P - 1, R_M)}{\sigma_M^2} = \frac{\mathbf{Cov}(Q, R_M)}{P\sigma_M^2}.$$

Substitute this in the pricing formula we then obtain the following **certainty equivalent pricing formula**:

$$P = \frac{1}{1+\mu_0}\Big\{\mathbf{E}(Q) - \frac{(\mu_M - \mu_0)\mathbf{Cov}(Q, R_M)}{\sigma_M^2}\Big\}.$$

Exercise 1.19. *Suppose the risk-free rate is* 3% *and the market portfolio's expected return rate is* 12%. *Consider the following assets*

Asset	a_1	a_2	a_3	a_4	a_5
β	0.65	1.00	1.20	−0.20	−0.60

Find for each asset, the expected return that the asset will be arriving under market equilibrium.

Suppose we find that the (historical) expected returns of these assets are $9\%, 11\%, 13.8\%, 2\%, -2\%$ *respectively. Find their Jensen Indexes.*

Exercise 1.20. *Given two random variables Y and X. The linear regression line of Y with respect to X is the line $y = \beta x + \alpha$ such that*

$$\mathbf{E}\Big((Y - \alpha - \beta X)^2\Big) = \min_{a,b\in\mathbb{R}} \mathbf{E}\Big((Y - a - bX)^2\Big).$$

Show that the linear regression line has slope $\beta = \mathbf{Cov}(Y, X)/\mathbf{Var}(X)$ and y-intercept $\alpha = \mathbf{E}(Y - \beta X)$.

Show that β_k satisfied (1.14).

Exercise 1.21. *Suppose our market system consists of a risk-free asset with return rate 3%, and three risky assets with the following parameters:*

	Mean Return	$\mathbf{Cov}(R_i, R_j)$		
Asset		a_1	a_2	a_3
a_1	*0.1*	*0.04*	*-0.006*	*0.016*
a_2	*0.2*	*-0.006*	*0.09*	*0.024*
a_3	*0.3*	*0.016*	*0.024*	*0.14*

(a) Calculate the weight of the market portfolio. Also, find the minimum risk for a 5% *return.*

(b) Fine β for the market portfolio, as well for each individual assets.

(c) Find the Sharp index for each asset and for the market portfolio. Verify that the Jensen index is zero for each asset.

1.5 Arbitrage Pricing Theory

The information required by the mean-variance theory grows substantially as the number m of assets increases. It requires m means and a total of $m(m+1)/2$ variances and covariances. If $m = 1000$, then we need $501,500$ parameters. It is a formidable task to obtain this amount of information directly. We need a simplified approach.

It is believed that one can sort out a few factors so that the returns of all assets can be traced back to these factors. A factor model that represents this connection between factors and individual returns leads to simplified structure and provides important insight into the relationship among assets.

The factor model framework leads to an alternative theory of asset pricing, termed **arbitrage pricing theory** (APT), originally devised by Ross [25]; for a practical application, see [2]. This theory does not require the assumption that investors evaluate portfolio on the basis of mean and variance; only that, when return are certain, investors prefer greater return to lesser return. In this sense the theory is much more satisfying than CAPM theory which relies on both the mean-variance framework and strong version of equilibrium—assuming everyone used the mean-variance framework.

1. Single–Factor model

Single-factor model assumes that there is a single factor that affects all assets's performance and all assets are correlated to each other through this single factor. Though simple, it illustrate the concept quite well.

Suppose there are m assets $\mathsf{a}_1, \cdots, \mathsf{a}_m$, whose returns are related by

$$R_i = b_i f + e_i, \qquad i = 1, \cdots, m \tag{1.15}$$

where b_i's are fixed constants, f is a random variable describes the system's overall behavior, and e_i are individual factors. It is assumed that all $e_1, \cdots, e_m, f$ are uncorrelated:

$$\mathbf{Cov}(e_i, e_j) = 0, \qquad \mathbf{Cov}(e_i, f) = 0 \qquad \forall\, i, j = 1, \cdots, n, j \neq i. \tag{1.16}$$

From these relations, one obtains

$$\begin{aligned} b_i &= \mathbf{Cov}(r_i, f)/\sigma_f^2, \qquad \sigma_f^2 := \mathbf{Var}(f), \\ \mu_i &= a_i + b_i \mu_f, \qquad \mu_i = \mathbf{E}(R_i), \mu_f := \mathbf{E}(f), a_i := \mathbf{E}(e_i), \\ \sigma_i^2 &= b_i^2 \sigma_f^2 + \sigma_{e_i}^2, \qquad \sigma_i^2 := \mathbf{Var}(R_i), \sigma_{e_i}^2 = \mathbf{Var}(e_i), \\ \sigma_{ij} &= b_i b_j \sigma_f^2, \qquad \sigma_{ij} := \mathbf{Cov}(R_i, R_j), \quad i \neq j, \end{aligned}$$

These equations reveal the primary advantage of a factor model: In the usual representation of asset returns, there are only a total of $3n + 2$ parameters, those of a_i's b_i's, $\sigma_{e_i}^2$'s, and μ_f and μ_f^2.

Now suppose a portfolio has weight $\mathbf{w} = (w_1, \cdots, w_m)$ where $\sum w_i = 1$. Then its return can be calculated by

$$R = \sum w_i R_i = \sum w_i b_i f + \sum w_i e_i = bf + e$$

where $b = \sum w_i b_i =: (\mathbf{w}, \mathbf{b})$ and $e = \sum w_i e_i = (\mathbf{w}, \mathbf{e})$. Consequently,

$$\begin{aligned} \mu &:= \mathbf{E}(R) = a + b\mu_f, \qquad a = \sum w_i \mathbf{E}(e_i), \quad b = \sum w_i b_i, \\ \sigma^2 &:= \mathbf{Var}(R) = b^2 \sigma_f^2 + \sigma_e^2, \\ \sigma_{e^2}^2 &:= \mathbf{Var}(\sum_i w_i w_i) = \sum_{i=1}^m w_i^2 \sigma_{e_i}^2. \end{aligned}$$

It is worthy noting that σ_e^2 should be quite small. For example, it we take $w_i = 1/m$ and assume that $\sigma_{e_i}^2 = s^2$ for all i. Then

$$\sigma_e^2 = s^2/m.$$

That is to say, by diversification, the non-system risk is more or less eliminated. Of course, the system risk $b^2\sigma_f^2$ cannot be eliminated since the factor f influences every asset. The risk due to the e_i's are independent and hence can be reduced by diversification.

We leave the corresponding Markowitz theory and CAPM theory as an exercise.
We have already seen that the CAMP model ends up

$$R_i = \beta_i R_M + e_i$$

where R_M is the return of the market portfolio. Thus, CAMP model can be regarded as a factor model.

2. Arbitrage Pricing Theory

Now assume that there are exactly n factors $f_1, \cdots, f_n$ that influence the return of each asset; that is we assume that the return R_i of asset a_i is given by

$$R_i = b_{i1} f_1 + \cdots, + b_{in} f_n + e_i \tag{1.17}$$

where same as before, all e_i's and f_j's are uncorrelated:

$$\mathbf{Cov}(e_i, e_j) = 0, \qquad \mathbf{Cov}(e_i, f_k) = 0, \qquad \mathbf{Cov}(f_k, f_l) = 0 \quad \forall\, i, j, k, l, i \neq j, k \neq l. \tag{1.18}$$

Here we remark that in application, the number of assets could be couple of thousands, whereas factors could be only a handful.

Theorem 1.6 **(Simple APT Theorem)** *Suppose there are m assets whose returns are governed by $n < m$ factors according to* (1.17) *where e_i are constants. Then there are $m+1$ constants $\mu_0, \lambda_1, \cdots, \lambda_n$ such that*

$$\mu_i = \mu_0 + b_{i1}\lambda_1 + \cdots + b_{in}\lambda_n \qquad \forall i = 1, \cdots, n.$$

This result is highly non-trivial since all constants $e_1, \cdots, e_m$ reduce to a single constant μ_0.

Proof. Set $\mathbf{1} = (1, \cdots, 1)$, $\mathbf{u} = (\mu_1, \cdots, \mu_m)$ and $\mathbf{b}_k = (b_{1k}, \cdots, b_{mk})$, $k = 1, \cdots, n$

Suppose $(\mathbf{w}, \mathbf{1}) = 0$ and $(\mathbf{w}, \mathbf{b}_k) = 0$ for all $k = 0, \cdots, n$. Consider the portfolio with weight $\mathbf{w}$. Its initial value if $(\mathbf{w}, \mathbf{1})V_0 = 0$ and is risk-free. Hence its return is also zero. This implies that $(\mathbf{w}, \mathbf{u}) = 0$. That $\mathbf{u}$ is perpendicular to every vector $\mathbf{w}$ that is perpendicular $\mathbf{1}, \mathbf{b}_1, \cdots, \mathbf{b}_n$. This implies that $\mathbf{u}$ is a linear combination of $\mathbf{1}, \mathbf{b}_1, \cdots, \mathbf{b}_n$. This concludes the proof. □

The existence of μ_0 is the beauty of the theory. Imaging there are thousands of different **well-diversified** portfolios (e.g. mutual funds), each being essentially no unsystematic risks. These portfolio form a collection of assets, the return on each satisfying a factor model with error. We therefore can apply APT to conclude that there are constants $\mu_0, \lambda_1, \cdots, \lambda_n$ such that the well-diversified portfolio having a rate

$$R = b_i f_i + \cdots + b_n f_n + e$$

with the expected return

$$\mu = \mu_0 + b_1\lambda_1 + \cdots + b_n\lambda_n.$$

Here the simple APT theorem applies since the mutual fund is so diversified that it is basically free of unsystematic risks, i.e. $\mathbf{Var}(e) \approx 0$.

Since various well-diversified portfolios can be formed with weights that differ on only a small number of assets, it follows that these individual assets must also satisfies

$$\mu_i = \mu_0 + b_{i1}\lambda_1 + \cdots + b_{in}\lambda_n.$$

(This argument is not completely rigorous, but can be articulated to make more convincing.)

Finally, if we embed the CAPM model into this multi-factor frame work, we have

$$R_i = b_{i1} f_1 + \cdots + b_{in} f_n + e_i$$

and

$$\mathbf{Cov}(R_M, R_i) = b_{i1}\mathbf{Cov}(R_M, f_1) + \cdots + b_{in}\mathbf{Cov}(R_M, f_n).$$

Here the term $\mathbf{Cov}(R_M, e_i)$ is dropped since if the market portfolio represents a well-diversified portfolio, it will essentially uncorrelated with non-system error e_i. Hence,

$$\beta_i = b_{i1}\beta_{f_1} + \cdots + b_{in}\beta_{f_n}, \qquad \beta_{f_i} := \mathbf{Cov}(R_M, f_i)/\sigma_M^2.$$

That is to say, the overall beta of the asset can be considered to be made up from underlying factor betas that do not depend on the particular asset. The weight of these factor betas in the overall asset is equal to the factor loading. Hence in this framework, the reason that different assets have different betas is that they have different loadings.

Example 1.11. Assume a single factor model and that the market portfolio consists of $w_1 = 20\%, w_2 = 30\%$ and $w_3 = 50\%$ of three assets $\mathsf{a}_1, \mathsf{a}_2, \mathsf{a}_3$ respectively. Suppose $\mu_0 = 0.05$, $\mu_M = 0.12$ and $\beta_1 = 2.0, \beta_2 = 0.5$ and $\beta_3 = 1.0$. Find the expected return μ_i of the asset a_i.

Solution. Assume the single factor is f. By scaling, we can assume that $\beta_f = 1$ so that $\beta_i = b_i$. By the Simple APT theorem, there exists λ such that

$$\mu_i = 0.05 + \beta_i\lambda, \quad i = 1, 2, 3.$$

Also, we know that

$$0.12 = \mu_M = w_1\mu_1 + w_2\mu_2 + w_3\mu_3 = 0.05 + \lambda \sum w_i\beta_i = 0.05 + 1.05\lambda.$$

Hence, $\lambda = 0.07/1.05 = 0.0667$ and

$$(\mu_1, \mu_2, \mu_3) = 0.05(1, 1, 1) + (\beta_2, \beta_2, \beta_3)\lambda = (18.3\%, 8.3\%, 11.67\%).$$

Exercise 1.22. *Assume the single factor model (i.e.* (1.15) *and* (1.16) *hold).*

(1) Calculate the Markowitz efficient frontier, as well the two funds in the Two Fund Theorem.

(2) Using the CAPM theory, calculate the market portfolio. Also, calculate the β_k for each asset a_k.

(3) Suppose $\sigma_i^2 < s^2$ for all i. Let R_M be the return of the market portfolio. Find the limit of R_M as $m \to \infty$. What is the relation between the market portfolio and the single factor?

Exercise 1.23. *Suppose risk-free rate is $\mu_0 = 10\%$ and two stocks are believed to satisfy the two-factor model*

$$R_1 = 0.01 + 2f_1 + f_2, \qquad R_2 = 0.02 + 3f_1 + 4f_2.$$

Find λ_1, λ_2 in the simple APT theorem.

Exercise 1.24. *Someone believes that the collection of all stocks satisfy a sing-factor model whose single factor is the market portfolio that gives information needed for three stocks A, B, C. Assume that risk-free rate is 5%, market portfolio's expected return is 12% with standard deviation 18%. Information on the portfolio is as follows:*

Stock	Beta	σ_{e_i}	weight
A	1.10	7.0%	20%
B	0.80	2.3%	50%
C	1.00	1.0%	30%

Find the portfolio's expected return and its standard deviation.

1.6 Models and Data

Mean-variance portfolio theory and the related models of the CAPM and APT are frequently applied to equity securities (i.e. publicly traded stocks). Typically when using mean-variance theory to construct a portfolio, a nominal investment period, or planning horizon, is chosen, say one year or one month, and the portfolio is optimized with respect to the mean and the variance for the period. However, to carry out this procedure, it is necessary to assign specific numerical values to the parameters of the model: the expected and the variance of those returns, and covariance between the returns of different securities. Where do we obtain these parameter values?

One obvious source is historical data of security returns. This method of extracting the basic parameters from historical return data is commonly used to structure mean-variance models. It is a convenient method since suitable sources are readily available. Are they reliable?

Here we shall investigate the **statistical limitation** in extracting parameters, which we call **blur of history**. It is important to understand the basic statistics of data processing and this fundamental limitation.

1.6.1 Basic Statistics

Suppose R is a random variable, with mean μ and variance σ^2. The purpose here is to use observations to estimate μ and σ. For this, we make n observations and record the values of R by $\{r_i\}_{i=1}^n$. It is a quite standard procedure that one uses the following as approximations of μ and σ^2:

$$\bar{\mu} := \frac{1}{n}\sum_{i=1}^n r_i, \qquad \bar{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (r_i - \bar{\mu})^2.$$

Now our question is how accurate is the estimation

$$\mu \approx \bar{\mu}, \qquad \sigma \approx \bar{\sigma}?$$

To answer such a question, let's suppose that the n observations are independent; more precisely, assume that $\{r_i\}_{i=1}^n$ are i.i.d. (independently identically distributed) random variables, with the same distribution as that R.

Given an interval $[a,b]$, we intend to calculate the probability

$$p \quad := \quad \text{Prob}(\mu - \bar{\mu} \in [a,b]) = \text{Prob}(\mu \in [\bar{\mu}+a, \bar{\mu}+b]).$$

The interval $[\bar{\mu}+a, \bar{\mu}+b]$ is called a **confidence interval** of μ with **confidence level** p.

Similarly, the confidence level of the interval $[\bar{\sigma}+a, \bar{\sigma}+b]$ for σ is

$$p \quad := \quad \text{Prob}(\sigma - \bar{\sigma} \in [a,b]) = \text{Prob}(\sigma \in [\bar{\sigma}+a, \bar{\sigma}+b]).$$

Quite often, one first chooses a confidence level p and then find an ε such that the p-confidence interval has length $b - a = 2\varepsilon$.

Now suppose R is normally distributed with mean μ and variance σ^2. Then $\bar{\mu}$ is normally distributed with mean μ and variance σ^2/n; that is, $\sqrt{n}(\mu-\bar{\mu})/\sigma$ is $N(0,1)$ (normal with mean zero and unit variance) distributed. Hence,

$$P(z) := \int_{-z}^{z} \frac{e^{-s^2/2}ds}{\sqrt{2\pi}} = \text{Prob}\Big(\frac{\sqrt{n}(\mu-\bar{\mu})}{\sigma} \in [-z,z]\Big) = \text{Prob}\Big(\mu \in [\bar{\mu} - z\sigma/\sqrt{n}, \bar{\mu} + z\sigma/\sqrt{n}]\Big).$$

That is to say, the $P(z)$-confidence interval for μ is $[\bar{\mu}-\varepsilon, \bar{\mu}+\varepsilon]$ where

$$\varepsilon = \frac{\sigma\, z}{\sqrt{n}}.$$

We list the relation between z and $P(z)$. Quite often, the value z is expressed as a function of q where $q = (1-p)/2$ is the area of the region under the curve $y = e^{-x^2/2}/\sqrt{2\pi}$ for x in $[z, \infty)$.

z_q	1	2	3	1.28	1.64	1.96	2.57	3.29	3.89	4.42
$P(z)$	68.3%	95.5%	99.7%	80%	90%	95 %	99% %	99.9 %	99.99%	99.999%
q				0.1	0.05	0.025	0.005	0.0005	0.00005	0.000005

Example 1.12. (1) Suppose $\sigma = 25$ and the mean of 20 samples is 112. Taking $p = 90\%$, the z-value is $z = z_{0.05} = 1.64$. Hence,

With 90% confidence, the mean is in the interval $112 \pm 1.64 * 25/\sqrt{20}$*, i.e.* $[103, 121]$.

(2) Suppose in a poll the 95 % confidence interval of percentage of population supporting a candidate is $39\% \pm 3\%$. We report as follows,

"the pool indicates that 39% population supports the candidate, where sample error is $\pm 3\%$*."*

(3) Consider a poll investigating the percentage of population supporting a bill. Suppose $\sigma = 0.2 = 20\%$ and $\varepsilon = 0.03 = 3\%$. To achieve a 95% confidence interval of width $2\varepsilon = 6\%$, the sample size needs to be, since $z = 1.96$, $n \geq N := (z\sigma/\varepsilon)^2 = (1.96 * 0.2/0.03)^2 = 171$. Namely, at least 171 people need to be asked to obtain the percentage of population supporting the bill, with sample error $\pm 3\%$.

Remark 1.1. 1. Here the z-test is based on the central limit theorem. The distribution for $z = (\bar{\mu}-\mu)/(\sigma\sqrt{n})$ is exactly $N(0,1)$ if $\xi_1, \cdots, \xi_n$ are normally distributed i.i.d random variables. If they are not normally distributed, we need to make sure n is not too small, so the deviation of the distribution of z from $N(0,1)$ is not a significant factor to our conclusions.

2. In most cases σ is not known. To find confidence interval, one uses the estimator $\bar{\sigma}$ to replace σ. It is shown by William Sealy Gosset in 1908 under the name of Student that the statistics

$$t := \frac{\bar{\mu}-\mu}{\bar{\sigma}/\sqrt{n}}$$

has the distribution now called student t-distribution, with $n-1$ degree of freedom. Hence, the p-confidence interval for μ is $[\bar{\mu}-\varepsilon, \bar{\mu}+\varepsilon]$ (or $\bar{\mu} \pm \varepsilon$) where

$$\varepsilon = \frac{\bar{\sigma}t}{\sqrt{n}}.$$

Here t is to be found from the Student-t distribution table. When $n \geq 10$, one can use the approximation $t \approx z$.

To find the confidence interval for the variance, we can use the Cochran's theorem which says that if $\{r_i\}_{i=1}^n$ are i.i.d $N(\mu, \sigma^2)$ distributed, then $(n-1)\bar{\sigma}^2/\sigma^2$ is Chi square distributed with degree of freedom $n-1$:

$$(n-1)\frac{\bar{\sigma}^2}{\sigma^2} \sim \chi^2_{n-1}.$$

The Chi square distribution with k freedom has density function

$$\frac{(1/2)^{k/2}}{\Gamma(k/2)} x^{k/2-1} e^{-x/2}, \qquad x \geqslant 0.$$

Fisher showed that if $X \sim \chi^2_k$, then $\sqrt{2X} - \sqrt{2k-1}$ is approximately $N(0,1)$ distributed when $k \gg 1$. Thus, when n is larger,

$$\sqrt{2(n-1)}\Big\{\frac{\bar{\sigma}}{\sigma} - \sqrt{\frac{2n-3}{2n-2}}\Big\} \sim N(0,1). \tag{1.19}$$

Hence, when $n \gg 1$, the p-confidence interval is approximately

$$\Big[\frac{\bar{\sigma}}{1+z/\sqrt{2n}}, \frac{\bar{\sigma}}{1-z/\sqrt{2n}}\Big]$$

where $z = z_q$ with $q = (1-p)/2$.

Example 1.13. If we use the approximation (1.19), we find that the p-confidence interval for σ is $[\bar{\sigma}/(1+z/\sqrt{2n}), \bar{\sigma}/(1+z/\sqrt{2n})]$.

Suppose the sample variance is calculated as $\bar{\sigma} = 0.20$. When $p = 95\%$, $z = z_{0.025} = 1.95996$. Hence, the 95% confidence interval for σ is

$$[0.1566,\ 0.2767] \text{ if } n = 25, \qquad [0.1839,\ 0.2192] \text{ if } n = 250, \quad [0.1946,\ 0.2057] \text{ if } n = 2500.$$

Similarly, if $p = 90\%$ so $z = z_{0.05} = 1.64449$, the confidence interval is

$$[0.162,\ 0.261] \text{ if } n = 26, \qquad [0.186,\ 0.216] \text{ if } n = 251, \quad [0.195,\ 0.205] \text{ if } n = 2500.$$

Using Mathematica's "Statistics ConfidenceIntervals" pacakege, we find the following:

Sqrt[ChiSquareCI[0.04, 25, ConfidenceLevel → 0.95]]={0.1569, 0.2761]},
Sqrt[ChiSquareCI[0.04, 250, ConfidenceLevel → 0.95]]={0.1839, 0.2192]},
Sqrt[ChiSquareCI[0.04, 2500, ConfidenceLevel → 0.95]]={0.1946, 0.2057]}.

This result is the same as above Fishes' approximation. Similar accuracy works also for confidence level = 90%.

It is convenient to write an interval $[a-\varepsilon, a+\varepsilon]$ as $a \pm \varepsilon$. Hence, when $n = 2500$ and confidence level 95%, we can say

$$\sigma = 0.200 \pm 0.006.$$

Finally, we consider the correlation coefficient $\bar{\rho}_{12}$ of two sample data $\{x_{1k}\}_{k=1}^{n}$ and $\{x_{2k}\}_{k=1}^{n}$

$$\bar{\rho}_{12} := \frac{S_{12}}{\sqrt{S_{11}S_{22}}}, \qquad S_{ij} := \frac{1}{n-1}\sum_{i=1}^{n}(x_{ik}-\bar{x}_i)(x_{jk}-\bar{x}_j), \quad \bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_{ik}.$$

Then $\bar{\rho}_{12}$ has the density

$$f(x;\rho) = \frac{\Gamma(n-1)\Gamma(n-2)}{\sqrt{2\pi}\Gamma(n-1/2)}(1-\rho^2)^{(n-1)/2}(1-\rho x)^{1/2-n}(1-x^2)^{n/2-2}{}_2F_1\left(\tfrac{1}{2},\tfrac{1}{2};n-\tfrac{1}{2};\tfrac{1}{2}(1+x\rho)\right)$$

where ${}_2F_1$ is hypergeometric function.

It is rather complicated to obtain the confidence interval from the above density function. Quite often, we use approximations. First let define

$$\bar{\xi} = \tanh^{-1}(\rho_{12} = \tfrac{1}{2}\log\frac{1+\bar{\rho}_{12}}{1-\bar{\rho}_{12}}, \qquad \xi := \tanh^{-1}(\rho).$$

Then

$$\sqrt{n-1}(\bar{\xi}-\xi) \longrightarrow N(0,1) \text{ as } n\to\infty.$$

Thus, the p-confidence interval for ξ is approximately $[\bar{\xi}-\frac{z}{\sqrt{n-1}}, \bar{\xi}+\frac{z}{\sqrt{z-1}}]$. Further results shows that a reasonably good asymptotic confidence interval for $\rho_{12} = \mathrm{Cor}[x_1,x_2]$ is

$$\left[\tanh\left(\tanh^{-1}(\bar{\rho}_{12}) - \frac{z}{\sqrt{n-3}}\right), \tanh\left(\tanh^{-1}(\bar{\rho}_{12}) + \frac{z}{\sqrt{n-3}}\right)\right]$$

where z is the z value for $N(0,1)$ distribution mentioned earlier.

Example 1.14. Suppose $\bar{\rho}_{12} = 0.9$ with $n = 3000$, we have 95% confidence interval

$$\rho_{12} \in \left[\tanh\left(\tanh^{-1}(0.9) - \frac{1.95996}{\sqrt{2997}}\right), \tanh\left(\tanh^{-1}(\rho_{12}) + \frac{1.95996}{\sqrt{2997}}\right)\right] = [0.8930, 0.9066].$$

1.6.2 Stock Returns

We investigate how we can extracting the expected return rate and variance of a security from historical data. For simplicity, we use continuously compounded rate.

Suppose the security under investigation time period $[0, T]$, with unit time being one year, has time invariant expected return rate ν and variance σ^2; that is, denote by S^t the unit share price of the security, we have

$$S_t = S_{t-1} e^{r(t)}, \qquad \mathbf{E}(r(t)\,) = \nu, \qquad \mathbf{Var}(r(t))\, = \sigma^2 \qquad \forall\, t \in [1, T].$$

Now suppose each year is divided into p periods of equal length and we have T years of historical data on the beginning and ending points of these periods. For convenience, we use $t_0, t_1, \cdots, t_{pT}$ for these dates. Then we can write

$$t_j = t_0 + j\Delta t, \qquad j = 0, 1, 2, \cdots, pT, \qquad \Delta t = \frac{1}{p}.$$

We denote the corresponding period return rate by $r_i, i = 1, \cdots, n$. Then

$$S_{t_i} = S_{t_{i-1}} e^{r_i} \quad \Big(\text{i.e. } r_i = \ln \frac{S_{t_i}}{S_{t_{i-1}}} \quad \Big), \qquad i = 1, 2, \cdots, pT$$

It is not unreasonable to assume that all r_i are i.i.d random variables. Hence, we use $\mu_{\Delta t}$ and $\sigma^2_{\Delta t}$ to denote the expected value and variance of each r_i. That is

$$\nu_{\Delta t} = \mathbf{E}(r_i), \qquad \sigma^2_{\Delta t} = \mathbf{Var}(r_i) \qquad \forall\, i = 1, 2, \cdots, pT.$$

We investigate the relation between ν, σ and $\nu_{\Delta t}, \sigma_{\Delta t}$. Pick any integer j such that $0 \leqslant p \leq j < pT$. We know

$$r(t_{j+p}) := \sum_{i=j+1}^{j+p} r_i = \sum_{i=j+1}^{j+p} \ln \frac{S_{t_i}}{S_{t_{i-1}}} = \ln \frac{S_{t_{j+p}}}{S_{t_j}} = \ln \frac{S_{1+t_j}}{S_{t_j}}.$$

represents the annual return rate. Hence,

$$\begin{aligned} \mu &= \mathbf{E}(r(t_{j+p})) = \mathbf{E}\Big(\sum_{i=j+1}^{j+p} r_i\Big) = \sum_{i=j+1}^{j+p} \mathbf{E}(r_i) = p\nu_{\Delta t} = \frac{\nu_{\Delta t}}{\Delta t}, \\ \sigma^2 &= \mathbf{Var}(R_{j+p}) = \mathbf{Var}\Big(\sum_{i=j+1}^{j+n} r_i\Big) = \sum_{j=i+1}^{j+p} \mathbf{Var}(r_i) = p\sigma^2_{\Delta t} = \frac{\sigma^2_{\Delta t}}{\Delta t}. \end{aligned}$$

Thus we have the following scaling law:

Lemma 1.1. *Suppose $r_1, r_2, \cdots, r_{pT}$ are i.i.d. random variables representing the return in $\Delta t = 1/p$ period in T units time. Let r be the corresponding return in unit time. Then the mean and variance obeys*

$$\begin{aligned} \nu_{\Delta t} &= \nu\Delta t, \qquad \mu := \mathbf{E}(r), \quad \mu_{\Delta t} := \mathbf{E}(r_i) \quad \forall\, i = 1, \cdots, pT, \\ \sigma^2_{\Delta t} &= \sigma^2\Delta t, \qquad \sigma^2 := \mathbf{Var}(r), \quad \sigma^2_{\Delta t} := \mathbf{Var}(r_i) \quad \forall i = 1, \cdots, pT. \end{aligned}$$

Now we see the error in using the following estimators for ν and ν_Δ:

$$\bar{\nu} := \frac{1}{T}\sum_{i=1}^{pT} r_i, \qquad \bar{\nu}_{\Delta t} = \frac{1}{pT}\sum_{i=1}^{pT} r_i = \bar{\nu}\Delta t.$$

We calculate,

$$\begin{aligned} \mathbf{E}(\bar{\nu}) &= \frac{1}{T}\sum_{i=1}^{pT} \mathbf{E}(r_i) = \nu, \\ \mathbf{Var}(\bar{\nu}) &= \frac{1}{T^2}\sum_{i=1}^{pT} \mathbf{Var}(r_i) = \frac{pT\sigma^2_{\Delta t}}{T^2} = \frac{\sigma^2}{T}, \\ \mathbf{SD}(\bar{\nu}) &= \sqrt{\mathbf{Var}(\bar{\nu})} = \frac{\sigma}{\sqrt{T}}. \end{aligned}$$

Similarly,

$$\bar{\nu}_{\Delta t} = \bar{\nu}\ \Delta t, \quad \mathbf{Var}(\bar{\nu}_{\Delta t}) = \mathbf{Var}(\bar{\nu})(\Delta t)^2 = \frac{\sigma^2}{p^2 T}, \qquad \mathbf{SD}(\bar{\nu}_{\Delta t}) = \frac{\sigma}{p\sqrt{T}} = \frac{\sigma_{\Delta t}}{\sqrt{T}}.$$

Now lets put a few numbers into these formulas.

Example 1.15. We take a quite typical annul return rate $\nu = 12\%$ and standard deviation $\sigma = 15\%$.

(1) Take $p = 12$ so $\Delta t = 1$ month. Hence, the monthly rate is $\nu_{\Delta t} = 1\%$ with deviation $\sigma_{\Delta t} = \sigma/\sqrt{12} = 4.33\%$. Thus, in a typically month, the monthly return rate is 1% subject to a 4.33% deviation, which is large than the expected rate itself.

(2) Suppose we take $p = 250$, the average number of trading days in a year, then the average daily return is $12\%/250 = 0.048\%$, whereas the standard deviation is $15\%/\sqrt{250} = .95\%$. This result is consistent with with out ordinary experience. On any given day, a stock value may easily move 0.5% or 2%, whereas the expected change is only about 0.05%. The daily mean is low compared to the daily fluctuation.

(3) Suppose we use one year of monthly data, i.e. $T = 1$ and $p = 12$. Then we have $\mathbf{SD}(\bar{\nu}_{\Delta t}) = \sigma/\sqrt{12} = 1.25\%$. We are only able to say the mean is 1% plus or minus 1.25%. The following is a sample list of 8 year's of average monthly return rate, with true mean being 1%:

$$3.02,\ .52,\ 1.67,\ 0.01,\ 1.76,\ 2.06,\ 1.37,\ .17 \quad \text{Average } : 1.37\%$$

(4) In order to obtain a reasonably good estimation, we need a standard derivation of about one-tenth of the mean value itself. This would require $T = (4.33 * 10)^2 = 1875$ month or 156 years of data, which is impossible since there is no way the expected monthly rate being a constant for such a long time!

(5) If we use 9 years of data to estimate ν by $\bar{\nu}$, the standard deviation is $\mathbf{SD}(\bar{\nu}) = \sigma/\sqrt{T} = 15\%/\sqrt{9} = 5\%$. Thus, even with 9 years of data, we still can only say (with very lower confidence), that the **expected annul return rate** is $12\%\pm5\%$.

From the example, we see that there is a statistical limitation on the measurement of data. The lack of reliability is not due to the faulty data or difficult computation, it is due to a fundamental limitation on the process of extracting estimates.

This is the historical blur problem for the measurement of ν. It is basically impossible to measure ν by $\bar{\nu}$ to within workable accuracy using historical data. Furthermore, the problem cannot be improved much by changing the period of length. If longer period are used, each sample is more reliable, but fewer independent samples are obtained in any year. Conversely, if smaller periods are used, more samples are available, but each is worse in terms of ratio of standard deviation to mean value. This problem of mean blur is a fundamental difficulty.

We remark that the variance $\sigma_{\Delta t}$ or $\sigma^2 = \sigma^2_{\Delta t}/\Delta t$ can be reasonably well approximated by the estimator

$$\bar{\sigma}^2_{\Delta t} = \frac{1}{pT-1}\sum_{i=1}^{pT}(r_i - \bar{\nu}_{\Delta t})^2, \qquad \bar{\sigma}^2 = \frac{\bar{\sigma}^2_{\Delta t}}{\Delta t} = \frac{p}{pT-1}\sum_{i=1}^{pT}(r_i - \bar{\nu}_{\Delta t})^2.$$

For example, suppose the return rate is a Brownian motion, then each r_i is normally distributed. We found that

$$\mathbf{SD}(\bar{\sigma}^2_{\Delta t}) = \frac{\sqrt{2}\sigma^2_{\Delta t}}{\sqrt{pT-1}} = \frac{\sqrt{2}\sigma^2}{p\sqrt{pT-1}}, \qquad \mathbf{SD}(\bar{\sigma}^2) = \frac{\sqrt{2}\sigma^2}{\sqrt{pT-1}}.$$

For example, if we take one year data, then the standard deviation of $\bar{\sigma}^2_{\Delta t}$ is $\sqrt{2}*(15\%)^2/(12*\sqrt{11}) = 0.0008$. Thus, we know that $\sigma^2_{\Delta t}$ is approximately 0.001875 ± 0.0008. This renders to the estimate monthly variance being or $4.33\% \pm 1\%$. If 9 years of data are used, we can say $\sigma_{\Delta t} = 4.33\% \pm 0.3\%$. In terms of $\sigma = \sigma_{\Delta t}\sqrt{12}$, we can say that $\sigma = 15\% \pm 1\%$. If daily returns are used, the estimate is even better.

Example 1.16. Tables 1.1–1.3 illustrate the mean and volatility and their confidence intervals, with various methods of sampling, for IBM. The unit for mean is 1/year and unit for volatility is $1/\sqrt{\text{year}}$.

Here for simplicity, we assume that there are 5 trading days per week, 21 trading days per month, and 252 trading days per year. Main codes, written in mathematica, are illustrated as follows:

```
⟨⟨ Statistics ‘ConfidenceIntervals‘
date={“12/29/2006”,“12/28/2006”,“12/26/2006”,“12/21/2006”, ···, “1/4/1962”,“1/2/1962”};
stockprice={96.68, 96.91, 95.37, 94.96, ···, 2.82, 2.8};
timeintervalname={“2001–2006”, “1997–2006”, “1987–2006”, “1967–2006”};
  timeinterval=252*{5, 10, 20, 40}; numberoftimeinterval=4;
methodname={“daily”, “weekly”, “monthly”, “annuly”}; numuerofmethod=4;
  jump={1, 5, 21, 252};
Do[ d=jump[[j]]; n=timeinterval[[i]]/d;
  return =Table[ Log[stockprice[[1+ k*d]]/stockprice[[1+(k+1)d]]  ],  {k, 0, n-1}];
  dt=d/252;   returnrate=return/dt;
  meanreturnrate[i,j]=Mean[returnrate];
  meanconfidenceinterval[i,j]=MeanCI[returnrate, ConfidenceLevel → 68%];
  standarddeviation = StandardDeviation[returnrate];
  varianceconfidenceinterval= VarianceCI[returnrate, ConfidenceLevel → 95%];
  volatility[i,j]=standardderivation * Sqrt[dt];
  volatilityconfidenceinterval[i,j]= Sqrt[varianceconfidenceinterval*dt],
  {i,1,numberofmethod},{j,1,numberoftimeinterval}]
```

One concludes from the table that the confidence interval for the volatility shrinks as the time interval of sampling shrinks. Nevertheless, the width of confidence interval for mean does not shrink as time interval shrinks.

Exercise 1.25. *Suppose* $r_i = \nu\Delta t + \sigma\sqrt{\Delta t}\; z_i$ *where* $z_1, z_2, \cdots,$ *are i.i.d. normally distributed random variables with mean zero and variance one. Take* $\nu = 10\%$ *and* $\sigma = 15\%$ *and* $\Delta t = 1/250$ *(e.g. one day).*

(i) Use a random number generator generating 10 years of daily rate of return: $r_1, r_2, \cdots, r_{2500}$.

(ii) Do statistics on the data, estimating annul rate ν *of return and standard deviation* σ. *Also find estimation intervals for* ν *and* σ *with a 70% confidence.*

Exercise 1.26. *Go through internet, find daily return rate of a particular stock or index, find its expected annul rate of return and variance. Present your result in terms of confidence intervals.*

Table 1.1: IBM volatility ($1/\sqrt{\text{year}}$) and Its 95% Confidence Interval

Samples	2001-2006	1997-2006	1987-2006	1067--2006
daily	0.249 { 0.240, 0.260}	0.327 { 0.320, 0.340}	0.302 { 0.300, 0.310}	0.263 { 0.260, 0.270}
weekly	0.268 { 0.250, 0.290}	0.319 { 0.300, 0.340}	0.291 { 0.280, 0.300}	0.254 { 0.250, 0.260}
monthly	0.288 { 0.240, 0.350}	0.328 { 0.290, 0.380}	0.308 { 0.280, 0.340}	0.267 { 0.250, 0.290}
anully	0.282 { 0.170, 0.810}	0.328 { 0.230, 0.600}	0.304 { 0.230, 0.440}	0.278 { 0.230, 0.360}

Table 1.2: IMB mean return rate (1/year) and its 95% Confidence Interval

Samples	2001-2006	1997-2006	1987-2006	1067--2006
daily	-0.038 {-0.256, 0.180}	0.099 {-0.104, 0.302}	0.078 {-0.055, 0.210}	0.083 { 0.001, 0.164}
weekly	-0.038 {-0.274, 0.198}	0.099 {-0.099, 0.297}	0.078 {-0.050, 0.205}	0.083 { 0.004, 0.162}
monthly	-0.038 {-0.296, 0.220}	0.099 {-0.107, 0.304}	0.078 {-0.058, 0.213}	0.083 {-0.000, 0.166}
anully	-0.038 {-0.388, 0.313}	0.099 {-0.136, 0.334}	0.078 {-0.064, 0.220}	0.083 {-0.006, 0.172}

Table 1.3: IMB mean return rate (1/year) and its 68% Confidence Interval

Samples	2001-2006	1997-2006	1987-2006	1067--2006
daily	-0.038 {-0.149, 0.073}	0.099 {-0.005, 0.202}	0.078 { 0.010, 0.145}	0.083 { 0.041, 0.124}
weekly	-0.038 {-0.158, 0.082}	0.099 {-0.002, 0.200}	0.078 { 0.013, 0.143}	0.083 { 0.043, 0.123}
monthly	-0.038 {-0.168, 0.092}	0.099 {-0.006, 0.203}	0.078 { 0.009, 0.147}	0.083 { 0.041, 0.125}
anully	-0.038 {-0.182, 0.106}	0.099 {-0.011, 0.209}	0.078 { 0.008, 0.148}	0.083 { 0.038, 0.127}

Summary

If everybody uses the mean-variance approach to invest and if everybody has the same estimates of the asset's expected returns, variances, and covariances, then everybody must invest in the same fund F of risky assets and in the risky-free asset. Because F is the same for everybody, it follows that, in equilibrium, F must correspond to the market portfolio M—the portfolio in which each asset is weighted by its proportion of total market capitalization. This observation is the basis for the capital asset pricing model.

If the market portfolio is the efficient portfolio of risky assets, it follows that the efficient frontier in the $\mu - \sigma$ diagram is a straight line that emanates from the risk-free point and passes through the point representing M. This line is the capital market line. Its slope is called the market price of risk, known as the Sharp index of market portfolio. Any efficient portfolio must lie on this line, i.e. has the Sharp index of market portfolio.

The CAPM can be represented graphically as a security market line:

$$\mu - \mu_0 = \beta(\mu_M - \mu_0)$$

where $\beta = \mathbf{Cov}(R, R_M)/\sigma_M^2$ is called the beta of the asset. Greater beta implies greater expected return. Also, the system risk of an asset is fully characterized by its beta.

The beta of risk-free asset is zero; the beta of the market portfolio is one. The betas of other stocks take other values, but the betas of most U.S stocks takes rang between 0.5 and 2.5. The beta of a portfolio of stocks is equal to the weighted average of the betas of the individual assets that make up the portfolio.

The CAPM can be converted to an explicit formula for the price of an asset. It is important to recognize that the pricing formula is linear.

One way to use mean-variance theory is to rely on the insight of the CAPM that if everyone followed the mean-variance approach and everyone agreed on the parameters, then the efficient fund of risky assets would be the market portfolio.

Using this idea, you need not compute anything; just purchase a mixture of the market portfolio and the risk-free asset. Many investors are not completely satisfied with this approach and believe that a superior solution can be computed by solving the Markowitz mean-variance portfolio problem directly, using appropriate parameters. We have seen, however, that it is fundamentally impossible to obtain accurate estimate of expected returns of common stocks using historical data. The standard deviation (or volatility) is just too great. Furthermore, the Markowitz mean-variance portfolio tends to be sensitive to these values. This, unfortunately, makes it essentially meaningless to computer the solution.

However, better estimates on a particular firm can be obtained in a variety of ways: (i) from detailed fundamental analyses of the firm, including an analysis of its future projects, its management, its financial condition, its competition, and projected market for its products or services, (2) as a composite of fellow analyst' conclusions and (3) from intuition and hunches based on news reports and personal experience. Such information can be systematically combined with the estimates derived from historical data to develop superior estimates.

After all, the single-period framework of Markowitz and CAPM are beautiful theories that ushered in an era of quantitative analysis and have provided an elegant foundation to support future work.

1.7 Project: Take Home Midterm

1. From internet find the adjusted stock prices for the following companies and indexes:
IMB, MicroSoft, Dell; Merck, Pfizer, Johnson & Johnson; SP500, DJ Ind., 13 Week US T. B.

   Obtain data from Jan 1st 1992 to Dec 31 2006 in the following categories

   Method: daily, weekly, monthly, annually.

   History Period: 5 year (02-06), 10 year (97-06), 15 year (92-06).

2. For each combination of period and method (total 12), find the mean return vector and the covariance matrix.

3. For each stock, each method, and each period (total 108), find confidence interval for the mean return and standard deviation. Write your answer in the form $a \pm \varepsilon$. Take confidence Level to be 68% for mean, and 90% for variance.

   From the 12 data for each company, provide your best guess to the mean return and variance that you think is reasonable. Using annual units.

4. Perform Markowitz mean-variance analysis, especially find two special funds: (i) the fund with minimum variance. (ii) the market portfolio with 5% risk-free interest rate.

   Compare the two funds from different historical data (12 of them.)

5. Assume that the annual risk free interest rate is 5%, perform the CAPM analysis. In particular, find the market portfolio and the beta value of each company.

6. Find, if possible, the total asset (stock price times total share) of each company. Compare the relative weight of the total asset (market share) and the weight of the market portfolio.

7. Provide your observations and conclusions towards the theory and the method of obtaining parameters in the model.

Chapter 2

Finite State Models

A **financial security** is *a legal contract that conveys ownership, credit, or right to ownership.*

For example, a financial security can be the ownership of a stock, the credit from a bond, or the right to ownership from an option. An option is the right, but not the obligation, to buy (or sell) an asset under specified terms.

A **derivative security**, or **contingent claim**, is *a security whose value depends on the values of other more basic securities, which in this case is called* **underlying security** *for the derivative.*

Derivative securities can be contingent on almost any variable, from the price of hogs to amount of snow falling at a certain ski resort. Perhaps this is why it has become common to refer to the underlying entity simply as the underlying.

In recent years, derivative securities have become increasingly important in the field of finance. Futures and options are now actively traded on many different exchanges. Forward contracts, swaps, and many different types of options are regularly traded outside of exchanges by financial institutions and their corporate clients in what are termed as the *over-the-counter* markets. Other more specialized derivative securities often form part of a bond or stock issue.

In this chapter, we shall study a **derivative pricing problem** which is *to determine a fair initial value of any derivative.* The difficulty is that the final value of the derivative is not known at time $t = 0$, since it generally depends on the final value of underlying asset. However, we shall assume that the final value of the underlying is a known random variable and so the set of possible final value of the asset is known. Consequently, the set of possible final values of the derivative is also known. Knowledge of this set along with the no-arbitrage principle is the key to derivative pricing.

The primary purpose of this chapter is to introduce a mathematical framework, the finite state model. The model is elegant and versatile, and applies to many kinds of important mathematical finance problems. In certain sense, finite state model to finance is like linear algebra to mathematics.

2.1 Examples

A **stock option** is *a contract between one party, called* **seller** *or* **writer***, and another party, called* **buyer***, that allows the buyer to buy from or sell to the writer in certain time limit a stock at a fixed price; an option is called a* **call option** *if the secured right is for the buyer to buy, and it is called a* **put option** *if the right is to sell. The fixed price in an option is called the* **exercise price** *or* **strike price***.*

The last date of option time period is called the **expiration data**. *An option is called a* **European option** *if the secured right can be exercised only on the expiration of the option; it is called an* **American option** *if the right can be exercised at any time on or before the expiration date.*

Notice that the values of options depend mostly on prices of the underlying stocks.

Example 2.1. Consider the a European call option on a stock with strike price E and expiration T. Let S_t be the stock price at t. Then the option has a cash value X at time T where

$$X = \max\{S_T - E, 0\}.$$

Indeed, if $S_T > E$, then one uses the option and cash E to buy one share of stock and sold it immediately at price S_T, gain a net profit of cash $S_T - E$, at time T. On the other hand, if $S_T \leqslant E$, then the option is voided automatically since there is no profit in exercising the option.

Similarly, for a European put option with strike price E and expiration T, the option has a cash value X at time T where

$$X = \max\{E - S_T, 0\}.$$

This is a typical example of derivative security, also called contingent claim, since the value of the option depends on the underlying stock. Here the stock option is a derivative security, with stock as its underlying. The central problem here is to find the current value of option; namely, determine the present value of a future payment, that depend on the stock price.

Example 2.2. Consider an American call option on a stock with strike price E and expiration T. Let S_t be the stock price at $t \in [0, T]$. Let $\tau \in [0, T]$ be the time that the option is exercise. Then the option provides at time τ the following cash value:

$$Y_\tau = \max\{S_\tau - E, 0\}.$$

Suppose the risk-free (continuously compounded) interest rate is a constant r. Then the present value X_0 and the future value X_T of the call option, if it is exercised at time τ, are respectively

$$X_0[\tau] = e^{-r\tau} Y_\tau = e^{-r\tau} \max\{S_\tau - E, 0\}, \qquad X_T[\tau] = e^{r(T-\tau)} \max\{S_\tau - E, 0\}.$$

Similarly, if an American put is exercised at time $\tau \in [0, T]$, it provides a time T cash value

$$X_T[\tau] = e^{r(T-\tau)} \max\{E - S_\tau, 0\}.$$

The problem here is to find an optimal time τ to exercise the option. Since a priori one does not exactly know the future behavior of stock price, finding optimal strategy is one of the key here.

Example 2.3. Consider the problem to price a European call option of duration $T = 1$ (month) with strike price $E = 180(\$)$ for a particular stock currently priced at 160. The implication of the option is that the option buyer can collect $\max\{S_T - 180, 0\}$ (\$) from the option writer at the expiration date $t = T$.

In the following analysis, we assume, for simplicity, that the risk-free interest rate r is zero; namely, there is no interest charge on lending and borrowing money.

1. Suppose we are certain that $S_T = 190$. Then the option is risk-free and the value of the option at time T is $S_T - E = 190 - 180 = 10$. Since the discount factor e^{-rT} is 1, its current value is also 10, so the initial value of the option is $(S_T - E)e^{-rT} = 10$. Of course, in general S_T is unknown and the state of matter is far complicated than this.

2. Suppose either $S_T = 140$ or $S_T = 200$, both with 50% chance. Then statistically, we would calculate the expected value to be

$$0.5 * \max\{140 - 180, 0\} + 0.5 * \max\{200 - 180, 0\} = 10.$$

Thus, the expected value is \$10. This is the value in the eyes of a gambler. Can really this option be sold at \$10 a piece?

3. Let's reconsider the above situation: either $S_T = 200$ or $S_T = 140$. Suppose on the market, such an option is sold for P (\$) per unit. Let's see if we can make money out of this. We consider the following portfolio:

$\mathbf{n}_0$: (i) x shares of stock, (ii) y units of the option, (iii) $-160x - yP$ (\$) of risk-free asset.

The initial value of this portfolio is zero. Now we calculate the value of this option at time T. There are two possible outcomes, so we consider the profit for each of these possibilities.

(a) If $S_T = 140$, the portfolio worths

$$V_1 = 140x + 0 * y - 160x - yP = -20x - yP.$$

(b) If $S_T = 200$, the worth of the portfolio is

$$V_2 = 200x + (200 - 180)y - 160x - yP = 40x + (20 - P)y.$$

Suppose, out of nothing, we want to get a profit V_1 in event $S_T = 140$ and profit V_2 in event $S_T = 200$. Can we do this? The answer depends on the solvability of x and y from the system

$$\begin{pmatrix} -20 & -P \\ 40 & 20 - P \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

We know that this system has a unique solution if and only the determinant of the matrix is non-zero, i.e. if and only if $P \neq \frac{20}{3}$. We now can conclude: *The price of the option should be* $P = \frac{20}{3} = \$6.67$.

Indeed, suppose $P \neq \frac{20}{3}$. Then we can set $V_1 = V_2 = 1$ and solve for x, y. This means that no matter what the market behaviors, we can make \$1 out of nothing. For example, suppose $P = 7$ (\$). Then setting $x = 1$ and $y = -3$ we see that no matter which outcome occurs, we are guaranteed to make $V_1 = V_2 = 1$ out of nothing.

In a financial market, an **arbitrage** opportunity is a chance that a positive gain can be made out of nothing. An **arbitrage-free** market is a market in which there are no arbitrage opportunities.

From this example, one sees that if the outcome of stock price S_T is narrowed down to two possible choices (i.e. either $S_T = 140$ or $S_T = 200$) then the price of the option is uniquely determined, under the no arbitrage assumption. The natural probability that each event occurs plays no role and the uncertainty is reduced to zero.

In view of the two analyses, we see the following phenomena. Suppose the option is sold at \$8.00. For a gambler, he or she would think it a good deal since the expectation value is \$10, namely, the \$8.00 option price is undervalued. On the other hand, for an economic scientist, an §8.00 option price is over-priced since its true value is \$6.67.

Example 2.4. Consider the same European call option as above: strike price $E = 180$ and expiration T. Suppose it is a common knowledge that there are only the following possible outcomes:

$$\begin{array}{ccccc} t=0 & & t=T/2 & & t=T \\ & & & & (\omega_{11})\ 200 \\ & & (\omega_1)\ 190 & \longrightarrow & \\ & & & & (\omega_{12})\ 180 \\ (\omega_0)\ 160 & \longrightarrow & & & \\ & & & & (\omega_{21})\ 170 \\ & & (\omega_2)\ 150 & \longrightarrow & \\ & & & & (\omega_{22})\ 140 \end{array}$$

That is, at current $t = 0$, we have the only event $\omega_0 : S_0 = 160$. At $t = T/2$, either event $\omega_1 : S_{T/2} = 190$ or event $\omega_2 : S_{T/2} = 150$. Similarly, at $t = T$, for each outcome ω_i occurred at time $T = T/2$, we have two sub-outcomes, denoted by ω_{ij}, $j = 1, 2$ respectively. We emphasize that ω_{ij} occurs only if ω_i occurs first. Also, we do not postulate any probability associated with each event.

We claim that *the price of the option should be set at* $P = 2.5$.

Consider, for easy explanation, the case that $P = 2$. We shall see how arbitrage exists. At time $t = 0$, out of nothing we create the following portfolio,

$\mathbf{n}_0$: *(i)* -1 *share stock, (ii) 4 unit option, (iii)* 152 *(\$) on risk-free asset — total value* $= 0$.

We shall manage it appropriately to make money. Namely, at time $t = T/2$, we monitor the market to make the following adjustment to the portfolio:

(ω_1) If ω_1 happens, the value of the portfolio $\mathbf{n}_0$ is $V_1 = -38(\$) + 4$ option. We adjust the portfolio $\mathbf{n}_0$ to a new portfolio $\mathbf{n}_1$ as follows

$\mathbf{n}_1$:*(i)* -4 *share stock, (ii) 4 option (iii)* 722 *(\$) on risk-free asset — total:* -38 *(\$)+ 4 option.*

At the time $t = T$, we have the following:

(ω_{11}) If ω_{11} happens, the value of the portfolio is $V_{12} = -4 * 200 + 722 + 4 * (200 - 180) = 2$ (\$).

(ω_{22}) If ω_{22} happens, the value of the portfolio is $V_{22} = -4 * 180 + 722 + 5 * 0 = 2$ (\$).

Hence, if ω_1 happens, one can adjusting the portfolio to make money (\$), regardless of the outcomes.

(ω_2) If ω_2 happens, the value of the portfolio $\mathbf{n}_0$ is 2 (\$) plus 4 options. Cashing the stock, investing the money in to the risk-free asset, and waiting to time T, we obtain 2 (\$) gain. Here the 4 options are automatically voided.

Thus, if the price of the option is 2 (\$) per unit, one can make money out of nothing; i.e. there is an arbitrage opportunity.

In a similar manner, if the price of option is bigger than 2.5, one can construct portfolios that make money out of nothing.

Example 2.5. Consider a Bermuda call option: The option can be exercised only at $t = T/2$ or $t = T$, with strike price $E = 180$. Assume the same dynamic behavior of stock as in Example 2.4, but half period (simple) interest rate is $R = 5\%$. Find its price.

Solution.

(1) Suppose time $t = T/2$ and we are at ω_1. Form a portfolio of x share of stock and y cash. Its current $t = T/2$ value is $190x + y$. We want this portfolio to match exactly the payment of the call option (if it is exercised at time T): $S_T x + 1.05y = \max\{S_T - 180, 0\}$. This renders to

$$200x + 1.05y = 20, \quad 180x + 1.05y = 0$$

Thus, $x = 1$, $y = -171.43$. Consequently, the option, if it is exercised at time T, has a cash value $xS_{T/2} + y = 190 - 171.43 = 18.57$ If one exercise it at $t = T/2$, it is also $190 - 180 = 10$. Hence, one should not exercise the option. Its value is $V_{T/2}(\omega_1) = \max\{10, 18.57\} = 18.57$.

(2) Similarly, if $t = T/2$ and we are at ω_2. It is easy to see that the option has value zero: $V_{T/1}(\omega_2) = 0$.

(3) Now one prepares a portfolio at time $t = 0$: x share of stock and y cash. The portfolio is made to match exactly the out come of the option at $t = T/2$: $xS_{T/3}(\omega_i) + 1.05y = V_{T/2}(\omega_i)$. Thus renders to

$$190x + 1.05y = 18.57, \qquad 150x + 1.05y = 0.$$

This gives $x = 0.46, y = -66.32$. Hence, the portfolio worth $V_0 = 160x + y = 7.96$.

In conclusion, the option should be sold at §7.96.

Example 2.6. Consider the Bermuda put option with same assumption as in the previous example. Find it price.

Solution. (1) Suppose $t = T/2$ and ω_1 happens. Then it is easy to see that the value $V_{T/2}(\omega_1)$ of the option is zero.

(2) Suppose $t = T/2$ and ω_2 happens. Form a portfolio of x share stock and y cash. We want to make it match the payment of the put option if it is exercised at T. For this, we need $xS_T + 1.05y = \max\{180 - S_T, 0\}$, i.e.

$$170x + 1.05y = 10, \qquad 140x + 1.05y = 40.$$

This gives $x = -1$ and $y = 171.38$. The value of the portfolio at $t = T/2$ is $150x + y = 21.42$. However, if one exercise the option, the option worth $180 - 150 = 30$. Hence, one should exercise the option and the value of the option is $V_{T/2}(\omega_2) = \max\{21.42, 30) = 30$.

Finally, at time $t = 0$, one makes a portfolio $xS_0 + y$ to match the payment $V_{T/2}$ of the option at $t = T/2$: $xS_{T/2} + 1.05y = V_{T/2}$. This renders to

$$190x + 1.05y = 0, \qquad 150x + 1.05y = 30.$$

We obtain $x = -0.75, y = 135.71$. Hence $V_0 = xS_0 + y = \$15.71$.

Answer. The price of the put option is \$15.71.

In summary, if we know certain combinations of the outcomes, certain derivative securities can be priced. The basic strategy is to find an appropriate portfolio and manage it optimally according to dynamics of the financial market so that at the end of day the values of the portfolio equals exactly that of the derivative security, regardless of the outcomes of the financial market. If we can find such a portfolio, the price of the derivative security is then the value of the portfolio at $t = 0$, under the mathematical finance law of no-arbitrage.

Exercise 2.1. *Consider a European call option with strike price* $E = 180$ *and expiration* T. *Suppose* $S_0 = 160$ *and either* $S_T = 190$ *or* $S_T = 140$. *Also assume that cash can be borrowed or lent at the risk-free interest rate* $R = 0.1$ *per* T *unit of time; that is, a* 1 *(\$) investment at* $t = 0$ *on the risk-free asset becomes* 1.1 *(\$) at time* T. *Find a portfolio that consists* x *share of stock and* y *(\$) cash that produces the exact payment as the option. From the portfolio, find the price of the option.*

Exercise 2.2. *Given the call option in Exercise 2.3, construct a portfolio such that one obtains a sure payment of* \$6.67 *at time* T.

Exercise 2.3. *Given the call option in Exercise 2.4, construct a portfolio such that one can make a sure payment of* $P = 2.5$ *at time* T.

Exercise 2.4. *Given the put option in Exercise 2.6, construct a portfolio such that one can make a sure payment of* $15.71 * 1.05^2 = \$17.32$ *at time* T.

Exercise 2.5. *Suppose one has the call option in in example 2.5. Construct portfolio such that one can make a sure cash value of* $(1.5)^2 * 7.96 = 8.77$ *at time* T.

Exercise 2.6. *Consider a European call option with strike price* \$180 *and expiration* T. *Calculate its price, assuming the following outcomes:*

$$
\begin{array}{cccc}
t=0 & t=T/3 & t=2T/3 & t=T \\
(\omega_0)\ 180 \longrightarrow &
\left\{\begin{array}{l}
(\omega_1)\ 190 \longrightarrow \left\{\begin{array}{l}
(\omega_{11})\ 200 \longrightarrow \left\{\begin{array}{l}(\omega_{111})\ 205\\(\omega_{112})\ 195\end{array}\right.\\
(\omega_{12})\ 185 \longrightarrow \left\{\begin{array}{l}(\omega_{121})\ 190\\(\omega_{122})\ 180\end{array}\right.
\end{array}\right.\\
(\omega_2)\ 175 \longrightarrow \left\{\begin{array}{l}
(\omega_{21})\ 180 \longrightarrow \left\{\begin{array}{l}(\omega_{211})\ 185\\(\omega_{212})\ 175\end{array}\right.\\
(\omega_{22})\ 170 \longrightarrow \left\{\begin{array}{l}(\omega_{221})\ 180\\(\omega_{222})\ 160\end{array}\right.
\end{array}\right.
\end{array}\right. & &
\end{array}
$$

2.2 A Single Period Finite State Model

1. States.

Suppose that there are a finite number of possible **states** that describe the possible outcomes of a specific investment situation. At the initial time it is known that only one of these will occur. At the end of the period, one specific state will be revealed. States describe certain physical phenomena and are constructed according to the needs. For example, we might define two weather states for tomorrow: sunny or rainy. We don't know today which of these will occur, but tomorrow this uncertainty will be resolved. Or as another example, the states may correspond to three economic events: high success, moderate success, and failure.

We use

$$\Omega = \{\omega_1, \cdots, \omega_n\}$$

to denote these states. States define uncertainty in a very basic manner. It is not even necessary to introduce probabilities of the states, although this will be done later. In an ideal case, we hope that we can find prices for derivative securities in a manner that real probabilities that each state occurs are irrelevant. Namely, by hedging, uncertainty is completely neutralized. As we shall see, artificial probabilities can be introduced on the occurrence of the states so that the price of a derivatives security is the expectation, under the artificial probability measure, of the payoff.

2. Security

A **security** is defined within the context of states as *a set of payoffs–one payoff for each possible state.*

Here again the payoffs of a security are not associated with probabilities. We denote a security by

$$\mathbf{s} = (s^1, \cdots, s^n) \in \mathbb{R}^n$$

where s^i is the payoff if state ω_i occurs.

In our earlier example, a security's payoff could be as follows: gain \$30 if it rains tomorrow and lost \$10 if it is sunny. It is not necessary to specify probabilities. This security is represented as $(30, -10)$.

The central issue here to price securities. For this, we assume that prices of a few basic securities are known and we shall derive formulas for other securities derived from these securities.

3. No-arbitrage assumption

An **arbitrage** opportunity is *a chance of earning money without investing anything.*

There are two types of arbitrages:

Type a arbitrage: *An investment produces an immediate positive reward with no future payoff.*

Type b arbitrage: *An investment has non-positive cost but has a positive probability of yielding a positive payoff and no probability of yielding a negative payoff.*

An example of type a arbitrage would be to buy two tickets for \$20 and immediately sell them at \$12 for each. An example of type b arbitrage would be a free lottery ticket–you pay nothing for the ticket, but have a chance of winning a prize.

In economic science, we shall assume that there is **no arbitrage**. As we shall see, no-arbitrage has profound consequences. Here is one of them.

Linear Pricing: *If there is no type a arbitrage, security prices are linear .*

Indeed, if $\mathbf{s}_1$ and $\mathbf{s}_2$ are securities with prices $P(\mathbf{s}_1)$ and $P(\mathbf{s}_2)$, the price of the security $\mathbf{s}_1 + \mathbf{s}_2$ must be $P(\mathbf{s}_1) + P(\mathbf{s}_2)$. For if $P(\mathbf{s}_1 + \mathbf{s}_2) < P(\mathbf{s}_1) + P(\mathbf{s}_2)$, we could purchase the combined security for $P(\mathbf{s}_1 + \mathbf{s}_2)$ and break it into $\mathbf{s}_1$ and $\mathbf{s}_2$ and sell them for $P(\mathbf{s}_1)$ and $P(\mathbf{s}_2)$, respectively. As a result we would obtain a profit $P(\mathbf{s}_1) + P(\mathbf{s}_2) - P(\mathbf{s}_1 + \mathbf{s}_2) > 0$ initially and no future payoff. Similarly, if $P(\mathbf{s}_1 + \mathbf{s}_2) > P(\mathbf{s}_1) + P(\mathbf{s}_2)$, we could buy $\mathbf{s}_1$ and $\mathbf{s}_2$ separately and then sell them together. Thus we must have $P(\mathbf{s}_1 + \mathbf{s}_2) = P(\mathbf{s}_1) + P(\mathbf{s}_2)$. We leave to the reader for the proof that $P(\alpha\mathbf{s}) = \alpha P(\mathbf{s})$ for any $\alpha \in \mathbb{R}$.

The above argument assumes an ideal functioning of the market:

securities can be arbitrarily divided into two pieces and that there are no transaction cost.

In practice these requirements are not met perfectly, but when dealing with large numbers of shares of traded securities in highly liquid markets, they are closely met.

4. State Prices

A special form of security is one that has a payoff only in one state. Thus, we define the **elementary state security** by $\mathbf{e}_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ where the 1 at the *ith* position. This security pays 1 ($) if enent ω_i occurs, and pay nothing otherwise.

Suppose we know the price for each state security. Denote $q_i = P(\mathbf{e}_i)$. Then for any security $\mathbf{s} = (s^1, \cdots, s^n)$, using linearity we have

$$P(\mathbf{s}) = P(\sum_{i=1}^{n} s_i \mathbf{e}_i) = \sum_{i=1}^{n} s^i P(\mathbf{e}_i) = \sum_{i=1}^{n} s^i q_i = (\mathbf{s}, \mathbf{q})$$

where $\mathbf{q} = (q_1, \cdots, q_n)$ and $(\cdot, \cdot)$ is the dot product.

Can we always derive the price of elementary state securities? The following result is fundamental.

Theorem 2.1 **(positive state prices theorem)** *A set of positive state prices exists if and only if there are no arbitrage opportunities.*

In mathematical language, the theorem states the following:

Suppose $\{\mathbf{s}_1, \cdots, \mathbf{s}_m\}$ *is a set of underlying securities with known prices* $P_1, \cdots, P_m$. *Then there is no arbitrage if and only if there exists a vector* $\mathbf{q} = (q_1, \cdots, q_n)$ *where* $q_i > 0$ *for each* i *such that*

$$P_j = (\mathbf{s}_j, \mathbf{q}) \qquad \forall\, j = 1, \cdots, m. \tag{2.1}$$

Before the proof, we state the following whose proof is left as an exercise.

Corollary 2.1. *(i) The state prices are unique if and only if the dimension of the vector space spanned by* $\mathbf{s}_1, \cdots, \mathbf{s}_m$ *has dimension* n.

(ii) If a security $\mathbf{s}$ *is derived from* $\mathbf{s}_1, \cdots, \mathbf{s}_n$, *e.g. a linear combination of them, then* $P(\mathbf{s}) = (\mathbf{s}, \mathbf{q})$.

Proof of Theorem 2.4. We prove it in two steps.

(a) Suppose there are positive state prices $q_i > 0$ for $\mathbf{e}_i$ for each $i = 1, \cdots, n$. We show that there is no-arbitrage. To see this, suppose a security $\mathbf{s} = (s^1, \cdots, s^n)$ can be constructed such that $s^i \geqslant 0$ for all i. That $\mathbf{q} = (q_1, \cdots, q_n)$ is called the state price vector means the price of the security is $P(\mathbf{s}) = (\mathbf{s}, \mathbf{q}) = \sum_{i=1}^{n} s^i q_i$. Since each q_i is positive, we see that $P(\mathbf{s}) > 0$ if $\mathbf{s} \neq \mathbf{0}$. Hence, there is no arbitrage possibility.

(b) Now suppose there is no-arbitrage. We show that there is at least one set of positive state prices. We shall provide two proofs. The one given here needs the following lemma and is a purely linear algebraic argument. The other, to be given later, is based on an important concept of utility function which has important applications.

Lemma 2.1. *Suppose* $\mathbf{S}$ *is a subspace of* $\mathbb{R}^n$ *and for each* $\mathbf{s} \in \mathbf{S} \setminus \{\mathbf{0}\}$, *at least one component of* $\mathbf{s}$ *is negative. Then there exists a vector* $\mathbf{y} = (y_1, \cdots, y_n) \in \mathbb{R}^n$ *such that* $y_i > 0$ *for each* i *and* $\mathbf{y}$ *is perpendicular to every vector in* $\mathbf{S}$.

We now use this lemma to complete our proof. Consider the set of all securities that have initial price zero and are derived from $\mathbf{s}_1, \cdots, \mathbf{s}_m$. This set can be written as

$$\mathbf{S} = \Big\{ \sum_{j=1}^{m} x^j \mathbf{s}_j \in \mathbb{R}^n \mid \sum_{j=1}^{m} x^j P_j = 0 \Big\}.$$

This is a subspace of $\mathbb{R}^m$. Since there is no-arbitrage, for every $\mathbf{s} \in \mathbf{S} \setminus \{\mathbf{0}\}$, at least one component of $\mathbf{s}$ is negative. Hence, there exists a positive vector $\mathbf{y} \in \mathbb{R}^n$ such that $\mathbf{y}$ is perpendicular to every vector in $\mathbf{S}$. If $P_j = 0$ for all $j = 1, \cdots, m$, setting $\mathbf{q} = \mathbf{y}$ we have for each j, $\mathbf{s}_j \in \mathbf{S}$ so that $(\mathbf{s}_j, \mathbf{y}) = 0 = P_j$ and we are done.

Now suppose there is at least one of $P_1, \cdots, P_m$ is non-zero. Without loss of generality, we assume that $P_1 \neq 0$. We consider two cases: (a) $(\mathbf{s}_1, \mathbf{y}) \neq 0$; (b) $(\mathbf{s}_1, \mathbf{y}) = 0$.

(a) Suppose $(\mathbf{s}_1, \mathbf{y}) \neq 0$. For each $j = 1, \cdots, m$, the security $\mathbf{s}_j - \frac{P_j}{P_1}\mathbf{s}_1$ has price zero so that as a vector, it is perpendicular to $\mathbf{y}$. It then follows that

$$0 = \Big(\mathbf{s}_j - \frac{P_j}{P_1}\mathbf{s}_1, \mathbf{y}\Big) = (\mathbf{s}_j, \mathbf{y}) - \frac{P_j}{P_1}(\mathbf{s}_1, \mathbf{y}) \quad \text{i.e.} \quad (\mathbf{s}_j, \mathbf{y}) = \frac{P_j}{P_1}(\mathbf{s}_1, \mathbf{y}) \quad \forall\, j = 1, \cdots, m.$$

Setting $\mathbf{q} = P_1\mathbf{y}/(\mathbf{s}_1, \mathbf{y})$ we then obtain $(\mathbf{s}_j, \mathbf{q}) = \frac{P_1}{(\mathbf{s}_1,\mathbf{y})}(\mathbf{s}_j, \mathbf{y}) = P_j$ for all $j = 1, \cdots, m$. We are done.

(b) Suppose $(\mathbf{s}_1, \mathbf{y}) = 0$. We can decompose $\mathbf{s}_1 = \hat{\mathbf{s}}_1 + \hat{\mathbf{s}}_1^\perp$ where $\hat{\mathbf{s}}_1 \in \mathbf{S}$ and $\hat{\mathbf{s}}_1^\perp \perp \mathbf{S}$. Since $\mathbf{s}_1 \notin \mathbf{S}$ (as $P_1 \neq 0$), we have $\hat{\mathbf{s}}_1^\perp \neq 0$. Take a small positive ε and consider the vector $\hat{\mathbf{y}} = \mathbf{y} + \varepsilon\mathbf{s}_1^\perp$. Since each component of $\mathbf{y}$ is positive, so is $\hat{\mathbf{y}}$ if we take ε sufficiently small, but not zero. As $\mathbf{s}_1^\perp \perp \mathbf{S}$, we also have $\hat{\mathbf{y}} \perp \mathbf{S}$. In addition, $(\mathbf{s}_1, \hat{\mathbf{y}}) = (\mathbf{s}_1, \mathbf{y} + \varepsilon\mathbf{s}_1^\perp) = \varepsilon\|\mathbf{s}_1^\perp\|^2 > 0$. Use this new $\hat{\mathbf{y}}$ and follow the same proof as in part (a) we then complete the proof. □

4. Risk-Neutral Probability

Now we introduce one of the most important idea in mathematical finance—the risk-neutral probability.

Let $\mathbf{q} = (q_1, \cdots, q_n)$ be a set of positive state prices. Set

$$\mathbf{p} = \frac{\mathbf{q}}{\sum_{i=1}^n q_i}, \qquad \mu_0 = \frac{1}{\sum_{i=1}^n q_i} - 1.$$

Then we have the following price formula

$$P(\mathbf{s}) = \frac{(\mathbf{s}, \mathbf{p})}{1 + \mu_0}. \tag{2.2}$$

Since $\mathbf{p}$ is a strongly positive vector and the sum of all its components is one, we can regard $\mathbf{p}$ as a probability measure on Ω, which we call the **risk-neutral probability**. Denote by $\mathbf{E}$ the expectation operator. Then the price of a security with final payoff $\mathbf{s}$ is

$$P(\mathbf{s}) = \frac{1}{1 + \mu_0}\mathbf{E}(\mathbf{s}).$$

We remark that if there is a risk-free asset, then μ_0 is exactly the risk-free return rate.

The risk-neutral probability has nothing to do with the actually probability that each state event occurs. It is introduced to neutralize risks. The artifice is deceptive in its simplicity; however, it has profound consequences. Indeed, the introduction of risk-neutral probability is the key to the success of Nobel prize winners Black and Scholes and Morton's theories on option pricing. The idea of using risk-neutral lays a foundation for modern economics.

6. Examples.

Example 2.7. Suppose Jazz Company needs to invest \$5,000 to host a show on August 15, 2006. If it rains, the company will lost all of its investment, otherwise get \$9,000 payoff. To ensure a guaranteed profit, the company seek for rain insurance. The insurance company, on the other hand, based on reliable information, found that the probability of rain on the day is 1/6. Hence, the insurance company set the price of ensuring \$3 payoff if it rains on Aug 15, 2006, for every dollar purchased. How much insurance should the company buy to maximize a sure return?

Suppose the company pays x dollar for insurance. Then we have the following calculation:
Initial Investment: $5000 + x$.
Payoff: (i) If it rains, the payoff is $3x$ from insurance, (ii) If sunny, the payoff is \$9,000 from the show.

Hence, we set $x = 3000$ to make a sure return of 9000. The initial cost will then be 8000. In this investment, the company get a sure return rate of $9000/8000 - 1 = 12.5\%$, risk-free.

It is worthy of noting the following:

(1) The company's risk of loss is totally neutralized. It has a guaranteed profit, regardless of the outcome. No probability is needed here. The 12.5% return rate is risk-free!.

(2) Suppose the company does not buy insurance. Then we have to use the probability of rain to analyze it expected return and variance. The return rate R is a random variable: $R = 9000/5000 - 1 = 80\%$ with probability 5/6 and $R = 0/5000 - 1 = -100\%$ with probability 1/6. Thus, the expected return rate is $\mu = 0.8 * 5/6 - 1/6 = 50\%$. The variance is $\sigma^2 = (0.8 - 0.5)^2 * 5/6 + (-1 - 0.5)^2/6 = 0.45$, standard-deviation $\sigma = \sqrt{0.45} = 66\%$. The expected return 50% is high, but the risk 66% is even higher.

(3) Now let's analyze the return of the insurance company. It is important to note that the insurance company's risk is not neutralized. It has a potential to lose \$6000 so we can assume that it needs \$6000 initial investment. The payoff is either 0 with probability 1/6 or 9000 with probability 5/6. That is, it has two possible return rates: if it rains: $R = 0/6000-1 = -1$; otherwise, $R = 9000/6000-1 = 0.5$. The expected return is $\mu = -1/6+0.5*5/6 = 25\%$, variance is $\sigma^2 = (-1-0.25)^2/6+(0.5-0.25)^2*5/6 = 0.3125$ so $\sigma = 56\%$. Consider as an investment, a 25% return rate with a 56% risk is sometimes not too bad.

Example 2.8. Let's put the previous example into the theory of a single-period state model.

We use $\Omega = \{\omega_1, \omega_2\}$ to denote the two possible states where ω_1 represents rain and ω_2 represents sunny, on Aug. 15, 2006. We have two securities:

(i) Investment on the show: $P(\mathbf{s}_1) = 5$, $\mathbf{s}_1 = (0, 9)$,

(ii) Rain insurance: $P(\mathbf{s}_2) = 1$, $\mathbf{s}_2 = (3, 0)$.

We wish to calculate the state price $\mathbf{q} = (q_1, q_2)$. Based on the two securities, we have the following system of equations

$$5 = 0 * q_1 + 9 * q_2, \qquad 1 = 3 * q_1 + 0 * q_2 \ .$$

This gives state prices: $q_1 = 1/3$ and $q_2 = 5/9$. As $q_1 + q_2 = 1/3 + 5/9 = 8/9$, we see that the price formula is

$$P(\mathbf{s}) = \frac{\frac{3}{8}s^1 + \frac{5}{8}s^2}{1 + 12.5\%} \qquad \forall \mathbf{s} = (s^1, s^2) \in \mathbb{R}^2.$$

From here we see the following:

(i) The equivalent risk-free rate is 12.5%.

(ii) The risk-neutral probability is 3/8 for ω_1 and 5/8 for ω_2. It has nothing to do with the natural probability 1/6 for ω_1 (rain) and 5/6 for ω_2 (sunny).

(iii) The company bought the security $\mathbf{s} = 1000\mathbf{s}_1 + 3000\mathbf{s}_2 = (9000, 9000)$ with initial cost

$$P(\mathbf{s}) = \frac{\frac{3}{8} * 9000 + \frac{5}{8} * 9000}{1 + 25\%} = 8000.$$

(iv) The insurance sold (shorted) the security $\mathbf{s} = -3000\mathbf{s}_2 = (-9000, 0)$ so its initial price is

$$P(\mathbf{s}) = \frac{\frac{3}{8} * (-9000) + \frac{5}{8} * 0}{1 + 25\%} = -3000;$$

namely, it has an initial income \$3000. This is not free, it bears future obligations: If if rains, pay 9000, otherwise, nothing.

(v) There is no arbitrage in the system.

A T-**bond** is a payment of unit cash at time T, nothing before that. The price of a T-bond at time $t < T$ is denoted by Z_t^T.

Example 2.9. Consider a single period of one year. There are two securities: a one (year) bond $\mathbf{s}_1$ and a two (year) bond $\mathbf{s}_2$, both having zero-coupon. Their prices are respectively $P(\mathbf{s}_1) = Z_0^1 = 0.95$ and $P(\mathbf{s}_2) = Z_0^2 = 0.90$ respectively. Suppose there are only two outcomes: ω_1 under which the price of one year bond (bought at the end of one year) is $Z_1^2(\omega_1) = 0.94$, and ω_2 under which $Z_1^2(\omega_2) = 0.96$. Calculate the risk neutral probability and the one year risk-free return rate.

Solution. At the end of one year period, the payment of 1-bond is

$$\mathbf{s}_1(\omega_1) = 1, \qquad \mathbf{s}_1(\omega_2) = 1.$$

For the 2-bond, at the end of one year period, its value is Z_1^2, since its payment is exactly the same as a new one-year bond. Hence, its payment is

$$\mathbf{s}_2(\omega_1) = Z_1^2(\omega_1), \qquad \mathbf{s}_2(\omega_2) = Z_1^2(\omega_2).$$

Now consider a generic security $\mathbf{s}$ whose payment is $\mathbf{s}(\omega_1) = v_1$ and $\mathbf{s}(\omega_2) = v_2$ where v_1, v_2 are arbitrary fixed constants. We look for x and y such that $\mathbf{s} = x\mathbf{s}_1 + y\mathbf{s}_2$, i.e. $x\mathbf{s}_1(\omega_i) + y\mathbf{s}_2(\omega_i) = v_i$ for $i = 1, 2$. This renders to the system

$$x + yZ_1^2(\omega_1) = v_1, \qquad x + yZ_1^2(\omega_2) = v_2.$$

Solving the system we obtain

$$x = \frac{Z_1^2(\omega_2)v_1 - Z_1^2(\omega_1)v_2}{Z_1^2(\omega_2) - Z_1^2(\omega_1)}, \qquad y = \frac{v_2 - v_1}{Z_1^2(\omega_2) - Z_1^2(\omega_1)}.$$

Consequently, the price of the security $\mathbf{s}$ is

$$\begin{aligned} P(\mathbf{s}) &= xP(\mathbf{s}_1) + yP(\mathbf{s}_2) = xZ_0^1 + yZ_0^2 \\ &= Z_0^1 \Big\{ \frac{Z_1^2(\omega_2) - Z_0^2/Z_0^1}{Z_1^2(\omega_2) - Z_1^2(\omega_1)}\, v_1 + \frac{Z_0^2/Z_0^1 - Z_1^2(\omega_1)}{Z_1^2(\omega_2) - Z_1^2(\omega_1)}\, v_2. \Big\} \\ &= \frac{1}{1+\mu_0} \Big\{ pv_1 + (1-p)v_2 \Big\}. \end{aligned}$$

Thus, the risk-neutral probability is given by

$$\text{Prob}(\omega_1) = \frac{Z_1^2(\omega_2) - Z_0^2/Z_0^2}{Z_1^2(\omega_2) - Z_1^2(\omega_1)} = 0.63, \qquad \text{Prob}(\omega_2) = \frac{Z_0^2/Z_0^1 - Z_1^2(\omega_1)}{Z_1^2(\omega_2) - Z_1^2(\omega_1)} = 0.37$$

The risk-free interest rate for the first period is

$$\mu_0 = \frac{1}{Z_0^1} - 1 = \frac{1}{0.95} - 1 = 0.05263.$$

Exercise 2.7. *Show that if one of the security is risk-free asset, then μ_0 in (2.2) is the risk-free return.*

Exercise 2.8. *Show that under no type a arbitrage assumption, security prices are linear.*

Exercise 2.9. *Prove Lemma 2.1. You may proceed as follows:*

1. *Denote $D = \{(r^1, \cdots, r^n) \mid \sum_{i=1}^n r_i = 1, r^1 \geqslant 0 \ \forall\, i = 1, \cdots, n\}$. Show that there exists $\mathbf{r}_* = (r_*^1, \cdots, r_*^n) \in D$ and $\mathbf{s}_* = (s_*^1, \cdots, s_*^n) \in \mathbf{S}$ such that $\|\mathbf{r}^* - \mathbf{s}^*\|$ is the distance between D and S.*

2. *Set $\mathbf{y} = \mathbf{r}_* - \mathbf{s}_*$. Show that $\mathbf{y} \perp \mathbf{S}$.*

3. *Using $(\mathbf{r}_* - \mathbf{s}_*, \mathbf{s}_*) = 0$ show that $K := \max\{r_*^i - s_*^i \mid r_i^* > 0\} > 0$.*

4. *By considering the distance from $\mathbf{s}_*$ to $\mathbf{r}_* - t\mathbf{e}_i + t\mathbf{e}_j$ show that $r_*^j - s_*^j \geqslant K$ for all $i = 1, \cdots, n$.*

Exercise 2.10. *Prove Corollary 2.1.*

Exercise 2.11. *(**Treasure Venture**) A company is seeking finance for a treasure adventure. It is estimated that there is a 0.3 probability of high success (ω_1), 0.4 probability of moderate success (ω_2) and 0.3 probability of failure (ω_3). The company hence issues the following two securities: each cost one dollar with the following pay-off:*

(i) \$3.00 if the adventure is a high success, \$1.00 if moderate success, and \$0.00 if failure.

(ii) \$6.00 if high success, \$0.00 otherwise.

Also, on the market there is a third security: (iii) a 20% risk-free return.

(a) Based on these three securities, find the state prices, the pricing formula, and risk–neutral probability.

(b) Suppose the company offer a new investment: (iv) If it is a high success, then the investment has a payoff of \$1000, otherwise refund all the money originally received (no interest). Can this investment be priced? If it can, how much it should be? How to achieve this from the existing securities?

(c) Consider the market system that just consists of these four securities. Using the CAPM model, find the market portfolio. Also, find the beta for each security.

Exercise 2.12. *Consider a system of only three stocks: S_1, S_2, S_3. Currently their unit prices are \$10, §30 and, §60, respectively. Suppose after one month, there are only the following three outcomes:*

$\omega_1 : \quad (S_1, S_2.S_3) = (11, 33, 56);$

$\omega_2 : \quad (S_1, S_2.S_3) = (11, 30, 60);$

$\omega_3 : \quad (S_1, S_2.S_3) = (11, 27, 63).$

(1) Find the state prices and neutral probabilities.

(2) Price the options with the following payoffs:

S_4 : A guaranteed right to by one share of stock S_2 for \$30;

S_5: A guaranteed right to sell one share of stock S_3 at \$60;

S_6: A guaranteed right to trade two shares of stock S_2 with one share of S_3;

S_7: A guaranteed right to either purchase one share of stock S_2 for \$30 or one share of stock S_3 for §60.

Also, for each option, by using only securities S_1, S_2, S_3, construct portfolios that achieve the same payoffs as the option.

(3) Suppose the natural probabilities associated with ω_1, ω_2 and ω_3 are all equal, being 1/3. Use CAPM model calculate the market portfolio for the system consisting of assets S_1, S_2, S_3 and their derivatives S_4, S_5, S_6, S_7.

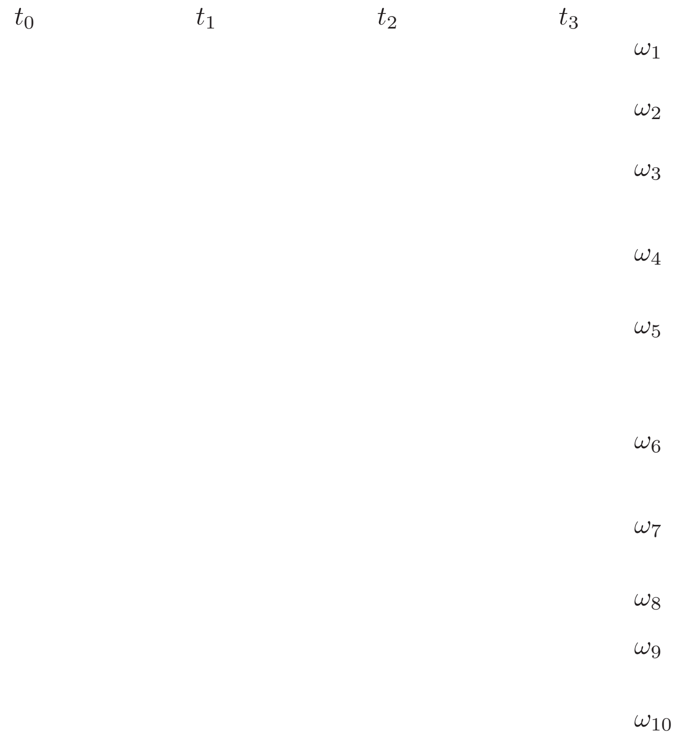


Figure 2.1: A Three period Tree Structure

2.3 Multi-Period Finite State Models

From the example presented in the previous section, we see that to price appropriately a derivative, we can try to find a portfolio and manage it optimally to yield an outcome equal to that of the derivative. We shall elaborate this idea into a mathematical framework, aiming at two aspects (i) the optimal management of a portfolio and (2) finding appropriate portfolios to price a derivative.

1. Trading Time

We begin with building a framework to represent securities in a multi-period setting in a finite number of states. For this we consider a financial market system in a time interval $[0, T]$. Assume that there are a finite number of time moments at which exchanges (trades) of assets can be made so that portfolios can be adjusted. We denote these dates by

$$\mathbf{T} = \{t_0, t_1, \cdots, t_K\}, \qquad 0 = t_0 < t_1 < \cdots < t_K = T.$$

For convenience, we use $t + \Delta t$ and $t - \Delta t$ to represent t_{i+1} and t_{i+1} respectively, when $t = t_i$.

2. State Space

The basic component of this multi-period framework is a tree structure defining a random process of state transitions. Figure 2.1 shows a three period tree structure.

The leftmost node, which we call root, represents the initial point of the process at time $t_0 = 0$. The process can then move to any of its successor nodes at time $t = t_1$. A probability can be assigned to each of the arrows. Each of the probability is non-negative and the sum of the probabilities for all arrows emanating from any particular node must be one.

The nodes of the tree can be thought of as representing various "states of the financial universe". They might be various sets of possible stock prices, or conditions of unemployment, or weather conditions that would affect agriculture and hence price of agricultural products. The graph must have enough branches to fully represent the financial problems of interest. Particular security processes are defined by assigning numerical values to the nodes.

It is always safest to make a full true, with no combined nodes, so that for any node, there is a unique path from the root to the node by following the arrows. If there are transition probabilities associated with each arrow, then the probability from any one node to any one of its successor is simply the multiples of the probabilities associated with the arrows on the path. In this way we never need to worry about possible path dependencies. For computation, on the other hand, we aggressively seek opportunities to combine nodes, so that we can devise a computationally efficient methods of solution. Here in this section we are developing theoretical tools, so for simplicity, we always assume that the tree is full, i.e., every node, except the root, has exactly one predecessor and every node before level T has at least one successor.

There are two ways to model states. One is to use a tree structure (or graph) as we described above. Another one is to first introduce a filtration defined as a sequence of σ-algebra $\{\mathcal{F}^i\}_{i=0}^K$ on a fixed set Ω begin the all possible outcomes, the subsequent one always contains its previous one, i.e. $\mathcal{F}^\tau \subset \mathcal{F}^t$ for all $\tau < t$, and then use martingales. For the convenience of explanation, we use true structure. For mathematical rigorous, we use a simplified version of filtration notation.

At each time $t \in \mathbf{T}$, the underlying economy can have various kinds of outcomes. As the example of sequential tossing of a die, the further we go, the more possible outcomes be there. Since subsequent events are based on the previous event, all these events form a tree structure. Mathematically, we can use an information structure. We collect all possible outcomes at $t = T$ by

$$\Omega = \{\omega_1, \omega_2, \cdots, \omega_{n(T)}\}$$

We can imagine each ω_i can be written as a word of T letters. At each time $t = t_i$, we can only read the first i letters, hence we can divide Ω into blocks according the first i letters. All these blocks form a partition of Ω.

A **partition** $\mathcal{P}$ of a set Ω is a collection of disjoint subsets whose union is Ω:

$$\mathcal{P} = \{B_1, \cdots, B_n\}, \qquad \emptyset \neq B_i \subset \Omega, B_i \cap B_j = \emptyset \forall\, i \neq j, \quad \bigcup_{i=1}^n B_i = \Omega$$

Refinement: *Let $\mathcal{P}$ and $\mathcal{Q}$ be any two partition of Ω. We say $\mathcal{Q}$ is a* **refinement** *of $\mathcal{P}$, written as $\mathcal{P} \prec \mathcal{Q}$ if every element in $\mathcal{Q}$ is a subset of certain element in $\mathcal{P}$.*

We now define state space as follows:

A **state space** (associated with a set Ω and trading time $\mathbf{T} = \{t_i\}_{i=1}^K$) is collection $\{\mathcal{P}_t\}_{t \in \mathbf{T}}$ of partitions of Ω that form a true structure:

$$\{\{\Omega\}\} = P_{t_0} \prec \mathcal{P}_{t_1} \prec \cdots \prec \mathcal{P}_{t_K} = \{\{\omega\} \mid \omega \in \Omega\}.$$

We remark that each block in $\mathcal{P}_t$ can be regarded as a node in the tree structure.

Example 2.10. The true structure in Figure 2.1 can be written as follows:

$$\begin{aligned}
\Omega &= \{\omega_i\ ;\ i=1,\cdots,10\},\\
\mathcal{P}_{t_0} &= \{\Omega\};\\
\mathcal{P}_{t_1} &= \{B^1_{t_1}, B^2_{t_1}\}, \quad B^1_{t_1} = \{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5\},\ B^2_{t_1} = \{\omega_6,\omega_7,\omega_8,\omega_9,\omega_{10}\},\\
\mathcal{P}_{t_2} &= \{B^1_{t_2} := \{\omega_1,\omega_2,\omega_3\}, B^2_{t_2} := \{\omega_4,\omega_5\}, B^3_{t_2} := \{\omega_6\}, B^4_{t_2} := \{\omega_7,\omega_8\}, B^5_{t_2} := \{\omega_9,\omega_{10}\}\},\\
\mathcal{P}_{t_3} &= \{B^i_{t_3}\ ;\ i=1,\cdots,10\}, \qquad B^i_{t_3} = \{\omega_i\}.
\end{aligned}$$

3. State Economy

We now consider a system with $(m+1)$ assets (such as stocks, bounds, forward contracts, etc), denoted by $\mathsf{a}^0,\cdots,\mathsf{a}^m$, where a_0 is short-term risk-free (to be explained later). Typically these assets are called the **underlying securities.** From these underlying securities, new securities such as options, swaps, futures, etc. can be constructed. The new securities are then called **derivative securities**. Our purpose is to set-up a framework, e.g. an enormously large information tree structure, that is able to accommodate financial needs of pricing derivative securities.

The state space is constructed in a way that allows one to project (precisely) values of our interest, which in the current case, are the unit prices of each asset. Hence we assume that the unit price of each individual asset is known at each of the nodes. For each $t \in \mathbf{T}$ and $B \in \mathcal{P}_t$, we use $S^i_t(B)$ to denote the unit price of one share of asset a_i. For convenience, we shall also regard S^i_t as a function on Ω so by default, we use the following notation:

$$S^i_t(\omega) = S^i_t(B) \qquad \omega \in B \in \mathcal{P}_t, \quad t \in \mathbf{T}.$$

One shall see such a convention makes difficult mathematical languages (such as martingales) easier to be understood. Also, we use vector $\mathbf{S}_t = (S^0_t,\cdots,S^m_t)$ to put all these prices in a compact form. Based on this, we define the following:

A **state economy** is *a collection* $\{\mathbf{S}_t\}_{t\in\mathbf{T}}$ *where* $\mathbf{S}_t$ *is a* $\mathbf{R}^{m+1}$ *valued random variable on* $(\Omega, \sigma(\mathcal{P}_t))$.

That $\mathbf{S}_t$ is measurable on $(\Omega,\sigma(\mathcal{P}_t))$ means that $\mathbf{S}_t$ is a constant vector (in $\mathbb{R}^{m+1}$) on each block in $\mathcal{P}_t$. This justifies our convention $\mathbf{S}_t(B) = \mathbf{S}_t(\omega)$ for every $\omega \in B \in \mathcal{P}_t$. Thus, at each node and for each asset, the price is known and unambiguous.

Thus regarding each block in $\mathcal{P}_t$ as a node in the tree structure, if B is a node of $\mathcal{P}_t$ and $\{B_1,\cdots,B_l\}$ are all those nodes in $\mathcal{P}_{t+1}$ that are emanated from B, then our tree structure is introduced to mean that if at time t the outcome B is revealed, then the asset's unit prices are given by $\mathbf{S}_t(B)$ and the next time, one and only one of the event $B_1,\cdots,B_l$ will occur, and if B_j occurs, then unit asset prices are given by $\mathbf{S}_{t+1}(B_j)$. This will become clear later whence we introduced the evaluation function.

In finance, a future value (say \$1000 cash to be received one year from now) has to be discounted to its present values. Here we introduce a short-term risk-free asset to perform this job.

Given a state space, an asset is **short-term risk-free** *if its unit price* S_t *at time t is a positive constant function on every* $B \in \mathcal{P}_{t-1}$*, i.e.* S^t *is positive and measurable on* $\sigma(\mathcal{P}_{t-1})$.

Here we use convention $\mathcal{P}_{-1} = \mathcal{P}_0 = \{\Omega\}$.

Example 2.11. Consider the tree structure in Figure 2.1 with $T = \{t_i\}_{i=0}^3$. Let $\{S_t\}_{t\in\mathbf{T}}$ be the unit price of an asset a defined by

$$\begin{aligned} S_{t_0}(\omega) &= 1 \qquad \forall\,\omega\in\Omega; \\ S_{t_1}(\omega) &= 1.1 \quad \forall\,\omega\in\Omega; \\ S_{t_2}(\omega_i) &= \begin{cases} 1.3 & i=1,\cdots,5, \\ 1.2 & i=6,\cdots,10, \end{cases} \\ S_{t_3}(\omega_i) &= \begin{cases} 1.4 & i=1,2,3 \\ 1.3 & i=4,5, \\ 1.2 & i=6, \\ 1.25 & i=7,8, \\ 1.3 & i=9,10. \end{cases} \end{aligned}$$

Then a is a short term risk-free asset since one period return of investment is know at the time of investment.

For example, at time $t = t_1$, suppose the system is at the node $B^1_{t_1} = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$. Then the return of investment in period $[t_1, t_2)$ is

$$\frac{S_{t_2}(\omega)}{S_{t_1}(\omega)} - 1 = \frac{1.3}{1.2} - 1 = 0.0933 \quad \forall\,\omega\in B^1_{t_1}.$$

Example 2.12. Consider the tree structure in Figure 2.1. Suppose $\{S_t\}_{t\in\mathbf{T}}$ is the price of an asset whose price is given by

$$S_{t_0}(\omega) = 1 \qquad \forall\,\omega\in\Omega, \qquad S_{t_1}(\omega_i) = \begin{cases} 1.3 & i=1,\cdots,5, \\ 1.2 & i=6,\cdots,10, \end{cases}$$

Then this is not a short term risk free asset, since the investment from t_0 to t_1 is unknown in advance. The return of the investment in $[t_0, t_1)$ can be either 0.2 or 0.3, depending on which outcome $B^1_{t_1}$ or $B^2_{t_1}$ occurs at time t_1.

Example 2.13. Consider the tree structure in Figure 2.1. Suppose $\{S_t\}_{t\in\mathbf{T}}$ is the price of an asset whose price is given by

$$S_{t_0}(\omega) = 1 \qquad \forall\,\omega\in\Omega, \qquad S_{t_1}(\omega_i) = \begin{cases} 1.3 & i=1,\cdots,4, \\ 1.2 & i=5,\cdots,10, \end{cases}$$

This security cannot be modelled by state space in Figure 2.1 since S_{t_1} is not $\mathcal{P}_{t_1}$ measurable. In other words, at time $t = t_1$, one does not always know the unit price of the asset.

We assume that a_0 is a short-term risk-free asset. Then for every $t \in \mathbf{T}$ and $B \in \mathcal{P}_{t-\Delta t}$ ($t-\Delta t = t_{i-1}$ if $t = t_i$), there exists a constant $r_t(B)$ such that

$$\frac{S^t_0(\omega)}{S^0_0(\Omega)} = e^{r_t(B)} \qquad \forall\omega\in B\in\mathcal{P}_{t-1}.$$

We call $r_t : \Omega \to \mathbb{R}$ (or more precisely $\mathcal{P}_{t-\Delta t} \to \mathbb{R}$) the **accumulated short-term risk-free rate**, which is measurable on $(\Omega, \sigma(\mathcal{P}_{t-\Delta t}))$ since it is a constant on every block in $\mathcal{P}_{t-\Delta t}$.

Example 2.14. Consider toss sequentially a coin three times. We can set

$$\begin{aligned}
\Omega &= \{000,001,010,011,100,101,110,100\}, \qquad \mathbf{T} = \{t_i\}_{i=0}^3,\\
\mathcal{P}_{t_0} &= \{\Omega\},\\
\mathcal{P}_{t_1} &= \{B_0, B_1\}, \qquad B_i = \{i00, i01, i10, i11\};\\
\mathcal{P}_{t_2} &= \{B_{00}, B_{11}, B_{10}, B_{11}\}, \qquad B_{ij} = \{ij0, ij1\},\\
\mathcal{P}_{t_3} &= \{\{\omega\} \mid \omega \in \Omega\}.
\end{aligned}$$

Then Ω together with $\{\mathcal{P}_{t_i}\}_{i=0}^3$ form a state space since $\mathcal{P}_{t_3} = \{\{\omega\}; \omega \in \Omega\}$ is a refinement of $\mathcal{P}_{t_2}$, a refinement of $\mathcal{P}_{t_1}$, a refinement of $\mathcal{P}_{t_0} = \{\Omega\}$.

Now consider the functions $\{S_t\}_{t\in\mathbf{T}}$ defined on Ω by

$$S_{t_0}(ijk) \equiv 1, \qquad S^0_{t_1}(ijk) = 2 * i, \qquad S_{t_2}(ijk) = 4 * i * j, \qquad S_{t_3}(ijk) = 8i * j * k \qquad \forall ijk \in \Omega.$$

Then $\{S_t\}_{t\in\mathbf{T}}$ can be unit prices of an asset. Note that S_{t_3} is not measurable on $\sigma(\mathcal{P}_{t_2})$ since it is not a constant on every block in $\mathcal{P}_{t_2}$. That is, after knowing the outcome of first two tosses of the coin, one still does not know the amount of award. Thus, the price S_t has to be calculable after all events at and before time t happened. For example, suppose S_t is the payoff of a bet. Then at time t, one knows the payoff of S_t. But before t, it may not be possible to know the exact value of S_t. Note also that the corresponding asset is not a short-term risk-free asset. The following is an example of prices of a short-term risk-free asset

$$S^0_{t_0}(ijk) = 1, \qquad S^0_{t_1}(ijk) = e^{R_0}, \qquad S^0_{t_2}(ijk) = e^{R_i}, \qquad S^0_{t_3}(ijk) = e^{R_{ij}}$$

where R_0, R_i, R_{ij} are constants. It asserts that from time $t = t_0$ to time $t = t_1$, the interest rate is R_0, whereas from any node B_i at time $t = t_1$ to time $t = t_2$, the risk-free rate is $R_i - R_0$ and from $t = t_2$ to $t = t_3$ and start from any node B_{ij}, the rate is R_{ij}. Different from ordinary assets, the prices of short-term risk–free asset are uniformly the same at their landing states.

Quote often, a state space is called an **information tree**. The deeper we go, the more information we have.

We have to emphasis that state economy is a model set-up at time $t = 0$. It does not need any information acquired in the future. In other words, it is a tabulation of the asset's prices under various kinds of outcomes regarded as possible. Typically it is constructed based on facts and theories.

3. Contingent claim.

Our goal is to price securities that are derivatives of the basic assets. By **price** we mean to determine an initial price for the derivative under the assumption that the market is free of arbitrage.

A **contingent claim** is *a real valued random variable on* $(\Omega, \sigma(\mathcal{P}^T))$. *At time* T *and under an event* $\omega \in \Omega$, *the value of a contingent claim* X *is* $X(\omega)$.

A contingent claim is quite often called a **derivative security.** Knowing the values of a claim under all possible outcomes does not necessarily tell us its current value, since an any time $t < T$, nobody knows exactly which event $\omega \in \Omega$ will actually happen, though as time progresses, the number of possible outcomes narrows down. Nevertheless, we still want to know how much does a derivative worth. As mentioned earlier, our strategy is try to find a portfolio that provides the price of a claim.

Example 2.15. Let $\{S_t^1\}_{t\in\mathbf{T}}$ and $\{S_t^2\}_{t\in\mathbf{T}}$ be the unit price of two assets a_1 and a_2 respectively. Consider a claim whose payment at time T is X given by

$$X(\omega) = \tfrac{1}{2}\{S_T^1(\omega) + S_T^2(\omega)\} \quad \forall\, \omega \in \Omega.$$

At time $t = 0$, we do not know that the payment X for sure since we do not know which event $\omega \in \Omega$ will happen at time T. Nevertheless, we know the initial price of X if $\frac{1}{2}(S_0^1 + S_0^1)$. This is so since one can form a portfolio consisting half share of each asset a_1 and a_2, pay no attention to the market until time T at which the value of the portfolio matches exactly the payment X. Of course, we need a no-arbitrage assumption to make this argument rigorous.

Next consider a payment Y at $t = T$ defined by

$$Y(\omega) = \max\{S_T^1(\omega),\ S_T^2(\omega)\} \qquad \forall\, \omega \in \Omega.$$

The is a claim since Y is known at time T. Clearly, the initial price of this claim is vary hard to calculate. We have to use combinations of assets and manage them optimally (e.g. regroup the proportions of each asset after revelation of an outcome at every trading time) to accumulate a wealth meet exactly the requirement of the contingent claim.

4. Portfolio and Trading Strategy

A **portfolio** at time $t \in \mathbf{T}$ is an $\mathbb{R}^{m+1}$ valued random variable $\mathbf{n}_t(\cdot) = (n_t^0, \cdots, n_t^m)$ on $(\Omega, \sigma(\mathcal{P}_t))$, where n_t^i is number of share of asset a_i in the portfolio.

A **trading strategy** is *a collection* $\{\mathbf{n}_t(\cdot)\}_{t\in\mathbf{T}}$ *where* $\mathbf{n}^t(\cdot)$ *is a portfolio at* $t \in \mathbf{T}$.

Consider the management of a portfolio. We use $n_t^i(\omega)$ to denote the number of units of asset a_i in the portfolio at time t and event $\omega \in \Omega$. We use a compact notation $\mathbf{n}_t = (n_t^0, n_t^1, \cdots, n_t^m)$ to denote the set of units of shares of all assets.

At each time $t \in \mathbf{T}$, we know only events in $\mathcal{P}_t$. Thus, the number of shares has to be a constant vector on every block in $\mathcal{P}_t$. In theory of measure, we say $\mathbf{n}_t(\cdot)$ is measurable on $(\Omega, \sigma(\mathcal{P}_t))$.

At each time $t \in \mathbf{T}$, one can trade assets, put in or take out money. Assume the market is perfect. Then a trade is equivalent to a liquidation followed immediately by an acquisition; that is, cash in all the assets and then purchase assets immediately according to the new weight wanted. Thus everything can be recorded by the numbers of units $\mathbf{n}_t$.

Note that at any time moment t in $\mathbf{T}$, there are two weights, one before the trading and one after. For clarity, in this chapter, all weights are **after trading**:

$$\mathbf{n}_s(\cdot) = \mathbf{n}_t(\cdot) \qquad \forall\, s \in [t, t + \Delta t), \qquad \mathbf{n}_{t-}(\cdot) := \mathbf{n}_{t-\Delta t}(\cdot).$$

We emphasize that for each $t \in \mathbf{T}$, $\mathbf{n}_t(\cdot)$ is a measurable function on $(\Omega, \sigma(\mathcal{P}_t))$, since any future information cannot be used for the design of trading strategy.

Literally, a strategy is a pre-arrangement of reactions on outcomes—if x happen, do plan X, it y happen , do plan Y, etc. One can think of it as a dynamical programming for a robot.

4. Valuation of Portfolios.

Suppose we are given a state space $(\Omega, \{\mathcal{P}_t\}_{t\in\mathbf{T}})$ and its economy $\mathbf{S}$. At each time $t \in \mathbf{T}$ and each outcome $B \in \mathcal{P}_t$, the value of a portfolio with unit $\mathbf{z} = (z^0, \cdots, z^m)$ share of assets is the $\mathbb{R}^{m+1}$ dot product of $\mathbf{z}$ and $\mathbf{S}_t(B)$:

$$\Big(\mathbf{S}_t(B), \mathbf{z}\Big) = \sum_{i=0}^{m} z^i S_t^i(B) = z^0 S_t^0(B) + \sum_{j=0}^{n} z^j S_t^j(B) \qquad \forall\, B \in \mathcal{P}_t.$$

We always normalize the risk-free asset price so $S_{t_0}^0 = 1$. Then S_t^0 is the time t value of unit capital investment on risk-free (bank deposit) asset. Hence, the value

$$\frac{(\mathbf{S}_t(B), \mathbf{z})}{S_t^0(B)} = z_0 + \sum_{i=1}^{n} z_i \frac{S_t^i(B)}{S_0^t(B)} \qquad \forall B \in \mathcal{P}_t$$

is called **the discounted value** of the portfolio; namely, the value of the portfolio translated into today's worth. One purpose of the introduction of a short-term risk-free asset here is to provide a reference for gain or loss.

Given a trading strategy $\mathbf{n}$, the value $(\mathbf{S}_t, \mathbf{n})$ of the corresponding portfolio depends on the outcome of the economy, and hence is a random variable.

The **valuation** associated with a state economy $\mathbf{S}$ is *the following family of operators* $\{\mathcal{V}_t\}_{t\in\mathbf{T}}$ *which map portfolios into real valued random variables:* $\mathcal{V}_t : \mathrm{RV}(\sigma(\mathcal{P}_t); \mathbb{R}^{m+1})) \to \mathrm{RV}(\sigma(\mathcal{P}_t), \mathbb{R})$,

$$\mathcal{V}_t[\mathbf{z}](\omega) = (\mathbf{S}_t(\omega), \mathbf{z}) \quad \forall\, \mathbf{z} \in \mathrm{RV}(\sigma(\mathcal{P}_t); \mathbb{R}^{m+1}), \omega \in \Omega, t \in \mathbf{T}.$$

Given a trading strategy $\mathbf{n}$ and state of economy $\mathbf{S}$, at time $t = t_i$, its value is

$$\begin{aligned} \text{before trading:} \qquad & \mathcal{V}_t[\mathbf{n}_{t-}] = \mathcal{V}_{t_i}[\mathbf{n}_{t_{i-1}}] = (\mathbf{S}_{t_i}, \mathbf{n}_{t_{i-1}}), \\ \text{after trading:} \qquad & \mathcal{V}_t[\mathbf{n}_t] = \mathcal{V}_t[\mathbf{n}_t] = (\mathbf{S}_t, \mathbf{n}_t). \end{aligned}$$

Note the following:

1. **Profit** of investment in time $[t_{i-1}, t_i)$: $\mathcal{V}_{t_i}[\mathbf{n}_{t_{i-1}}] - \mathcal{V}_{t_{i-1}}[\mathbf{n}_{t_{i-1}}] = (\mathbf{S}_{t_i} - \mathbf{S}_{t_{i-1}}, \mathbf{n}_{t_{i-1}})$.

   That is, the profit of investment comes from the unit price change of assets.

2. **Extra Capital** needed to perform the trading: $\mathcal{V}_t[\mathbf{n}_{t-}] - \mathcal{V}_t[\mathbf{n}_t] = (\mathbf{S}_t, \mathbf{n}_t - \mathbf{n}_{t-})$.

   That is, extra money needed to perform trading is due to the change of portfolio (the number of units of shares for each assets).

If no money is put in or taken out, the values before and after each trading are the same.

A **self-financing trading strategy** under a state economy $\mathbf{S}$ is one $\{\mathbf{n}_t\}_{t\in\mathbf{T}}$ that satisfies

$$\Big(\mathbf{S}_t(B), \mathbf{n}_{t-}(B)\Big) = \Big(\mathbf{S}_t(B), \mathbf{n}_t(B))\Big) \qquad \forall\, t \in \mathbf{T}, \qquad B \in \mathcal{P}_t.$$

Since $\mathcal{P}_{t_i}$ is a refinement of $\mathcal{P}_{t_{i-1}}$, $\mathbf{n}_{t_{i-1}}$ is constant on every block of $\mathcal{P}_{t_i}$. In the sequel, we pay attention only on trading strategies that are self-financing.

At this moment, we introduce two terminologies most often used in trading.

Roll Over. A rolling over at time t means the number of shares of each assets are kept fixed for the rest of time:

$$\mathbf{n}_s(\cdot) = \mathbf{n}_t(\cdot) \quad \forall\, s \in [t, T].$$

After taking the roll over strategy, the number of shares on each asset in the portfolio is fixed, but the portfolio's final value depends on final outcomes of the economy.

Lock in. A lock in at time t means the portfolio is liquidated and kept for the rest of the time:

$$\mathbf{n}_s(\cdot) = (n_t^0(\cdot), 0, \cdots, 0) \quad \forall s \in [t, T].$$

After lock in, the portfolios discounted value will never change.

6. Replication.

Our goal is to find portfolios that are equivalent to contingent claims. Hence, we introduce

A **replicating strategy** for a contingent claim X is *a self-financing trading strategy* $\mathbf{n}$ *such that*

$$\mathcal{V}_T[\mathbf{n}_T](\omega) = X(\omega) \qquad \forall\, \omega \in \Omega.$$

An **attainable contingent claim** is *a claim that has at least one replicating strategy.*

A state model is said to be **complete** if *every contingent claim is attainable.*

7. The Finite State Model

We summarize our discussion as follows.

A **state model** consists of the following:

1. Trading dates $\mathbf{T} = \{t_i\}_{i=0}^K$ and state space $(\Omega, \{\mathcal{P}_t\}_{t\in\mathbf{T}})$;
2. Assets (underlying securities) $\mathsf{a}_0, \cdots, \mathsf{a}_m$ and their prices $\{\mathbf{S}_t := (S_0^t, \cdots, S_m^t)\}_{t\in\mathbf{T}}$ at the states.

In a state model, portfolios, trading strategies, and evaluation of portfolios can be consequently defined (by the defaults we discussed). Finally, the problem of pricing contingent claims (i.e. derivative securities) can be studied by searching replicating strategies.

Example 2.16. Consider a three period investment, starting from 3 ($) cash. Suppose for some $\hat{\omega} \in \Omega$, $\hat{\omega}$ is the actual event that happened and the unit share price $S_t^i(\hat{\omega})$ of asset a_i at time t are observed as follows:

	$S_t^0(\hat{\omega})$	$S_t^1(\hat{\omega})$	$S_t^2(\hat{\omega})$
t_{-1}	0.9	0.9	0.9
$t_0 = 0$	1	1	1
t_1	1.1	1.2	1.3
t_2	1.2	1.1	1.6
$t_3 = T$	1.3	1.5	1.2

Find the final value of the portfolio with the following self-financing strategies:

1. Static portfolio, starting from equal capital distribution on each assets.
2. Dynamical portfolio with equal capital distribution on each assets.
3. Always invest all capital on the asset that performed the best in most recent period.
4. Invest twice capital on the best performed in most recent period and short the needed capital on the worst performed asset in most recent period.
5. Invest twice capital on the worst performed in most recent period and short the needed capital on the best performed asset in most recent period.

1. First consider a trading strategy of equal share. That is,

$$\mathbf{n}_t(\omega) = (1, 1, 1) \quad \forall \omega \in \Omega, \quad t \in \mathbf{T}.$$

This is a self-finance trading strategy and the value V_t of the portfolio at each trading time t is

$$\begin{aligned} V_{t_0}(\hat{\omega}) &= \mathcal{V}_{t_0}[\mathbf{n}_{t_0}](\hat{\omega}) = 1 + 1 + 1 = 3, \\ V_{t_1}(\hat{\omega}) &= \mathcal{V}_1[\mathbf{n}_{t_0}](\hat{\omega}) = 1.1 + 1.2 + 1.3 = 3.6, \\ V_{t_2}(\hat{\omega}) &= \mathcal{V}_{t_2}[\mathbf{n}_{t_1}](\hat{\omega}) = 1.2 + 1.1 + 1.6 = 3.9 \\ V_T(\hat{\omega}) &= \mathcal{V}_T[\mathbf{n}_{t_2}](\hat{\omega}) = 1.3 + 1.5 + 1.2 = 4.0(\$). \end{aligned}$$

2. Consider a self financing trading strategy: Equal weights at each trading time:

$$n_t^0 S_t^0 = n_t^1 S_t^1 = n_t^2 S_t^2 \quad \forall\, t \in \mathbf{T}.$$

Initially, we form a portfolio $\mathbf{n}_0 = (1, 1, 1)$. Its value at time t_1 is $V_{t_1}(\hat{\omega}) = 1.1 + 1.2 + 1.3 = 3.6$.

Even distribution on the three assets means investing (\$)1.2 on each asset. Hence, the new portfolio at time $t = t_1$ is $\mathbf{n}_{t_1} = (1.2/1.1, 1.2/1.2, 1.2/1.3)$.

The value of the portfolio at time $t = t_2$ is

$$V_{t_2}(\hat{\omega}) = \mathcal{V}_{t_2}[\mathbf{n}_{t_1}](\hat{\omega}) = \frac{1.2}{1.1} * 1.2 + \frac{1.2}{1.2} * 1.1 + \frac{1.2}{1.3} * 1.6 = 3.886.$$

Evenly distributing the capital on three asset means investing $3.866/3 = 1.2954$ on each asset. Thus the portfolio at $t = t_2$ is $\mathbf{n}_{t_2} = (1.2954/1.2, 1.2954/1.1, 1.2954/1.6)$.

The final value of the portfolio is

$$V_T(\hat{\omega}) = \mathcal{V}_T[\mathbf{n}_{t_2}](\hat{\omega}) = \frac{1.2954}{1.2} * 1.3 + \frac{1.2954}{1.1} * 1.5 + \frac{1.2954}{1.6} * 1.2 = 4.14(\$).$$

3. Suppose we use the self-financing strategy of investing all capital evenly on assets that performed the best in most recent period.

Since in the time period $[t_{-1}, t_0)$, the three asset performs equally well, we set $\mathbf{n}_0 = (1, 1, 1)$.

At time $t = t_1$, the value of the portfolio is $V_{t_1}(\hat{\omega}) = \mathcal{V}_{t_1}[\mathbf{n}_0](\hat{\omega}) = 3.6$. Since the best performed asset in $[t_0, t_1)$ is a_2, we put all money on this asset, so $\mathbf{n}_{t_1} = (0, 0, \frac{3.6}{1.3})$.

At time $t = t_2$, the value of the portfolio is $V_{t_2}(\hat{\omega}) = \mathcal{V}_{t_2}[\mathbf{n}_{t_1}](\hat{\omega}) = \frac{3.6}{1.3} * 1.6 = 4.43$. Since in $[t_1, t_2)$ period, the best performed asset is again a_2, we put all the money on a_2 by setting $\mathbf{n}_{t_2} = (0, 0, \frac{4.43}{1.6})$.

At time T, the value of the portfolio is

$$V_T(\hat{\omega}) = \mathcal{V}_T[\mathbf{n}_{t_2}](\hat{\omega}) = \frac{4.43}{1.6} * 1.1 = 3.05(\$).$$

4. Now we consider the trading strategy of shorting money equal available money on the worst performed asset and invest all acquired money on the best performed asset.

Initially, $\mathbf{n}_0 = (1,1,1)$, so $V_{t_1}(\hat{\omega}) = \mathcal{V}_{t_1}[\mathbf{n}_0](\hat{\omega}) = 1.1 + 1.2 + 1.3 = 3.6$. In $[t_0, t_1]$, the best perform asset is a_2 and the worst perform asset is a_0. Hence we short \$3.6 from asset a_0 and invest $3.6+3.6 = \$7.2$ on a_2; namely, we put $\mathbf{n}_{t_1} = (-3.6/1.1, 0, 7.2/1.3)$.

At time $t = t_2$, the value of the portfolio is

$$V_{t_2}(\hat{\omega}) = \mathcal{V}_{t_2}[\mathbf{n}_{t_1}](\hat{\omega}) = -\frac{3.6}{1.1} * 1.2 + \frac{7.2}{1.3} * 1.6 = 4.934(\$).$$

In the period $[t_1, t_2)$, the best and worst performed assets are a_2 and a_1 respectively. Hence, we short 4.934 on a_1 and invest $2 * 4.934 = 9.869$ on the second asset. That is, $\mathbf{n}_{t_2} = (0, -4.934/1.1, 9.869/1.6)$.

The final value of the portfolio is

$$V_T(\hat{\omega}) = \mathcal{V}_T[\mathbf{n}_{t_2}](\hat{\omega}) = -\frac{4.934}{1.1} * 1.5 + \frac{9.869}{1.6} * 1.2 = 0.67(\$).$$

5. Finally we consider the strategy of invest twice available capital on worst performed asset and short the needed capital on the best performed in the most recent period.

Initially, we have $\mathbf{n}_0 = (1,1,1)$ so $V_{t_1}(\hat{\omega}) = 3.6$. In $[t_0, t_1)$, the worst and best performed assets are a_0 and a_2 respectively, so $\mathbf{n}_{t_1} = (7.2/1.1, 0, -3.6/1.3)$.

AT time $t = t_2$, the value of portfolio is $V_{t_2}(\hat{\omega}) = 3.42$. The portfolio at $t = t_2$ is $\mathbf{n}_{t_2} = (0, 6.84/1.1, -3.42/1.6)$. The final value is

$$V_T(\hat{\omega}) = \mathcal{V}_T[\mathbf{n}_{t_2}](\hat{\omega}) = \frac{6.84}{1.1} * 1.5 - \frac{3.42}{1.6} * 1.2 = 6.76(\$).$$

Exercise 2.13. *Show that the set of all attainable contingent claims is a vector space.*

Exercise 2.14. *Show that each of the strategy described in example 2.16 is a self-financing trading strategy.*

Exercise 2.15. *Perform the same calculation as in example 2.16 using the following observed data:*

	$S_t^0(\hat{\omega})$	$S_t^1(\hat{\omega})$	$S_t^2(\hat{\omega})$	$S_t^3(\hat{\omega})$
t_{-1}	*0.9*	*0.89*	*0.88*	*0.87*
$t_0 = 0$	*1*	*1*	*1*	*1*
t_1	*1.1*	*1.2*	*1.3*	*1.05*
t_2	*1.2*	*1.5*	*1.6*	*1.15*
$t_3 = T$	*1.3*	*1.5*	*1.3*	*1.3*

Exercise 2.16. *Consider the game of tossing a fair coin T times. Assume $T = 3$.*

(1) Built a state model accounting all possible outcomes and the following two assets (stocks):

a_1*: one share worth \$8.00, doubles after a head (H) but reduce by half after a tail (T).*

a_2*: one share worth \$27.00, triples after each T but reduces to 1/3 after each H.*

(2) Based on the two securities a_1 and a_2, make a short-term risk free-asset $\mathsf{a}_0 = \theta_1 \mathsf{a}_1 + \theta_2 \mathsf{a}_2$. Is the asset also long-term risk-free, i.e. its future value depends only on time, but not on the outcomes?

(3) Compare the values of the following claims:

(i) X_1: Get \$8.00 for three Hs, nothing otherwise.

(ii) X_2: Get \$6.00 for at least two Hs, nothing otherwise.

(iii) X_2: Get \$2.00 multiplied by the total number of H appeared.

(iv) X_4: Get \$8 if the sequence is THT.

(v) An option gives the right, but not obligation, to by one share of a_2 at price 27 after three tosses.

2.4 Arbitrage and Risk-Neutral Probability

1. Arbitrage

For a finite state model to be well-formed, we need the no arbitrage assumption.

A state model $\{\mathbf{T}, \Omega, \{\mathcal{P}_t\}_{t\in\mathbf{T}}, \{\mathbf{S}_t\}_{t\in\mathbf{T}}\}$ *is called* **arbitrage-free** *if there does not exist any self-financing strategy* $\{\mathbf{n}_t\}_{t\in\mathbf{T}}$ *satisfying*

$$\mathcal{V}_0[\mathbf{n}_0] \leqslant 0, \qquad \mathcal{V}_T[\mathbf{n}_T](\omega) \geqslant 0 \quad \forall \omega \in \Omega, \qquad \sum_{\omega\in\Omega} \mathcal{V}_T[\mathbf{n}_T](\omega) - \mathcal{V}_0[\mathbf{n}_0](\Omega) > 0.$$

These inequality says the opposite of no-arbitrage. An arbitrage is an opportunity or the existence of a self-financing trading strategy such that there is no initial cost, i.e. $\mathcal{V}_0[\mathbf{n}_0] \leqslant 0$ and no obligation of any future payment, i.e. $\mathcal{V}_T[\mathbf{n}_T](\omega) \geqslant 0$ for all $\omega \in \Omega$, but $\sum_{\omega\in\Omega} \mathcal{V}_T[\mathbf{n}_T](\omega) - \mathcal{V}_0[\mathbf{n}_0](\Omega) > 0$ which means either $\mathcal{V}_0[\mathbf{n}_0](\Omega) < 0$, i.e. there is an initial profit, or $\sum_{\omega\in\Omega} \mathcal{V}_T[\mathbf{n}_T](\omega) > 0$ meaning there is a positive probability of getting something. In view of this, we can divide arbitrage into two types:

Type A Arbitrage: $\mathcal{V}^0[\mathbf{n}^0](\Omega)] < 0, \mathcal{V}^T[\mathbf{n}^T](\omega) \geqslant 0$ for every $\omega \in \Omega$;
Type B Arbitrage: $\mathcal{V}[\mathbf{n}^0](\Omega) \leqslant 0$, $\mathcal{V}^T[\mathbf{n}(T)](\omega) \geqslant 0$ for all $\omega \in \Omega$, $\sum_{\omega\in\Omega} \mathcal{V}^T[\mathbf{n}^T](\omega) > 0$.

If a state model has no arbitrage, then we can define price of replicable contingent claims.

Theorem 2.2 **(law of one price)** *In a arbitrage-free state model, there is no difference among the initial values of portfolios of all self-financing trading strategies that replicate a same contingent claim. Consequently, the price of any attainable contingent claim can be defined as the initial value of the portfolio of any self-financing trading strategy that replicates it.*

Proof. First by the definition of the valuation function, we see that $\mathcal{V}_t[\cdot](\omega)$ is a linear functional on portfolios. Now suppose both self-financing trading strategy $\mathbf{n}^1$ and $\mathbf{n}^2$ replicate a contingent claim X. Then $\mathcal{V}_T[\mathbf{n}_T^1] \equiv X \equiv \mathcal{V}_T[\mathbf{n}_T^2]$. It follows that for $\mathbf{n} = \mathbf{n}_1 - \mathbf{n}_2$, we have $\mathcal{V}_T[\mathbf{n}_T] \equiv 0$. Since there is no type A arbitrage, we must have $\mathcal{V}^0[\mathbf{n}^0](\Omega) = 0$. This implies that $\mathcal{V}_0[\mathbf{n}_0^1] = \mathcal{V}_0[\mathbf{n}_0^0]$; namely, the initial value of portfolio $\mathbf{n}^1$ and portfolio $\mathbf{n}^2$ are the same. This completes the proof. □

We point out that since the valuation operator $\mathcal{V}$ is linear, the pricing for all attainable contingent claims are linear also.

2. Conditional Probability

Here we introduce some popular notations.

Suppose $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $|\Omega|$ is finite. We say that $\mathbb{P}$ is **strongly positive** if

$$\mathcal{F} = 2^\Omega, \qquad \mathbb{P}(\{\omega\}) > 0 \qquad \forall\, \omega \in \Omega.$$

Suppose $B \subset \Omega$ is a measurable set and has positive measure. We use $B \cap \mathcal{F}$ to denote the set $\{B \cap A \mid A \in \mathcal{F}\}$. The **conditional probability** $\mathbb{P}(\cdot \mid B)$ is defined by

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \qquad \forall A \in \mathcal{F}.$$

This means that under the occurrence of event B, the probability that A is observed is $\mathbb{P}(A|B)$. Note that $(B, B\cap\mathcal{F}, \mathbb{P}(\cdot \mid B)$ is a probability space.

Now suppose ξ is a random variable. The expectation of ξ under the measure $\mathbb{P}(\cdot \mid B)$ is called the **conditional expectation** and is denoted by

$$\mathbf{E}(\xi \mid B) := \frac{\int_B \xi(\omega)\mathbb{P}(d\omega)}{\mathbb{P}(B)}.$$

Next suppose $\mathcal{P} = \{B_1, \cdots, \mathcal{B}_k\}$ is a partition of Ω where each B_i are measurable and has positive measure. Then $B_i \in \mathcal{P} \to \mathbf{E}(\xi \mid B_i)$ is a random variable on $\mathcal{P}$, which we can regard as a random variable on Ω. We use notation $\mathbf{E}(\xi \mid \mathcal{P})$ to denote such random variable. More precisely,

$$\mathbf{E}(\xi \mid \mathcal{P})(\omega) := \mathbf{E}(\xi \mid B) = \frac{\int_B \xi(\hat{\omega})\mathbb{P}(d\hat{\omega})}{\mathbb{P}(B)} \qquad \forall\, B \in \mathcal{P},\, \omega \in B.$$

We remark again that $\mathbf{E}(\xi \mid \mathcal{P})$ is constant on each block in $\mathcal{P}$ and hence can be regarded both as a function from Ω to $\mathbb{R}$ and a function from $\mathcal{P}$ to $\mathbb{R}$, depending on our needs and interpretations.

A **filtration** is a sequence $\{\mathcal{P}_t\}_{t\in\mathbf{T}}$ of partitions of Ω such that $\mathcal{P}_{t_{+1}}$ is a refinement of $\mathcal{P}_{t_i}$ for each $i = 0, \cdots, K-1$. It is called an **information tree** if $\mathcal{P}_0 = \{\Omega\}$ and $\mathcal{P}_T = \{\{\omega\} \mid \omega \in \Omega\}$. Suppose ξ_t is a random variable on $(\Omega, \sigma(\mathcal{P}_t))$ (e.g. ξ_t is constant on every block in $\mathcal{P}_t$). The collection $\{\xi_t\}_{t\in\mathbf{T}}$ is called a **martingale** adapted to the filtration $\{\mathcal{P}_t\}_{t\in\mathbf{T}}$ if

$$\xi_{t_i} = \mathbf{E}(\xi_{t_{i+1}}|\mathcal{P}_{t_i}) \quad \forall i = 0, \cdots, K-1,$$

that is

$$\xi_{t_i}(\omega) := \xi_{t_i}(B) = \int_B \xi_{t_{i+1}}(\hat{\omega})\, \frac{\mathbb{P}(d\hat{\omega})}{\mathbb{P}(B)} \qquad \forall\, B \in \mathcal{P}_{t_i}, \omega \in B, i = 0, \cdots, K-1.$$

In the sequel, we use notation $t + \Delta t = t_{i+1}$ when $t = t_i$.

3. Risk-Neutral Probability

The most important structure hidden in arbitrage-free state model is the risk-neutral probability.

In a state model $(\mathbf{T}, \{\mathcal{P}_t\}_{t\in\mathbf{T}}, \{\mathbf{S}_t\})$, **a risk-neutral probability** *is a strongly positive probability measure* $\mathbb{P}$ *on* $(\Omega, 2^\Omega)$ *such that* $\{\mathbf{S}_t/S_t^0\}_{t\in\mathbf{T}}$ *is a martingale; that is,*

$$\frac{\mathbf{S}_t}{S_t^0} = \mathbf{E}\Big(\frac{\mathbf{S}_{t+\Delta t}}{S_{t+\Delta t}^0} \;\Big|\; \mathcal{P}_s\Big) \qquad \forall t = t_i, t+\Delta t = t_{i+1}, i = 0, \cdots, K-1,$$

or equivalently, for each asset a_i*,* $i = 0, \cdots, m$ *and evert time* $t = 0, \cdots, T-1$*, the unit share prices of* a_i *satisfies*

$$S_t^i(B) = e^{r_t(B) - r_{t+\Delta t}(B)}\mathbf{E}(S_{t+\Delta t}^i|\mathcal{P}_t) = e^{r_t(B)-r_{t+\Delta t}(B)} \int_B S_{t+\Delta t}^i(\omega)\frac{\mathbb{P}(d\omega)}{\mathbb{P}(B)} \qquad \forall\, B \in \mathcal{P}_t.$$

We notice that $S_{t+\Delta t}^0 = e^{R_{t+\Delta}(\cdot)}$ is constant on each block B in $\mathcal{P}_t$ so that

$$\mathbf{E}\Big(\frac{\mathbf{S}_{t+\Delta t}}{S_{t+\Delta}^0} \;\Big|\; \mathcal{P}_t\Big) = \frac{1}{S_{t+\Delta t}^0}\mathbf{E}(\mathbf{S}_{t+\Delta t} \mid \mathcal{P}_t) = e^{-r_{t+\Delta t}(\cdot)}\mathbf{E}(\mathbf{S}_{t+\Delta t} \mid \mathcal{P}_t\}.$$

4. The risk-neutral transition probabilities

The collection $\{\mathbb{P}(A \mid B) \mid B \in \mathcal{P}_t, A \in \mathcal{P}_{t+\Delta t}, A \subset B, t \in \mathbf{T}\}$ are called risk-neutral **transition probabilities**. In the tree structure, $P(A \mid B)$ assigns positive probability of each transition from the node B to its successor A. One can show that to find a measure $\mathbb{P}$ it is equivalent to find all transition measures, since our tree is full, namely, from any node, there is a unique path back the root.

The risk-neutral transition probability is similar to the risk-neutral probability in the one period state model. Here we provided a little bit detail to see how we can find them.

Let's fix a time $t \in \mathbf{T}$.

If $t = T$, then $\mathbf{E}(\cdot \mid \mathcal{P}_t) = \mathbf{E}(\cdot \mid \mathcal{P}_T)$. Since $\mathcal{P}_T = \{\{\omega\} \mid \omega \in \Omega\}$, We have

$$\mathbf{E}(\xi \mid \mathcal{P}_T) = \xi$$

for every function ξ defined on Ω.

Now we assume $t < T$. Let $B \in \mathcal{P}_t$ be any block. For simplicity, let's denote $\mathbf{S}_t(B) = \mathbf{s} = (S_t^0(B), S_t^1(B), \cdots, S_t^m(B))$. Suppose the node B has n successors, denoted by $\{B_1, \cdots, B_n\}$. We use $S_{t+\Delta t}(B_j)$ to denote the value of $S_{t+\Delta t}^i$ on B_j. Also, we denote

$$e^r := \frac{S_{t+\Delta t}^0(B)}{S_t^0(B)} = \frac{S_{t+\Delta t}^0(B_1)}{S_t^0(B)} = \cdots = \frac{S_{t+\Delta t}^0(B_n)}{S_t^0(B)} = e^{r_{t+\Delta t}(B)}$$

the discount factor. Finally, we denote the transition probability from B to B_j by

$$p_j := \mathbb{P}(B_j \mid B) = \frac{\mathbb{P}(B_j)}{\mathbb{P}(B)}.$$

That $\mathbb{P}$ is risk neutral requires $\mathbf{p} := (p_1, \cdots, p_n)$ be strongly positive and satisfy the following system

$$e^r \, S_t^i(B) = \sum_{j=1}^{n} p_j S_{t+\Delta t}^i(B_j) \qquad \forall i = 0, 1, \cdots, m \tag{2.3}$$

since $S_{t+\Delta t}^0(B_1) = \cdots = S_{t+\Delta}^0(B_n) = e^r S_t^0(B)$, we see that the equation for $i = 0$ is equivalent to $\sum_{j=1}^n p_j = 1$.

From one period state model, we know that no-arbitrage implies the existence of at least one strongly positive transition measure, even when $m \geqslant n$. We shall provide more details in the next section.

Here we emphasize again that risk-neutral probability has nothing to do with the natural probability, i.e, the probability that each event B_j occurs at time $t+1$ under the condition that B occurs at time t.

5. The Pricing Formula

Once a risk-neutral probability is found, the price of any attainable contingent claims can be easily calculated.

Theorem 2.3 **(Finite State Model Pricing Formula)** *Suppose a state model admits a risk-neutral probability $\mathbb{P}$. Then the initial price of any contingent claim X is*

$$P(X) = \mathbf{E}\Big(\frac{S_0^0 X}{S_T^0}\Big) = \int_\Omega \frac{S_0^0 X(\omega)}{S_T^0(\omega)} \; \mathbb{P}(d\,\omega). \tag{2.4}$$

We remark that although risk-neutral probabilities may not be unique, the price formula is unique for all attainable contingent claims.

Proof. Let $t \in \mathbf{T} \setminus \{T\}$ be arbitrary. Pick an arbitrary $B \in \mathcal{P}_t$. Let $\{B_1, \cdots, B_n\}$ be all successors of B at time $t+\Delta t$. Denote by $(p_1, \cdots, p_n)$ the risk-neutral transition probability from B to $B_1, \cdots, B_n$.

Suppose $\{\mathbf{n}_t\}_{t\in\mathbf{T}}$ is a self-financing trading strategy. Then

$$\begin{aligned}\frac{\mathcal{V}_t[\mathbf{n}_t](B)}{S_t^0(B)} &= \frac{(\mathbf{n}_t(B), \mathbf{S}_t(B))}{S_t^0(B)} = \frac{(\mathbf{n}_t(B), \sum_{j=1}^n p_j \mathbf{S}_{t+\Delta t}(B_j))}{S_0^{t+\Delta t}(B)} = \frac{\sum_{i=1}^n p_j(\mathbf{n}_t(B), \mathbf{S}_{t+\Delta t}(B_j))}{S_{t+\Delta t}^0(B)} \\ &= \frac{\sum_{j=1}^n p_j(\mathbf{n}_{t+\Delta t}(B_j), \mathbf{S}_{t+\Delta t}(B_j)}{S_{t+1}^0(B)} = \sum_{j=1}^n p_j \frac{\mathcal{V}_{t+\Delta t}[\mathbf{n}_{t+\Delta t}](B_j)}{S_{t+\Delta t}^0(B_j)} \\ &= \frac{1}{\mathbb{P}(B)} \int_B \frac{\mathcal{V}_{t+\Delta t}[\mathbf{n}_{t+\Delta t}]}{S_{t+\Delta t}^0} \ \mathbb{P}(d\omega).\end{aligned}$$

It then follows

$$\int_B \frac{\mathcal{V}_t[\mathbf{n}_t]}{S_t^0} \mathbb{P}(d\omega) = \mathbb{P}(B) \frac{\mathcal{V}_t[\mathbf{n}_t](B)}{S_t^0(B)} = \int_B \frac{\mathcal{V}_{t+\Delta t}[\mathbf{n}_{t+\Delta t}]}{S_{t+\Delta}^0} \mathbb{P}(d\omega) \qquad \forall\, B \in \mathcal{P}_t.$$

After adding over all $B \in \mathcal{P}_t$, we the obtain

$$\int_\Omega \frac{\mathcal{V}_t[\mathbf{n}_t]}{S_t^0} \mathbb{P}(d\omega) = \int_\Omega \frac{\mathcal{V}_{t+\Delta t}[\mathbf{n}_{t+\Delta t}]}{S_{t+\Delta t}^0} \mathbb{P}(d\omega) \qquad \forall t = t_i, i = 0, 1, \cdots, K-1.$$

Therefore, for any self-financing trading strategy $\{\mathbf{n}^t\}_{t\in\mathbf{T}}$,

$$\frac{\mathcal{V}_{t_0} 0[\mathbf{n}_{t_0}]}{S_{t_0}^0} = \int_\Omega \frac{\mathcal{V}_0[\mathbf{n}_0]}{S_0^0} \ \mathbb{P}(d\omega) = \int_\Omega \frac{\mathcal{V}_{t_1}[\mathbf{n}_{t_1}]}{S_{t_1}^0} \ \mathbb{P}(d\omega) = \cdots = \int_\Omega \frac{\mathcal{V}_T[\mathbf{n}_T]}{S_T^0} \ \mathbb{P}(d\omega).$$

As we shall see in the next section, the existence of a risk-neutral probability implies the state model has no arbitrage, so that price of contingent claims is well-defined. Now if X is a attainable contingent claim and $\{\mathbf{n}_t\}_{t\in\mathbf{T}}$ is a replicating self-financing trading strategy, then the value of X is

$$P(X) := \mathcal{V}_{t_0}[\mathbf{n}_{t_0}] = S_0^0 \int_\Omega \frac{\mathcal{V}_T[\mathbf{n}_T]}{S_T^0} \ \mathbb{P}(d\omega) = \int_\Omega \frac{S_0^0 X(\omega)}{S_T^0(\omega)} \mathbb{P}(d\omega) = \mathbf{E}\Big(\frac{S_0^0 X}{S_T^0}\Big).$$

This completes the proof.

Example 2.17. Consider the tree structure in Exercise 2.6. We have

$$\Omega = \{\omega_{ijk} \mid i,j,k = 1,2\}, \qquad \mathbf{T} = \{t_0 := 0, t_1 := T/3, t_2 := 2T/3, t_3 := T\}.$$

$$\begin{aligned}\mathcal{P}_0 &= \{\omega_0\}, \qquad \omega_0 := \Omega, \\ \mathcal{P}_{t_1} &= \{\omega_1, \omega_2\}, \qquad \omega_i := \{\omega_{ijk}\ ;\ j,k = 1,2\}, \quad i = 1,2, \\ \mathcal{P}_{t_2} &= \{\omega_{11}, \omega_{12}, \omega_{21}, \omega_{22}\}, \qquad \omega_{ij} := \{\omega_{ij1}, \omega_{ij2}\}, \quad i,j = 1,2., \\ \mathcal{P}_T &= \{\{\omega\} \mid \omega \in \Omega\}.\end{aligned}$$

Denote by $S_t(\omega)$ the stock price at time t and event ω. We have

	ω_{111}	ω_{112}	ω_{121}	ω_{122}	ω_{211}	ω_{212}	ω_{221}	ω_{222}
S_{t_0}	180	180	180	180	180	180	180	180
S_{t_1}	190	190	190	190	175	175	175	175
S_{t_2}	200	200	185	185	180	180	170	170
S_T	205	195	190	180	185	175	180	160

Assume that a_0 is an risk free asset whose one period ($\Delta t = T/3$) return is $e^r = 1.02$. Hence,

$$S^0_{t_i}(\omega) = e^{i\,r} = 1.02^i \quad \forall\, i = 0, 1, 2, 3, \quad \omega \in \Omega.$$

Now consider a sub-tree $B \to \{B_1, B_2\}$. Denote by p_i the risk-neutral transition probability from B to B_i. Since $S^0_{t+\Delta t}/S^0_t = e^r = 1.02$, the system (2.3) becomes

$$1 = p_1 + p_2, \qquad e^r\; S_t(B) = p_1 S_{t+\Delta t}(B_1) + p_2 S_{t+\Delta t}(B_2).$$

Thus,

$$p(B \to B_1) = p_1 = \frac{e^r S_t(B) - S_{t+\Delta t}(B_2)}{S_{t+\Delta t}(B_1) - S_{t+\Delta t}(B_2)}, \quad p(B \to B_2) = 1 - p_1.$$

We find the following

$$\begin{aligned}
p(\omega_0 \to \omega_1) &= \frac{1.02 * 180 - 175}{190 - 175} = 0.573, \qquad p(\omega_0 \to \omega_2) = 0.427,\\
p(\omega_1 \to \omega_{11}) &= \frac{1.02 * 190 - 185}{200 - 185} = 0.587, \qquad p(\omega_1 \to \omega_{12}) = 0.413,\\
p(\omega_2 \to \omega_{21}) &= \frac{1.02 * 175 - 170}{180 - 170} = 0.85, \qquad p(\omega_2 \to \omega_{22}) = 0.15,\\
p(\omega_{11} \to \omega_{111}) &= \frac{1.02 * 200 - 195}{205 - 195} = 0.90, \qquad p(\omega_{11} \to \omega_{112}) = 0.10,\\
p(\omega_{12} \to \omega_{121}) &= \frac{1.02 * 1850 - 180}{190 - 180} = 0.87, \qquad p(\omega_{12} \to \omega_{122}) = 0.13,\\
p(\omega_{21} \to \omega_{211}) &= \frac{1.02 * 180 - 175}{185 - 175} = 0.86, \qquad p(\omega_{21} \to \omega_{212}) = 0.14,\\
p(\omega_{22} \to \omega_{221}) &= \frac{1.02 * 170 - 165}{180 - 160} = 0.67, \qquad p(\omega_{22} \to \omega_{222}) = 0.33,
\end{aligned}$$

From the transition probability, we find the risk-neutral probability

$$\begin{aligned}
\text{Prob}(\omega_{111}) &= p(\omega_0 \to \omega_1)p(\omega_1 \to \omega_{11})p(\omega_{11} \to \omega_{111}) = 0.286,\\
\text{Prob}(\omega_{112}) &= p(\omega_0 \to \omega_1)p(\omega_1 \to \omega_{11})p(\omega_{11} \to \omega_{112}) = 0.05,\\
\text{Prob}(\omega_{121}) &= p(\omega_0 \to \omega_1)p(\omega_1 \to \omega_{12})p(\omega_{12} \to \omega_{121}) = 0.206,\\
\text{Prob}(\omega_{122}) &= p(\omega_0 \to \omega_1)p(\omega_1 \to \omega_{12})p(\omega_{12} \to \omega_{122}) = 0.031,\\
\text{Prob}(\omega_{211}) &= p(\omega_0 \to \omega_2)p(\omega_2 \to \omega_{21})p(\omega_{21} \to \omega_{211}) = 0.312,\\
\text{Prob}(\omega_{212}) &= p(\omega_0 \to \omega_2)p(\omega_2 \to \omega_{21})p(\omega_{21} \to \omega_{212}) = 0.051,\\
\text{Prob}(\omega_{221}) &= p(\omega_0 \to \omega_2)p(\omega_2 \to \omega_{22})p(\omega_{22} \to \omega_{221}) = 0.043,\\
\text{Prob}(\omega_{222}) &= p(\omega_0 \to \omega_2)p(\omega_2 \to \omega_{22})p(\omega_{22} \to \omega_{222}) = 0.021,
\end{aligned}$$

Finally, consider a claim at time T whose payment is

$$X(\omega) = \begin{cases} 100 & \text{if } \; S_T(\omega) > 187, \\ 0 & \text{otherwise} \end{cases}$$

The initial price of such a claim is

$$P(X) = \mathbb{E}\Big(\frac{S^0_0 X}{S^0_T}\Big) = [1.02]^{-3}\mathbb{E}[X] = 1.02^{-3} * 100 * (0.286 + 0.05 + 0.206 + 0.31) = 54.00(\$).$$

Exercise 2.17. *Take the risk-free interest rate by* $e^r = 1.01$, *recalculate the risk-neutral (transition) probability and the initial price of the claim in Example 2.17.*

Also calculate the initial prices of European call and put options with strike price 180.

How about the initial prices of the American call and put options with strike price 180. *(Hint: In the subtree* $B \to \{B_1, B_2\}$, *the value* V_t *of the American call option is*

$$V_t(B) = \max\Big\{S_t(B) - 180, e^{-r}[p(B \to B_1)V_{t+\Delta t}(B_1) + p(B \to B_2)V_{t+\Delta t}(B_2)]\Big\}.$$

The value of the American put is

$$V_t(B) = \max\Big\{\ 180 - S_t(B), e^{-r}[p(B \to B_1)V_{t+\Delta t}(B_1) + p(B \to B_2)V_{t+\Delta t}(B_2)]\ \Big\}.$$

2.5 The Fundamental Theorem of Asset Pricing

Knowing a risk-neutral probability allows us to evaluate a contingent claim rather efficiently. Concerning the existence and uniqueness of risk-neutral probabilities, we have the following fundamental result.

Theorem 2.4 **(Fundamental Theorem of Asset Pricing)** *(1) In a multi-period finite state model, there is no arbitrage opportunity if and only if there is a risk-neutral probability. (2) In a finite arbitrage free model, the risk–neutral probability is unique if and only if the model is* **complete**.

Proof. Due to the if and only if's, we divide the proof into four parts.

1. Suppose there is a risk-neutral probability. We show that there is no arbitrage.

Let $\{\mathbf{n}_t\}_{t\in\mathbf{T}}$ be any self-financing trading strategy such that $\mathcal{V}_T[\mathbf{n}_T](\omega) \geqslant 0$ for every $\omega \in \Omega$. Then from the pricing formula (2.4) we see that $\mathcal{V}_0[\mathbf{n}_0] \geq 0$. In addition, since $\mathbb{P}$ is a strongly positive measure, we see that either $\mathcal{V}_T[\mathbf{n}_T] \equiv 0 = \mathcal{V}_0[\mathbf{n}_0]$ or $\mathcal{V}_0[\mathbf{n}_0] > 0$. Thus, there is no arbitrage.

2. Next assume that the model is arbitrage free. We show there exists a risk-neutral probability.

In the tree structure, take any one period subtree say $\{B - - > B_1, \cdots, B_k\}$ where $B \in \mathcal{P}_t$. We claim that it is also arbitrary free. Indeed, if there is an arbitrage, we do nothing before t, and if B reveals at time t, we can make free profit and lock in at time $t+\Delta t$. In this way we find there is arbitrage for the original model.

Hence, any one period subtree is arbitrage free. The theory on one-period state model then provides a risk-neutral probability for the subtree, which gives a transition probability from B to all its immediate successors. Do this for every one-period subtree we then obtain a complete set of transition probabilities, from which we obtain a risk-neutral probability for the original model. We leave details to the readers.

3. Suppose the model is arbitrage-free and complete. We show risk-neutral probability is unique.

Let $\mathbb{P}$ and $\tilde{\mathbb{P}}$ be two risk-neutral probabilities. Since the model is complete, for any elementary state security $\mathbf{e}_i$ defined by $\mathbf{e}_i(\omega_j) = \delta_{ij}$ for every $\omega_j \in \Omega$, we have, by the pricing formula,

$$P(\mathbf{e}_i) = \int_\Omega \frac{S_0^0 \mathbf{e}_i}{S_T^0} \mathbb{P}(d\omega) = \frac{S_0^0}{S_T^0(\omega_i)} \mathbb{P}(\omega_i).$$

In a similar manner, we can find that $P(\mathbf{e}_i)$ is given by the same formula with $\mathbb{P}(\omega_i)$ replaced by $\tilde{\mathbb{P}}(\omega_i)$. Hence, we must have $\mathbb{P}(\omega_i) = \tilde{\mathbb{P}}(\omega_i)$. As $\omega_i \in \Omega$ is arbitrary, we then know that $\mathbb{P} = \tilde{\mathbb{P}}$. That is, risk-neutral probabilities are unique.

4. Finally suppose there is no arbitrage and there is a unique risk-neutral probability. We want to show that the model is complete, i.e. every contingent claim has a replicating trading strategy.

To this purpose, we write Ω as $\Omega = \{\omega_1, \cdots, \omega_n\}$ and set

$$\pi_i = \frac{S_0^0 \mathbb{P}(\omega_i)}{S_T^0(\omega_i)} \qquad \forall\, i = 1, \cdots, n.$$

Suppose the model is not complete. Then the dimension of $\mathcal{M}$ of all attainable contingent claims is less that n. Consequently, the space $\tilde{\mathcal{M}} := \{(X(\omega_1)\pi_1, \cdots, X(\omega_n)\pi_n)\} \mid (x_1, \cdots, x_n) \in \mathcal{M}\}$ also has dimension less than n. Hence, there exists a non-zero vector $\mathbf{y} = (y_1, \cdots, y_n)$ in the orthogonal complement of $\bar{\mathcal{M}}$. That is,

$$0 = \sum_{i=1}^{n} y_i \pi_i X(\omega_i) = 0 \qquad \forall X \in \mathcal{M}.$$

Now take a sufficiently small positive ε and define a new measure $\mathbb{P}^*$ on Ω by

$$\mathbb{P}^*(\omega_i) = (1 + \varepsilon y_i)\mathbb{P}(\omega_i). \qquad \forall \omega_i \in \Omega.$$

Since ε is small but not zero we have a strongly positive measure $\mathbb{P}^*$ on Ω that is different from $\mathbb{P}$. We now show that $\mathbb{P}^*$ is also a risk-neutral probability, thereby derive a contradiction to the assumption that the risk-neutral probability is unique.

We use $\mathbf{E}^*$ to denote the expectation under the measure $\mathbb{P}^*$. First of all, for any $\mathbf{X} \in \mathcal{M}$, we have

$$\begin{aligned} \mathbf{E}^*(S_0^0 \mathbf{X}/S_T^0) &= \sum_{i=1}^{n} \frac{S_0^0 X_i(\omega_i)}{S_T^0(\omega_i)} \mathbb{P}^*(\omega_i) = \sum_{i=1}^{n} \frac{S_0^0 X_i(\omega_i)}{S_T^0(\omega_i)} (1 + \varepsilon y_i)\mathbb{P}(\omega_i) \\ = \quad & \mathbf{E}(S_0^0 X/S_T^0) + \varepsilon \sum_{i=1}^{n} y_i \pi_i X(\omega_i) = \mathbf{E}(S_0^0 X/S_T^0) = P(X) \qquad \forall X \in \mathcal{M}. \end{aligned}$$

Next, let $\mathbf{n}_t = (1, 0, \cdots, 0)$ for all t which corresponds to one share of short-term risk-free asset in the portfolio for all time. Then its payoff is $S_T^0(\omega)$. Take this as the contingent claim X we obtain by the definition of price that

$$S_0^0 = P(\mathbf{X}) = P(S_T^0) = \mathbf{E}^*(S_0^0) = S_0^0 \sum_{i=1}^{n} \mathbb{P}^*(\omega_i).$$

This implies that $\mathbb{P}^*$ is a probability measure.

It remains to show that $\{\mathbf{S}_t\}$ is a martingale under $\mathbb{P}^*$. For this, let $t \in \mathbf{T} \setminus \{T\}$ be arbitrary. Fix any $B \in \mathcal{P}_t$. Consider the following trading strategy. Do noting if B does not happen; if at t, B happen, borrow money from short-term risk-free asset just enough to buy one share of asset a_j, and wait one period to time $t + \Delta t$ and then lock in. This is a self-financing trading strategy, whose portfolios can be written as

$$\begin{aligned} \mathbf{n}_\tau = \mathbf{0} \quad \forall\, \omega \notin B, \tau \in \mathbf{T}, \qquad & \mathbf{n}_\tau = \mathbf{0} \quad \forall \omega \in B, \tau < t, \\ \mathbf{n}_t(B) = \Big(-\tfrac{S_t^j(B)}{S_t^0(B)}, 0, \cdots, 0, 1, 0, \cdots, 0 \Big) \qquad & (1 \text{ is at } j\text{th position}), \\ \mathbf{n}_\tau(\omega) = \Big(-\tfrac{S_t^j(B)}{S_t^0(B)} + \tfrac{S_{t+\Delta t}^j(\omega)}{S_{t+\Delta t}^0(B)}, 0, \cdots, 0 \Big) \qquad & \forall\, \omega \in B, \tau \geq t + \Delta t. \end{aligned}$$

Since the initial price of this portfolio is zero, the price of the contingent claim it replicates is also zero. Consequently,

$$0 = P(\mathcal{V}_T(\mathbf{n}_T)) = \mathbf{E}^*(S_0^0 \mathbf{n}_T/S_T) = \Big\{ -\frac{S_t^j(B)}{S_t^0(B)} \mathbb{P}^*(B) + \int_B \frac{S_{t+\Delta t}^j}{S_{t+\Delta t}^0} \mathbb{P}^*(d\omega) \Big\}$$

Hence,

$$\frac{S_t^j(B)}{S_t^0(B)} = \int_B \frac{S_{t+\Delta t}^j}{S_{t+\Delta t}^0} \frac{\mathbb{P}^*(d\omega)}{\mathbb{P}(B)} = \mathbf{E}^*\Big(\frac{S_{t+\Delta t}^j}{S_{t+\Delta t}^0} \;\Big|\; B\Big).$$

As B is arbitrary, we then have $S_t^j/S_t^0 = \mathbf{E}^*(S_{t+1}^j/S_{t+1}^0 \mid \mathcal{P}_t)$. Since t is also arbitrary, $\mathbb{P}^*$ therefore is a risk-neutral probability. Thus we obtain a contradiction. This contradiction shows that the model is complete. This completes the proof.

Most of the content here can be find from [6].

Exercise 2.18 (**A Trinomial Tree**). *A certain underlying state model is a tree where each node has three successor nodes, indexed a, b, c. There are two assets defined on this tree. At certain period the prices of the two assets are multiplied by factors, depending on the successor node. The factors are shown below:*

Security successor nodes	a	b	c
a_1	*1.2*	*1.0*	*0.8*
a_2	*1.2*	*1.3*	x

(A) Assume that $x = 1.4$. Is there a short term risk-free asset for this period?

(B) Assume that $x = 1.4$. Is it possible to construct an arbitrage?

(C) Assume that $x = 0.7$ and risk-free interest rate is 10%. Also assume that the above factors are valid for all time. Show that the model is complete. Suppose the initial unit share price of a_1 and a_2 are both \$100. Calculate the price of the following claims at $t = 2$.

(i) A contingent claim with the following payoff:

$$X(\{aa, ab, ac, ba, bb, bc, ca, cb, cc\}) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}.$$

(ii) The option, not obligation, to exchange one share of a_2 to one share of a_1,

(iii) The option, not obligation, to exchange one share of a_1 with one share of a_2.

(iv) A European call option for a_2 at strike price $K = 100$.

(v) An American put option for a_1 at strike price $K = 100$.

(vi) An Asian call option for asset a_1 with strike price $S = (S_1^0 + S_1^1 + S_1^2)/3$ where S_1^i is the price of asset a_1 at time $t = i$.

(vii) A European call option for either one share of more expansive asset, at strike price \$100.

(viii) An American put option for one share of cheaper asset at strike price \$100.

Exercise 2.19 (**Node Separation**). *Consider a short-term risk-free rate on a binomial tree. At $t = 0$, the interest rate is 10%. At $t = 1$, the rate for upper node is 10% and lower node is 5%. Trace out the growth of \$1 invested on short-term risk-free asset at $t = 0$ and rolled over at time $t = 1$. Show that a full tree is needed.*

Exercise 2.20 (**Trinomial Tree and Lattice**). *Assume that that risk-free interest rate is 10% per period and a stock price has three possibilities: (a) increases by 20%, (b) remains the same, and (c) decreases by 10%.*

(i) Construct a trinomial tree with $T = 2$.

(ii) Find all possible risk-neutral probabilities.

(iii) Find all contingent claims at $t = 2$ that are replicable.

(v) Construct a trinomial lattice with states being the stock's prices. Find all possible risk-neutral probabilities. Also find all possible contingent claims at $t = 2$ that are replicable. Are there differences between a tree model and a lattice model?

2.6 Cash Flow

A **cash flow** is *a string of capitals to be received in a set of fixed dates.* If the fixed dates are $\mathbf{T} = \{t_0, t_1, \cdots, t_K\}$ $(0 = t_0 < t_1 < \cdots < t_K = T)$ and the capital to be received at t_i is $d(t_i)$, then the cash flow can be denoted by $\{(t_0, d(t_0)), \cdots, (t_K, d(t_K))\}$ or simply $\{(t, d(t))\}_{t\in\mathbf{T}}$.

1. Present Value and Future Value

Suppose the (continuously compounded) interest rate is a constant ν. Then the balance of depositing $P(0)$ at time $t = 0$ "grows" to

$$P(t) = P(0)e^{\nu t}$$

at time t. That is to say, the present value of a future payment $P(t)$ at time t is

$$P(0) = P(t)e^{-\nu t}.$$

Thus, for a cash flow $\text{CF} = \{(t, d(t))\}_{t\in\mathbf{T}}$ where $\mathbf{T} = \{0 = t_0, t_1, \cdots, t_K = T\}$, its **present value** is

$$\text{PV(CF)} := \sum_{t\in\mathbf{T}} d(t)e^{-\nu t}.$$

Similarly, its **future value** at time T is

$$\text{FV(CF)} = \sum_{t\in\mathbf{T}} d(t)e^{\nu(T-t)} = e^{\nu T}\text{PV(CF)}.$$

Example 2.18. Consider a \$100,000 home mortgage on a 15 year term with a so-called 6%/year, but indeed $R = 0.5\%$/month interest rate. The mortgage is payed back monthly, with the first payment due at the last day of the first month after receiving the mortgage. Find the monthly mortgage payment P and the balance $M(t)$ that is needed to payoff the mortgage at time t.

(i) Let's use month as our unit time. Denote by $\mathbf{T} = \{i\}_{i=0}^{15*12=180}$ the times of mortgage payments. Then the cash flow of the home owner can be written as follows:

$$\text{CF}_1 := \{(0, 100000), (1, -P), (2, -P), \cdots, (180, -P)\}.$$

Similarly, the loaner has cash flow

$$\text{CF}_2 := \{(0, -100000), (1, P), (2, P), \cdots, (180, P)\}.$$

(ii) Let's use $M(t)$ to denote the principal that the mortgage borrower owe to the loaner at time t.

From time $t - 1$ to t, the interest on the principal is $RM(t - 1)$, minus the payment P, so new principle is

$$M(t) = M(t - 1) + RM(t - 1) - P = (1 + R)M(t - 1) - P.$$

Multiplying both sides by $[1 + R]^{-t}$ gives

$$M(t)[1 + R]^{-t} = M(t - 1)[1 + R]^{-(t-1)} - P[1 + R]^{-t}.$$

By induction, we then obtain

$$M(t)[1 + R]^{-t} \quad = \quad M(0) - \sum_{i=1}^{t} P[1 + R]^{-t} = M(0) - P\frac{1 - (1 + R)^{-t}}{R}.$$

Since $M(T) = 0$, we find that the monthly payment P and consequently the principal $M(t)$ are

$$P = \frac{R}{1 - [1+R]^{-T}} M(0), \qquad M(t) = \frac{1 - [1+R]^{t-T}}{1 - [1+R]^{-T}} M(0).$$

In our example, $R = 0.005, T = 180, M(0) = 100,000$ so the monthly payment is

$$P = 100,000 * 0.005/[1 - (1.005)^{-180}] = \$843.85$$

(iii) The present value of the mortgage borrower's cash flow, under constant interest rate $R = e^{\nu} - 1$ is

$$\begin{aligned} PV(\text{CF}) &= M(0) + \sum_{t=1}^{T} (-P)[1+R]^{-t} \\ &= M(0) - P\frac{1 - [1+R]^{-T}}{R} = 0. \end{aligned}$$

The future value of the cash flow of the mortgage loaner at time T, under the constant interest rate R, is

$$\begin{aligned} \text{FV(CF)} &= -M(0)[1+R]^{T} + \sum_{t=1}^{T} P[1+R]^{T-t} \\ &= [1+R]^{T}\Big\{P\frac{1 - [1+R]^{-T}}{R} - M(0)\Big\} = 0. \end{aligned}$$

Quite often, the interest rate is a random variable, evaluating the value of a mortgage contract is very difficult.

2. US Government Bonds

There are many kinds of government bonds. Short termed ones (0-1 year) are quite often called **bills**, whereas medium termed (1–10 year) ones are often called **notes**. They are typically issued at auctions. How to price a bond thus becomes a personal matter.

A typical US **bond** is issued with coupons that pay off semiannually the interest from the principle. For example, a 10 year \$1000 bond with 12%/year interest rate will pay \$1,000 at the last day of the tenth year, plus 20 coupons each of which pays \$60; these coupons (sent in mail in old days) are payed one at a time, with half year period, starting at the last day of the first half year. Using half year as our unit time, the cash flow can be represented by

$$\{(0, -P), (1, 60), (2, 60), \cdots, (19, 60), (20, 1060)\}$$

where P is the cost to obtain such a bond. The value of P is driven by the financial market.

When a bond is detached from coupons, it is called a **stripped bond** or **zero-coupon bond**; it is one of the simplest fundamental security used in mathematical finance. For example, when coupons are "lost" (or kept by the owner as a not "for sale" object), the bond in the above example becomes zero coupon bond in the financial market, it has the simplest case flow

$$\{(0, 0), (1, 0), \cdots, (19, 0), (20, 1000)\}.$$

Of course, each coupon in the above example can also be regarded as a zero-coupon bond. The mathematical usage of a zero-coupon bond is that it represents an absolute future value, regardless of inflation or deflation. It is quite popular to use zero-coupon bond as collateral in many financial arrangements.

Example 2.19. Consider a financial system in which there are two kinds of zero-coupon bonds available:

B_1. A half year zero-coupon bond of face value \$100 currently sold at \$96;

B_2. A one year zero-coupon bond of face value \$100 currently sold at \$90.

Suppose we have a total of \$10,000, 10 year term, bonds which bear 12%/year interest payed semiannually. These bonds are issued 9 years ago. How much do they worth now?

To solve the problem, first we notice the cash flow of the security (the total of these 10 year bonds):

$$\{(0, -P), (1, 600), (2, 10600)\}.$$

We want to replicate this cash flow by that from B_1 and B_2. The cash flow of B_1 and B_2 can be represented as

$$\text{CF}_1 := \{(0, -96), (1, 100), (2, 0)\}, \quad \text{CF}_2 = \{(0, -90), (1, 0), (2, 100)\}.$$

We now use the principle of replicating portfolio and no arbitrage. Suppose x units of B_1 and y units of B_2 can produce the cash flow of B. Then we need

$$P = x * 96 + y * 90, \qquad x * 100 = 600, \qquad y * 100 = 10600.$$

Thus, $x = 6$ units of bond B_1 plus $y = 106$ units of bond B_2 will produce the exact payment of B. Hence, the present value of these 9 year old ten year term bonds is

$$P = 6 * 96 + 106 * 90 = \$10116.$$

Finally, if the last interest (\$600) from these bonds has not been collected yet, then they worth $10116 + 600 = \$10716$.

In an actual financial market, interest rates fluctuate with time. To make sense of them and to derive necessary information on short term risk-free interest rates needed in a state model is a complicated matter. We omit detailed study here.

3. Pricing a Security with Dividend

Suppose we are to price a derivative security which has a cash flow $\{(t, d(t, \cdot)\}_{t\in\mathbf{T}}$ where d is a random variable (say the dividend from a stock or bond). In a complete finite state model, such a problem can be easily handled.

Consider the time from t to $t+\Delta t$. Suppose at time t, the state is at $B \in \mathcal{P}_t$. Denote its descendent states by $\{B_1, \cdots, B_n\}$, i.e. $B = \cup_{i=1}^n B_i$ and $B_i \in \mathcal{P}_{t+\Delta t}$. Let $\mathbf{p} = (p_1, \cdots, p_n)$ be the risk-neutral transition probability from state B to states $(B_1, \cdots, B_n)$.

Let $V_t(\omega) = V_t(B)$ be the value of the derivative security at time t, after the dividend payment. Then by the pricing formula, its value can be calculated by

$$V_t(B) = e^{r_t(B)-r_{t+\Delta t}(B)} \sum_{i=1}^n \Big\{ d(t+\Delta t, B_i) + V_{t+\Delta t}(B_i) \Big\} p_i.$$

Exercise 2.21. *Using the two bonds in Example* 2.18 *calculate the (market expected) short term interest rates, for the first half and second half year respectively. Construct portfolios that support your conclusion.*

Exercise 2.22. *Suppose a stock, currently \$100 per share, pays a fixed dividend of* $d = \$1$ *per share per month. Assume that the stock price (after the dividend) obeys a binomial tree model: after each month, its price increases either by a factor* $u = 1.04$ *or a factor of* $d = 0.94$*. Suppose the risk-free interest rate is* $R = e^{\nu_0} - 1 = 0.01$ *per month. Calculate the price of a European put option with strike price* $K = \$100$ *and 3 month duration. Construct a replication portfolio to support your conclusion.*

Exercise 2.23. *Discuss the difference between using zero-coupon bonds and using cash as a collateral in keeping an account (with zero interest) in a stock market performing trades on futures.*

Exercise 2.24. *Consider a 30-year mortgage of* 10000 *with monthly interest rate* $R = e^{\nu} - 1 = 0.7\%$*. Find the monthly payment. Also use spreadsheet tabulate the balance at any time moment.*

Redo the calculation for $R = 0.6\%, 0.5\%, 0.4\%$ *respectively.*

Exercise 2.25. *Consider a 30-year mortgage of* 10000 *with monthly interest rate* $R = e^{\nu} - 1$*. Find the monthly payment. Assume that the zero coupon* T*-bond is sold at* $Z_0^T = e^{-0.005T - 0.00001T^2}$ *(*T *measured in month). Find the present value of the cash flow of the mortgage payment for* $R = 0.6\%$ *and* $R = 0.5\%$ *respectively. Also fine the critical* R *such that the present value of the mortgage payment is equal to* 10000.

Hint: Since one can by a time T *payment of* P *at current price of* PZ_0^T*, the present value of the mortgage payment is*

$$PV(CF) = \sum_{i=1}^{360} PZ_{t_0}^{t_i}.$$

Chapter 3

Asset Dynamics

True multi-period investments fluctuate in values, distribute random dividends, exist in an environment of variable interest rates, and are subject to a continuing variety of uncertainties. By asset dynamics it means the change of values of assets with time. Two primary model types are used to represent asset dynamics: trees and stochastic processes.

Binomial tree models are finite state models based on the assumption that there are only two possible outcomes between each single period. Consider a binomial tree model for a stock price. If we use U for up and D for down, then in a n-period binary tree model, each state can be represented by a word of n letters of only two symbols "U" and "D". At the final time moment T, there are 2^n possible states. Quite often tree branches can be combined to form lattices. For example, if each U and D represent the increment of stock price by a factor of $u > 1$ and $d < 1$, respectively. Then after k ups and $n - k$ downs, the stock price at time T is $u^k d^{n-k}$ fold of its initial price. While there are C_n^k many ways to reach this price, if only the stock prices are relevant to the problem, then we can combine all those nodes which give the same price. In this ways, the 2^n binary states can be replaced by an equivalent $n + 1$ states, resulting the commonly used binomial lattice model.

Stochastic process, or continuum model, on the other hand, are more realistic than binomial tree models in the sense that they have a continuum of possible stock prices at each period, not just two. The continuum models allow sophisticated analytical tools get involved so some problems can be solved analytically, as well as computationally. They also provide the foundation for constructing binomial lattice models in a clear and consistent manner (once the necessary mathematical tools are assessed). Stochastic process, particularly the Ito process, models are fundamental to dynamic problems.

Binomial tree models are conceptual easier to understand and analytically simpler than the Ito process. They provide an excellent basis for computational work associated with investment problems. We shall present a binary tree model known as the Cox-Ross-Rubinstein (CRR) Model [3] which is a specific example of finite state model present in the previous chapter.

From the CRR model, we shall take a limit to derive the famous Black–Scholes model [1]. Historically the paper of Black–Scholes came first and initialed the modern approach to option's evaluation. They discovered the rational option price derived from risk-neutral probability, an astonishing consequence of the seemingly trivial no-arbitrage assumption. Another earlier significant contribution was Merton [18, 17]. The simplified approach using binomial lattice was first presented by Sharpe in [28] and later developed by Cox, Ross, Robinstein [3] and also Rendleman and Bartter [23].

3.1 Binomial Tree Model

To define a binomial tree model, a basic period length of time is established (such as one week or one day). According to the traditional Cox–Ross–Robinstein (CRR) model, if the price is known at the beginning of a period, the price at the beginning of the next period is only one of two possible values—a multiple of u for up and a multiple of d for down, here $u > 1 > d > 0$. The probabilities of these possibilities are q and $1 - q$ respectively. Therefore, if the initial price is S, at the beginning of nth period, there are only $n + 1$ possible stock prices, $Su^k d^{n-k}$, $k = 0, 1, \cdots, n$ with a binomial probability distribution of parameter $q \in (0, 1)$. The state model based on a tree of size $\sum_{i=0}^{n} 2^i = 2^{n+1} - 1$ can be collapsed to a binary lattice $\{(t_k, u^i d^{k-i}) \mid k = 0, 1, \cdots, i = 0, \cdots, k\}$ of size $\sum_{i=0}^{n}(i + 1) = (n + 1)(n + 2)/2$. From each node, there are two outgoing arrows, one for up and one for down. For a tree, each node (except root) has only one incoming arrow; for lattice, each node (except the root and the first level) has two incoming arrows. Hence, for a tree structure history is unique, whereas for lattice, one cannot trace the history. In applications, if history is not needed, using lattice saves computation time. For most cases, a tree structure is much clearer and much more versatile than a lattice structure and therefore is strongly recommended.

		uuu			u^3
	uu			u^2	
u		uud udu duu	u		du^2
	ud du			ud	
d		udd dud ddu	d		ud^2
	dd			d^2	
		ddd			d^3

Statistically one can gather historical data to estimate two of the most important parameters: the expected return $\mathbf{E}(R)$ and variance $\mathbf{Var}(R)$. Hence, in applying the theoretical model to reality, the model parameter (u, d, q) has to satisfy the **matching condition**

$$qu + (1 - q)d = 1 + \mathbf{E}(R), \qquad q(1 - q)(u - d)^2 = \mathbf{Var}(R).$$

There are three parameters for two equations, so one of the parameter is free. One can show that if the single period is sufficiently short, then the choice of the free parameter is not very much relevant. In application, one adds in one of the following equations to fix the parameter:

(i) $q = 1/2$; (ii) $ud = 1$; (iii) $p = 1/2$

where p is the risk-netural probability (to be explained later). Each choice has its own advantages.

Here we shall take a simpler but more fundamental and theoretical approach than the general binomial tree or lattice approach described above. We shall simulate the stock price by using the digitized **Brownian motion**: in each time period, the position of a particle either move to the left or to the right by exactly one unit length, both with probability 1/2. Such digitized Brownian motion seems very special, nevertheless, in its limit, it can form most of the known stochastic processes. This phenomena has its root from the central limit theorem which asserts that almost all averages of i.i.d random variables are empirically normally distributed.

1. Trading Dates.

The time interval in our consideration is $[0, T]$ which is divided into n subinterval of equal length. Times of trading are at the end points of these subintervals. Hence, we set

$$\mathbf{T} = \{t_0, t_1, \cdots, t_n\}, \qquad t_i = i\Delta t \quad \forall\, i = 0, 1, \cdots, n, \qquad \Delta t = \frac{T}{n}.$$

In most applications, people use Excel spreadsheet, taking n ranging from 5 to 100; that is, Δt can be one day, one week, one month, or even one year. One keeps in mind that without accurate input (e.g. evaluation of important parameters), taking large n but not the limit of $n \to \infty$, may not help much. Nevertheless, for mathematical beauty and deep theoretical analysis, we would like to take the limit as $\Delta t \to 0$ to obtain continuous models, for which, calculus will be very useful.

2. Assets.

In our consideration there are two assets:

a_0 : a risk-free asset with constant continuous compounded interest rate ν_0;

a_1 : a security (e.g. stock) whose unit share price is S^t, a random variable satisfying

$$S^0 = S, \qquad S^{t_i} = S^{t_{i-1}} e^{\nu\Delta t + \sigma\sqrt{\Delta t}\,\epsilon_i} \qquad \forall\, i = 1, 2, \cdots, n$$

where ϵ_i is a random variable with probability distribution

$$\text{Prob}(\epsilon_i = 1) = \text{Prob}(\epsilon_i = -1) = \tfrac{1}{2} \qquad \forall\, i = 1, 2, \cdots, n.$$

It is assumed that $\epsilon_1, \epsilon_2, \cdots, \epsilon_n$ are independent, identically distributed random variables.

The random variable $\nu\Delta t + \sigma\sqrt{\Delta t}\ \epsilon_i$ is the continuously compounded return rate of the stock in a single period. The conventional one period return rate R_i introduced in the mean-variance theory is given by $R_i := e^{\nu\Delta + \sigma\sqrt{\Delta t}\epsilon_i} - 1$. Hence, in a single period, the expected return rate $\mathbf{E}(R_i)$ and variance $\mathbf{Var}(R_i)$ are given by

$$\mathbf{E}(R_i) = \tfrac{1}{2}e^{\nu\Delta t + \sigma\sqrt{\Delta t}} + \tfrac{1}{2}e^{\nu\Delta + \sigma\sqrt{\Delta t}} - 1 = \Big(\nu + \frac{\sigma^2}{2}\Big)\Delta t + O(\Delta t^{3/2}),$$
$$\mathbf{Var}(R_i) = \tfrac{1}{4}(e^{\nu\Delta t + \sigma\sqrt{\Delta t}} - e^{\nu\Delta t - \sigma\sqrt{\Delta t}})^2 = \sigma^2\Delta t + O(\Delta t^{3/2}).$$

It is very important to notice that

when interests are compounded continuously, the expected growth rate ν differ from the expected return rate μ by approximately half of the variance.

3. State Space

Base on the behavior of the stock price, we build a state space as follows. We use

$$\Omega = B^n := \{(z_1, \cdots, z_n) \mid |z_i| = 1 \quad \forall\, i = 1, \cdots, n\}, \qquad B = \{-1, 1\}.$$

The information tree $\{\mathcal{P}^t\}_{t\in\mathbf{T}}$ is then defined by

$$\mathcal{P}^{t_i} = \{\{(z_1, \cdots, z_i) \times B^{n-i}\} \mid |z_k| = 1 \text{ for all } k = 1, \cdots, i\}, \qquad i = 0, 1, \cdots, n.$$

4. State Economy.

For the risk-free asset, its unit share price S_0^t at time t is easily calculated to be

$$S_0^t(\omega) = S_0^0 e^{\nu_0 t} \quad \forall\omega \in \Omega,\, t \in \mathbf{T}.$$

The stock price S^t can be calculated by, for every $\omega = (z_1, \cdots, z_n) \in \Omega$,

$$S^{t_k}(\omega) = S^{t_{k-1}} e^{\nu\Delta t + \sigma\sqrt{\Delta t}\, z_k} = \cdots = S e^{\nu t_k + \sigma\sqrt{\Delta t} \sum_{i=1}^k z_i} \qquad \forall k = 1, \cdots, n.$$

We remark that according to our construction, the natural probability is

$$\text{Prob}(\{\omega\}) = 2^{-n} = \frac{1}{|\Omega|} \qquad \forall\, \omega \in \Omega.$$

Consequently, using combinatory, for any $k = 0, 1, \cdots, n$,

$$q_k := \text{Prob}\Big(S^T = S e^{\nu T + \sigma\sqrt{\Delta t}(2k-n)}\Big) = 2^{-n} \begin{pmatrix} n \\ k \end{pmatrix} = 2^{-n} C_n^k = \frac{1}{2^n} \frac{n!}{k!(n-k)!},$$

That is to say,

$$\frac{\ln S^T - \nu T}{\sigma\sqrt{\Delta t}} \text{ is binomially distributed .}$$

We remark that the binary tree can be collapsed to binary lattice by combing the nodes with same sum $\sum_{i=1}^k z_i$, for every $k = 1, \cdots, n$. This is allowed as far as only spot stock prices are concerned.

Here we remark that the possibility of collapsing of a binomial tree model to a binary lattice model relies on the constancy of μ and σ. In sophisticated models, both μ and σ are functions of S^t and t and one realizes it to be very hard to collapse a binary tree model to a binary lattice. Nevertheless, one can use a trinomial lattice model.

5. Risk-Neutral Probability.

First we calculate the risk-neutral transition probability. For any node $(z_1, \cdots, z_k) \times B^{n-k} \in \mathcal{P}^{t_k}$, there are two immediate successors, $(z_1, \cdots, z_k, 1) \times B^{n-k-1}$ and $(z_1, \cdots, z_k, -1) \times B^{n-k-1}$. Denote by p the probability for up and by $1 - p$ the probability for down. It is easy to derive the equation for p:

$$e^{\nu_0 \Delta t} = p\, e^{\nu\Delta t + \sigma\sqrt{\Delta t}} + (1-p)\, e^{\nu\Delta t - \sigma\sqrt{\Delta t}} \qquad \Longrightarrow \qquad p \;=\; \frac{e^{(\nu_0 - \nu)\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}.$$

For the model to be good, i.e., arbitrage free, we need $0 < p < 1$. This is equivalent to

$$\sigma > |\nu - \nu_0|\sqrt{\Delta t}.$$

Under this condition, we have unique risk-neutral probability. That is, the state model is arbitrage free and complete.

From the transition probabilities, we can derive the risk-neutral probability

$$\mathbb{P}(\{\omega\}) = p^{\frac{1}{2}\sum_{i=1}^{n}(1+z_i)}(1-p)^{\frac{1}{2}\sum_{i=1}^{n}(1-z_i)} \qquad \forall \omega = (z_1, \cdots, z_n) \in \Omega.$$

From which, we can also calculate the risk-neutral probability distribution of S^T:

$$p_k := \mathbb{P}\Big(\{\omega \mid S^T(\omega) = Se^{\nu T + \sigma\sqrt{\Delta t}(2k-n)}\}\Big) = \frac{n!\, p^k(1-p)^{n-k}}{k!\,(n-k)!}, \quad k = 0, \cdots, n.$$

As we know from the previous chapter, in pricing contingent claims, one has to use the risk-neutral probability distribution. This is a revolutionary idea of Black and Scholes. In old days, people used natural probabilities pricing derivative securities, resulting mismatch with reality.

5. Pricing Formula

Suppose we have a derivative security whose payoff is given by

$$X = f(S^T) \quad \text{only at time} \quad T$$

where $f : \mathbb{R} \to \mathbb{R}$ is a given smooth function. Then by our pricing formula, the price for the derivative security is

$$\begin{aligned} P(X) &= \mathbf{E}\Big(\frac{S_0^0 X}{S_0^T}\Big) = e^{-\nu_0 T}\mathbf{E}(X) = e^{-\nu_0 T}\mathbf{E}(f(S^T)) = e^{-\nu_0 T}\sum_{\omega\in\Omega} f(S^T(\omega))\ \mathbb{P}(\{\omega\}) \\ &= \sum_{k=0}^{n} p_k\, e^{-\nu_0 T}\ f(Se^{\nu T + \sigma\sqrt{\Delta t}(2k-n)}) \\ &= \sum_{k=0}^{n} \frac{n!\ \ p^k(1-p)^{n-k}}{k!(n-k)!}\ e^{-\nu_0 T}\ f(Se^{\nu T + \sigma\sqrt{T}\ (2k-n)/\sqrt{n}}). \end{aligned}$$

We can summarize our calculation as follows:

Theorem 3.1 **(Contingent Claim Price Formula)** *Assume that the risk-free continuously compounded interest rate is ν_0 and the log of the underlying security price obeys the digtized Brownian Motion with mean $\nu\Delta t$ and variance $\sigma^2\Delta t$:*

$$\ln S^{t_k} - \ln S^{t_{k-1}} = \nu\Delta t + \sigma\sqrt{\Delta t}\ \epsilon_k, \qquad t_k = k\Delta t \qquad \forall k \in \mathbb{N},$$

where $\epsilon_1, \epsilon_2, \cdots$ are independent, identically distributed random variables satisfying $\text{Prob}(\epsilon_k = 1) = \text{Prob}(\epsilon_k = -1) = 1/2$ *for all $k \in \mathbb{N}$. Also assume that $\sigma > |\nu - \nu_0|\sqrt{\Delta t}$.*

Then for any derivative security X with payoff $f(S^T)$ at time $T = n\Delta t$, its price at initial time is

$$P(X) = P_{\Delta t}(S, T) := \sum_{k=0}^{n} \frac{n!\ \ p^k(1-p)^{n-k}}{k!(n-k)!}\ e^{-\nu_0 T}\ f\Big(S\, e^{\nu T + \sigma(2k-n)\sqrt{\Delta t}}\Big)$$

where S is the current price of the underlying security and $p = \dfrac{e^{(\nu_0-\nu)\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}$.

Exercise 3.1. *Under the assumption of Theorem 3.1, show that if the current time is $t = T - kn\Delta t$ and spot security price is S, then the price of a contingent claim with payoff $f(S^T)$ at time T is*

$$P_{\Delta t}(S, T-t) := \sum_{i=0}^{m} \frac{m!\;\, p^i(1-p)^{m-i}}{i!(m-i)!}\; e^{-\nu_0(T-t)}\; f\Big(S\, e^{\nu(T-t)+\sigma(2i-m)\sqrt{\Delta t}}\Big) \tag{3.1}$$

3.2 Pricing Options

An **option** is *the right, but not obligation, to buy or sell an asset under specific terms.*

A **call option** is the one that *gives the right to purchase something.*

An **put option** is the one that *gives the right to sell something.*

To **exercise** an option means *the actually buying or selling asset according to option terms.*

An option **buyer** or **holder** has the right to exercise an option according to the option terms.

An option **writer** or **seller** has the obligation to fulfil buyer's right.

The specifications of an option include, but may not limited to the following:

1. A clear description of what can be bought (for a call) or sold (for a put).

   For options on stock, each option is usually for 100 shares of a specified stock. Mathematically, one option means for one share of stock.

2. The **exercise price** or **strike price** for the underlying asset to be sold or bought at.

3. The period of time that the option is valid. This is typically defined as **expiration date**.

There are quite a number of options types:

1. **European option** In this option, the right of the option can be exercise only on the expiration date. Strike price is fixed.

2. **American option** The option right can be exercised any time on or before the expiration data. Strike price is fixed.

3. **Bermudan option** The exercise dates are restricted, in some case to specific dates, in other cases to specific periods within the lifetime of the option.

4. **Asian option** The payoff depend on the average price S_{avg} of the underlying asset during the period of the option. There are basically two ways that the average can be used.

   (i) S_{avg} is served as strike price; the payoff for a call is $\max\{S^T - S_{avg}, 0\}$.

   (ii) S_{avg} is served as the final asset price; the payoff for a call is $\max\{S_{avg} - K, 0\}$.

5. **Look-back option** The effective strike price is determined by the minimum (in the case of call) or maximum (in the case of put) of the price of the underlying asset during the period of the option. For example, a European look-back call has payoff = $\max\{S^T - S_{min}\}$ where S_{min} is the minimum value of the price S over the period from initiation to termination T.

6. **Cross-ratio option** These are foreign-currency options denominated in another foreign currency; for example, a call for 100 US dollars with an exercise price of 95 euros.

7. **Exchange option** Such option gives one the right to exchange one specified security for another.

8. **Compound option** A compound is an option on an option.

9. **Forward start option** These are options paid for one date, but do not begin, until a later date.

10. **"As you like it" option** The holder can, at a specific time, declare the option to be either a put or a call.

PS: The words European, American, Bermudan, Asian are words for the structure of the option, no matter where they are issued.

Option prices can be calculated by using tree structures, if the dynamics of the price of underlying asset can be described by a finite state model. In the sequel, we provide a few examples demonstrating how to calculate prices of options by using a binomial tree structure. The key here is we have to *use risk-neutral probability, not natural probability!*

Assume that risk-free interest rate is $\nu_0 = 0.10$. Let's consider of a stock of initial price $S^0 = 100$, growth rate $\nu = 0.12$/year, and volatility $\sigma = 0.20/\sqrt{\text{year}}$. We shall consider options of duration three months, so we construct a three month tree structure. For simplicity, de take $\Delta t = 1$ (month) $= 1/12$ (year). We find

$$u = e^{\nu\Delta t+\sigma\sqrt{\Delta t}} = 1.0701, \quad d = e^{\nu\Delta-\sigma\sqrt{\Delta t}} = 0.9534, \qquad p = \frac{e^{(\nu_0-\nu)\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta}}} = 0.4712$$

The tree is displayed as follows:

states	*ddd*	*ddu*	*dud*	*duu*	*udd*	*udu*	*uud*	*uuu*
p_k	.1479	.1318	.1318	0.1174	.1318	.1174	.1174	0.1046
$S^3(k)$	86.66	97.26	97.26	109.17	97.26	109.17	109.17	122.53
$S^2(k)$	90.89		102.02		102.02		114.51	
$S^1(k)$	95.33				107.01			
S^0				100.00				

Here the second line displays all possible states, and the top line provides the risk-neutral probability of these states, being $p_k = p^k(1-p)^{3-k}$, $k = 0, 1, 2, 3$. The lines after are possible prices of the stock. Since we have constant ν, ν_0, σ, the risk-neutral probabilities are constants in the sense that probability p for up and $1-p$ for down.

We now consider the following options. Their durations are all three months.

1. **A European call option** with strike price $K = \$100$. The caller can buy a stock from seller at price 100 and sell it on the market at price S^T, so the payoff is $S^t - 100$. Of course, if $S^t \leqslant 100$, the caller just let the option void. Hence the payoff is $X = \max\{S^3 - 100, 0\}$. From this, we can use the risk-neutral probability to determine its price

$$\begin{aligned} EC &= e^{-\nu_0 T}\sum_{k=1}^{8} p_k X(k) = e^{-\mathbf{u}_0 T}\sum_{i=0}^{8} p_k \max\{S^3(k) - 100, \ \ 0\} \\ &= e^{-0.1*3/12}\Big\{0 + 3*0 + 3*9.17*0.1174 + 22.53*0.1046\Big\} = \$5.44\ . \end{aligned}$$

Thus, one European call option worths \$5.44.

Note that for European call options, we need only a lattice structure.

2. **A European put option** with strike price $K = \$100$. The option holder can by one share of stock at price S^T and sell to the writer at $K = 100$, with profit $100 - S^T$. Surely, if $S^T \geqslant 100$, the option expires quietly. Hence the payoff of is $X = \max\{100 - S^3\}$. Using risk-neutral probability, we calculate the option's price as

$$\begin{aligned} EP &= e^{-\nu_0 T} \sum_{i=0}^{8} p_k \max\{100 - S^3(k), \ \ 0\} \\ &= e^{-0.1*3/12} \Big\{ 13.34 * 0.1479 + 3 * 2.74 * 0.1318 + 3 * 0 + 0 \Big\} = 2.97. \end{aligned}$$

We check the **put-call option parity formula** $EC - EP = S^0 - Ke^{-\nu_0 T}$:

$$EC - EP = 5.44 - 2.97 = 2.47, \qquad S - Ke^{-\nu_0 T} = 100 - 100e^{-0.1*3/12} = 2.47.$$

3. **An Asian call option** with strick price $K = (\sum_{i=1}^{3} S^t)/3$ (The average of past three month). Then the payoff is $X = \max\{0\ ,\ \ S^3 - (\sum_{i=1}^{3} S^t)/3)\}$. Tracking the history, we find value's of X and the corresponding probability as follows:

history	*uuu*	*uud*	*udu*	*duu*	*udd*	*dud*	*ddu*	*ddd*
K	114.68	110.23	106.07	102.18	102.10	98.21	94.50	90.96
S^3	122.53	109.17	109.17	109.17	97.26	97.26	97.26	86.66
p	0.1046	0.1174	.1174	0.1174	0.1317	0.1317	0.1317	0.1479

Here K is calculated by taking the average of three stock prices in the last three month. For example,

$$K(udu) = \Big\{ S^0 u + S^0 ud + S^0 udu \Big\} / 3 = 106.07.$$

Thus, the price of the Asian call option is

$$\begin{aligned} AC &= e^{-0.1*3/12} \{ (122.53 - 114.68) * 0.1046 + (109.17 - 106.07) * 0.1174 \\ &\quad + (109.17 - 102.18) * 0.1174 + (97.26 - 94.50) * 0.1317) = \$2.31\ . \end{aligned}$$

4. **An Asian Put Option** with strick price being the average of past three month. Then the payoff is $X = \max\{(\sum_{i=1}^{3} S^t)/3) - S^3, 0\}$. Tracking the history, we find value's of X and calculate its price by

$$\begin{aligned} AP &= e^{-0.1*3/12} \{ (110.23 - 109.17) * 0.1174 + (102.10 - 97.26) * 0.1317 \\ &\quad + (98.21 - 97.26) * 0.1174 + (90.96 - 86.66) * 0.1479) = \$1.41\ . \end{aligned}$$

Thus, the Asian put option should have a price of \$1.41.

5. **An American Call Option** with strike price $K = 100$. It can be argued that the best strategy for American call is to exercise the right at the last day. There is no advantage to exercise earlier! We leave the detailed calculation as an exercise.

6. **American Put.** This is a much hard problem. We have to work backwards to obtain its solution.

At the end of time, we know the value of the put, denoted by $P^3 = (K - S^3)^+ := \max\{K - S^3, 0\}$.

S^3	122.53	109.17	97.26	86.66
$P^3 = (K - S^3)^+$	0	0	2.74	13.34

At $t = 2$, we first use the risk-neutral probability calculate the price of the claim. It's value is

$$\hat{P}^2(\omega) = e^{-r\Delta t}\{p\, P^3(\omega u) + (1-p)P^3(\omega d)\}$$

One can show that a replicating portfolio at time $t = 2$ can be prepared to pay the P^3 exactly at time $t = 3$, and the cost of the portfolio is $\hat{P}^2$.

Since we have option to exercise the right valued at $\max\{K - S^2\}$. If we know the value $\hat{P}^2$ is smaller this, then we should exercise the option. Hence, we take the maximum of $(K - S^2)^+$ and $\hat{P}^2$. The calculation is as follows:

S^2	114.51	102.02	90.89
$(K - S^2)^+$	0	0	9.11
$e^{-r\Delta t}[P^3(i\omega u) + (1-p)P^3(\omega d)]$	0	1.43	8.28
P^2	0	1.43	**9.11**

In the next step, we perform the same calculation

$$P^1(\omega) = \max\Big\{(K - S^1(\omega))^+ \ \ , \ \ e^{-r\Delta t}(pP^2(\omega d) + (1-p)P^2(\omega d)\Big\}$$

S^1	107.01	95.34
$(K - S^1)^+$	0	4.66
$e^{-r\Delta t}[pP^2[i] + (1-p)P^3(i+1)]$	0.75	5.46
P^1	0.75	5.46

Finally, $P^0 = \max\{(K - S^0), e^{-r\Delta t}(pP^1(u) + (1-p)P^1(d)\} = 3.21$. Thus, the American put has value \$3.21.

7. **Summary.** From these example, we can summarize the method as follows:

1. If the claim has only a final payment X. Then the value at earlier time t_k can be evaluated by

$$P^{t_k}(\omega) = p\ P^{t_{k+1}}(\omega u) + (1-p)\ P^{t_{k+1}}(\omega d).$$

2. If there are **options** *to exercise a right at time t_k with value $X^{t_k}(S^{t_k})$ after the revelation of stock price S^{t_k} at time t_k, then*

$$P^{t_k}(\omega) = \max\Big\{X^{t_k}(S^{t_k}(\omega)) \ \ , \ \ p\ P^{t_{k+1}}(\omega u) + (1-p)\ P^{t_{k+1}}(\omega d)\Big\}$$

Finally, we have to emphasize that the value we calculated is obtained by taking $\Delta t = 1$ month. It is too big for the value to be accurate. Theoretically we should take $\Delta t = dt$, an infinitesimal.

Exercise 3.2. *For the American call option, follow the process as the American put option, go backward step by step find the value of the American call option. Show that at each time, the value of the call option is always bigger than $S^t - K$, so option holder should not exercise the option right. Consequently, the value of the American option is the same as that of the European put option.*

Exercise 3.3. *A stock has an initial price $S^0 = \$100$ and we are considering a four month security. Take $\Delta t =$ one moth, $\nu_0 = 0.1$/year, $\nu = 0.15$/year and $\sigma^2 = 0.04$/year, construct both a binomial tree and binomial lattice. Computer the risk-neutral probabilities.*

(1) Use the lattice to calculate the prices of a European put option, a European call option, an American call option, and an American put option. Here strike prices are the same as the current stock price and duration is fourth month.

(2) Use the tree to calculate the price for an Asian call option to be called at the beginning of the fifth month with strike price is $K = (S^2 + S^3 + S^4)/3$ where S^t is the price at the first day of $(t+1)$th month.

3.3 Replicating Portfolio for Derivative Security

As a special example of the general finite model, here we see how a portfolio is managed to provide the exact payment of a contingent claim, thereby providing the price of the claim. We continue to use the notations in the previous section.

Suppose X is a contingent claim with payoff $f(S^T)$. Here we shall use the binary tree model. Let $\{(n_0^t(\cdot), n_1^t(\cdot))\}_{t\in\mathbf{T}}$ be the trading strategy, where n_0^t is the number of shares of risk-free asset whose unit price is $e^{\nu_0 t}$ and $n_1^t(\cdot)$ is the number of shares of the underlying security whose unit price is $S^t(\cdot)$.

We denote by $V^t(\omega)$ the value of the portfolio at time t and event $\omega \in \Omega$. Hence,

$$V^t(\cdot) = n_0^t(\cdot)e^{\nu_0 t} + n_1^t(\cdot)S^t(\cdot).$$

Clearly, at time T, we require

$$V^T(\omega) = f(S^T(\omega)) \qquad \forall\, \omega \in \Omega.$$

We now investigate how a portfolio can be established at time t_{n-1} so that regardless of the outcome at time T, the updated value generated by the portfolio matches, with 100% certainty, with the needs, e.g. pay the claim off at time T.

Hence, assume that we are at time $t = t_{n-1}$ and at a state $(z_1, \cdots, z_{n-1}) \times B$. For convenience we use the following notations

$$\mathbf{z} = (z_1, \cdots, z_{n-1},\ ?\ ) \in \mathcal{P}^{n-1},$$
$$\mathbf{z}_U = (z_1, \cdots, z_{n-1}, 1), \qquad \mathbf{z}_D = (z_1, \cdots, z_{n-1}, -1).$$
$$s = S^{t_{n-1}}(\mathbf{z}), \qquad s_U = S^T(\mathbf{z}_U) = su, \qquad s_D = S^T(\mathbf{z}_D) = sd$$

where

$$u = e^{\nu\Delta t + \sigma\sqrt{\Delta t}}, \qquad d = e^{\nu\Delta t - \sigma\sqrt{\Delta t}}.$$

At current time $t = t_{n-1}$, we know event $\mathbf{z}$ happened and we have $n_0^t(\mathbf{z})$ shares of risk-free asset whose unit price is $e^{\nu_0 t}$ and $n_1^t(\mathbf{z})$ shares of security whose unit price is $s = S^t(\mathbf{z})$. At the end period (e.g. at time T), the outcome is either $\mathbf{z}_U$ or $\mathbf{z}_D$. For the value of the portfolio to prepare for the payment of the claim X in either outcome, it is necessary and sufficient to have

$$\left.\begin{array}{l} f(s_T) = n_0^t(\mathbf{z})e^{\nu_0 T} + n_1^t(\mathbf{z})s_U \\ f(s_D) = n_0^t(\mathbf{z})e^{\nu_0 T} + n_1^t(\mathbf{z})s_D \end{array}\right| \qquad \begin{array}{l} t = t_{n-1}, s = S^t(\mathbf{z}) \\ s_U = su, \quad s_D = sd \end{array}$$

This system has a unique solution given by, for $t = t_{n-1}$,

$$n_0^t(\mathbf{z}) = e^{-\nu_0 T}\frac{f(sd)u - f(su)d}{u - d}, \qquad n_1^t(\mathbf{z}) = \frac{f(su) - f(sd)}{su - sd}.$$

Notice that the value of the portfolio at time $t = t_{n-1}$ is

$$V^t(\mathbf{z}) = n_0^t(\mathbf{z})e^{\nu_0 t} + n_1^t(\mathbf{z})S^t(\mathbf{z}) = e^{\nu_0(t-T)}\{p\, f(su) + (1-p)\, f(sd)\}.$$

Thus, the portfolio at time $t = t_{n-1}$ is completely determined by the claim.

Now we consider the general time period from $t = t_k$ to $t_{k+1} = t + \Delta t$. Similar to the above discussion, we denote a general block in $\mathcal{P}^k$ by

$$\begin{gathered}\mathbf{z} = (z_1, \cdots, z_k, ?, \cdots, ?), \in \mathcal{P}^k,\\ \mathbf{z}_U = (z_1, \cdots, z_k, 1, ?, \cdots, ?) \in \mathcal{P}^{k+1},\\ \mathbf{z}_D = (z_1, \cdots, z_k, -1, ?, \cdots, ?) \in \mathcal{P}^{k+1},\\ s = S^t(\mathbf{z}), \quad s_U = S^{t+\Delta t}(\mathbf{z}_U) = su, \quad s_D = S^{t+\Delta t}(\mathbf{z}_D) = sd.\end{gathered}$$

Suppose we know the value of $V^{t+\Delta t}(\mathbf{z}_U) = V^{t+\Delta t}(su)$ and $V^{t+\Delta t}(\mathbf{z}_D) = V^{t+\Delta t}(sd)$. Then for the portfolio constructed at time $t = t_{n-1}$ to match exactly with the needed portfolio at time $t + \Delta t$, we need

$$V^{t+\Delta t}(\mathbf{z}_U) = e^{\nu_0(t+\Delta t)} n_0^t(\mathbf{z}) + s\, u\, n_1^t(\mathbf{z}), \qquad V^{t+\Delta t}(\mathbf{z}_D) = e^{\nu_0(t+\Delta t)} n_0^t(\mathbf{z}) + s\, d\, n_1^t(\mathbf{z}).$$

It follows that there are a unique portfolio (n_0^t, n_1^t) that provides the need for payment at at time t_{k+1}:

$$\begin{aligned} n_0^t(\mathbf{z}) &= e^{-\nu_0(t+\Delta t)} \frac{V^{t+\Delta t}(\mathbf{z}_D)u - V^{t+\Delta t}(\mathbf{z}_U)d}{u-d},\\ n_1^t(\mathbf{z}) &= \frac{V^{t+\Delta t}(\mathbf{z}_U) - V^{t+\Delta t}(\mathbf{z}_D)}{su - sd}, \end{aligned}$$

The value of the portfolio at $t = t_k$ is

$$V^t(\mathbf{z}) = e^{-\nu_0 t}\Big(p\, V^{t+\Delta t}(\mathbf{z}_U) + (1-p)\, V^{t+\Delta}(\mathbf{z}_D)\Big).$$

With this reduction formula, we then use an induction to derive the following formula, at $t = t_{n-m} = T - m\Delta t$ and spot security price $s = S^t(\mathbf{z})$,

$$V^t(\mathbf{z}) = \sum_{i=0}^{m} \frac{m!\, p^i\, (1-p)^{m-i}}{i!\, (m-i)!} e^{\nu_0(t-T)}\ f\Big(s\, e^{\nu(T-t)+\sigma\sqrt{\Delta t}(2i-m)}\Big).$$

It is important to observe that the right-hand sides depends only on s, i.e., it depends only on $\sum_{i=1}^m z_i$; namely, the binomial tree model can be replaced by a binomial lattice model, whose states are

$$\{(t_k, Su^i d^{k-i}) \mid i = 0, \cdots, k, k = 0, \cdots, n\}.$$

Hence, we can summarize our result in terms of the lattice model setting as follows.

Theorem 3.2 **(Portfolio Replication Theorem)** *Assume the condition of Theorem 3.1. Then at any time $t \in \mathbf{T}$ and spot underlying security price s, the value $V = V(s,t)$ of the contingent claim is*

$$V(s,t) = \sum_{i=0}^{m} \frac{m!\, p^i\, (1-p)^{m-i}}{i!\,(m-i)!} e^{\nu_0(t-T)}\, f\Big(s\, e^{\nu(T-t)+\sigma\sqrt{\Delta t}(2i-m)}\Big), \qquad m = \frac{T-t}{\Delta t}.$$

The portfolio replicating the contingent claim is unique. At any time $t-\Delta t$ and spot security price s, it consists of $n_{rf}(s,t-\Delta t)$ shares of risk-free asset and $n_S(s,t-\Delta t))$ shares of security, where

$$n_{rf}(s,t-\Delta t) = e^{-r\nu_0\Delta t}\frac{V(sd,t)\,u - V(su,t)\,d}{u-d},$$

$$n_S(s,t-\Delta t) = \frac{V(su,t)-V(sd,t)}{su-sd},$$

$$u = e^{\nu\Delta t+\sigma\sqrt{\Delta t}}, \qquad d = e^{\nu\Delta t-\sigma\sqrt{\Delta t}}, \qquad p = \frac{e^{(\nu_0-\nu)\Delta t}-e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}}-e^{-\sigma\sqrt{\Delta t}}} = \frac{e^{\nu_0\Delta t}-d}{u-d}.$$

Example 3.1. Consider a European call option with duration 3 month and strike price 19. The risk-free interest rate is $\nu_0 = 0.1$/year, current stock price is 20 with variance $\sigma^2 = 0.02$/year and return rate $\nu + \sigma^2/2 = 0.22$/year. Set $\Delta t = 1/12$ we have $n = 3$ and

$$u = 1.079, \qquad d = 0.961, \qquad p = 0.399.$$

The stock price, value of the call, and number of shares are given as follows:

time	stock	price		
0	20.00			
1	19.23	21.58		
2	18.48	20.75	23.28	
3	17.77	19.95	22.39	25.13

time	value			
0	1.73			
1	0.98	2.89		
2	0.37	1.90	4.44	
3	0	0.95	3.39	6.13

time	shares			
0	(-14.56, 0.815)			
1	(-11.93, 0.676)	(-18.53, 1.000)		
2	(-7.53, 0.435)	(-18.53, 1.000)	(-18.53, 1.000)	
3	(-7.53, 0.435)	(-18.53, 1.000)	(-18.53, 1.000)	(-18.53, 1.000)

Here the table for value is constructed as follows: (i) Last row is the payoff $= \max\{S-19, 0\}$. (ii) From each row up, $v(i,j) = [0.399 * v(i+1,j+1) + (1-0.399) * v(i+1,j)] * e^{-0.1/12}$. For example, third row last number: $[6.13 * p + 3.39 * (1-p)]e^{-r\Delta t} = 4.44$. The initial price of the call is 1.73.

Now we see how portfolio is constructed, i.e. how a writer of the call prepares for the call.

At time $t = 0$: the writer sells an option for 1.73 and sets up a portfolio consisting of -14.56 cash (share of risk-free asset) and 0.815 share of stock. We can check the balance $-14.56+0.815*20.00 = 1.73$.

At time $t = 1$ month, if stock price drops to 19.23, then the portfolio is worth 0.98. The writer responds by shorting $11.93 * e^{0.1/12}$ cash and longing 0.676 share of stock, total $-11.93 * e^{0.1/12} + 0.676 * 19.23 = 0.98$. If stock price rises up to 21.58, then the portfolio is worth 2.89. The writer responds by shorting 18.53 share of risk-free asset and longing 1.000 share of stock, totaling $-18.53 * e^{0.1/12} + 1.000 * 21.58 = 2.89$.

At time $t = 2$ month, if stock price is 20.75 or 23.28, the writer rolls over the portfolio. If price drops to 18.48, the writer shorts $7.53e^{0.1*2/12}$ cash and buys 0.435 share of stock, net value $-7.53 * e^{0.1*2/12} + 0.453 * 18.46 = 0.37$.

At t=3, in each situation, the portfolio pays the call exactly. Note that if stock price drop from 18.48 to 17.77, the portfolio worths $-7.53 * e^{0.1*3/12} + 0.435 * 17.77 = 0$, no need to answer call. If it moves up from 18.48 to 19.95, then the portfolio worths $-7.50 * e^{0.1*3/12} + 0.435 * 19.95 = 0.95$ which, adding caller's 19, is just enough to buy one share of stock at 19.95/share to give it to the caller. In other situations, -18.53 share of risk-free becomes -19.00 cash. Hence, the portfolio has -19.00 cash balance and one share of stock, just enough to pay the caller.

We emphasize again that we take $\Delta t = 1$ (month) is for illustration only. If possible, smaller Δt is preferred.

Exercise 3.4. *Using the parameters the example 3.1, calculate the price and corresponding replicating portfolio for (i) six month European call, strike price 19. (ii) Six month European put with strike price 19.*

Also, study the American call and put options with duration 3 months and strike price 20.

3.4 Certain Mathematical Tools

We review necessary mathematical tools from probability that are needed for the study of stochastic process.

> A **probability space** is a triple $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is a set, $\mathcal{F}$ is a σ-algebra on Ω, and $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$.
>
> A **random variable** is a measurable function on a probability space.
>
> A **stochastic process** is a collection $\{S_t\}_{t \in \mathbf{T}}$ of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$; here $\mathbf{T}$ is a set for time such as $\mathbf{T} = [0, T]$, $\mathbf{T} = [0, \infty), \mathbf{T} = \{0, 1, 2, \cdots\}$, or $\mathbf{T} = \{t_0, t_1, t_2, \cdots, t_n\}$, etc.

Here by $\mathcal{F}$ being a $\boldsymbol{\sigma}$**-algebra** on Ω it means that $\mathcal{F}$ is a non-empty collection of subsets of Ω that is closed under the operation of compliment and countable union.

Also, by a **probability measure**, it means that $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a non-negative function satisfying $P(\Omega) = 1$ and for every countable disjoint sets $A_1, A_2, \cdots$ in $\mathcal{F}$,

$$\mathbb{P}\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

Finally, a function $f : \Omega \to \mathbb{R}$ is called $\mathcal{F}$ **measurable** if

$$\{\omega \in \Omega \mid f(\omega) < r\} \in \mathcal{F} \quad \forall r \in \mathbb{R}.$$

Two random variables X and Y are called equal and write $X = Y$ if $\mathbb{P}(\{\omega \in \Omega \mid X(\omega) \neq Y(\omega)\}) = 0$.

The simplest measurable function is the **characteristic function** $\mathbf{1}_A$ of a measurable set $A \in \mathcal{F}$ defined by

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A, \\ 0 & \text{if } \omega \notin A. \end{cases}$$

A simple function is a linear combination of finitely many characteristic functions. For a simple function $\sum_{i=1}^n c_i \mathbf{1}_{A_i}$, its integral is defined as

$$\int_\Omega \left(\sum_{i=1}^n c_i \mathbf{1}_{A_i}(\omega) \right) \mathbb{P}(d\omega) := \sum_{i=1}^n c_i \mathbb{P}(A_i).$$

The **integral** of a general measurable function is defined as the limit (if it exists) of integrals of an approximation sequence of simple functions.

In a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, each $\omega \in \Omega$ is called a sample event, and each $A \in \mathcal{F}$ is called an **observable event** with observable probability $\mathbb{P}(A)$. Similarly, if A is not measurable, then A is called a **non-observable event**.

Let Ω be a set and S be a collection of subsets of Ω. We denote by $\sigma(S)$ the smallest σ-algebra that contains S. $\sigma(S)$ is called the σ-algebra **generated** by S.

In $\mathbb{R}$, the σ-algebra $\mathcal{B}$ generated by all open intervals is called the Borel σ-algebra and each set in $\mathcal{B}$ is called a **Borel set**.

If X is a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ and B is a Borel set of $\mathbb{R}$, then

$$X^{-1}(B) := \{\omega \in \Omega \mid X(\omega) \in B\}$$

is a measurable set. In the sequel, we use notation, for every Borel set B in $\mathbb{R}$,

$$\text{Prob}(X \in B) := \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\}).$$

Note that the mapping:

$$\mathbb{P}X^{-1} : \; B \in \mathcal{B} \longrightarrow \mathbb{P}(X^{-1}(B))$$

defines a probability measure on $(\mathbb{R}, \mathcal{B})$; namely, $(\mathbb{R}, \mathcal{B}, \mathbb{P}X^{-1})$ is a probability space.

$$\mathbb{R} \overset{\mathbb{P}}{\longleftarrow} \mathcal{F} \overset{X^{-1}}{\longleftarrow} \mathcal{B}$$

In the study of a single random variable X, all properties of the random variable are observed through the probability space $(\mathbb{R}, \mathcal{B}, \mathbb{P}X^{-1})$, since X is experimentally observed by measuring the probability of the outcome $X^{-1}(B)$ for every $B \in \mathcal{B}$.

Given a random variable X, its **distribution function** is defined by

$$F(x) := \text{Prob}(X \leqslant x) := \text{Prob}(X \in (-\infty, x]) := \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \leqslant x\}) \quad \forall\, x \in \mathbb{R}.$$

The **distribution density** is defined as the derivative (if it exists) of F:

$$\rho(x) = \frac{dF(x)}{dx} \qquad \forall\, x \in \mathbb{R}.$$

A random variable is called $N(\mu, \sigma^2)$ ($\mu \in \mathbb{R}, \sigma > 0$) distributed, or with mean μ and standard deviation σ (i.e. variance σ^2) if it has the probability density

$$\rho(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big).$$

When $\sigma = 0$, a $N(\mu, 0)$ random variable X becomes a deterministic constant function $X(\omega) \equiv \mu$ for (almost) all $\omega \in \Omega$.

In the sequel, we use $\mathbb{E}$ for **expectation** and **Var** for the **variance**

$$\begin{aligned}
\mathbb{E}[X] &:= \int_\Omega X(\omega)\mathbb{P}(d\omega), \\
\mathbf{Var}[X] &:= \int_\Omega \Big(X(\omega) - \mathbb{E}(X)\Big)^2 \mathbb{P}(d\omega) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.
\end{aligned}$$

The Law of the unconscious statistician

Given a random variable X with probability density function ρ and an itegrable real function f on $(\mathbb{R}, \mathcal{B})$, the expectation of $f(X)$ is

$$\mathbb{E}[f(X)] := \int_\Omega f(X(\omega))\mathbb{P}(d\omega) = \int_{-\infty}^{\infty} f(x)\rho(x)\,dx.$$

As far as a single random variable X is concerned, in certain sense all relevant information in $\mathbb{P}$ is contained in the measure $\mathbb{P}X^{-1}$, i.e. the distribution function F. Of course in such a study we lost track of the model underlying the random variable. Nothing is lost just so long as we are interested in one random variable.

If however we have two random variables X and Y on $(\Omega, \mathcal{F}, \mathbb{P})$, then two distribution functions for $\mathbb{P}X^{-1}$ and $\mathbb{P}Y^{-1}$ are not by themselves sufficient to say all there is about X, Y and their interrelation. We need to consider $Z = (X, Y)$ as a map of Ω into $\mathbb{R}^2$ and define a measure by $\mathbb{P}Z^{-1}$. This measure on the Borel sets of the plane defines what is usually called the joint distribution of X and Y, and is sufficient for a complete study of X, Y and their interrelations. This idea of course extends to any finite number of random variables.

The concept of a stochastic process is now a straightforward generalization of these ideas. For any index set $\mathbf{T}$ of time, a **stochastic process** on $\mathbf{T}$ is a collection of random variables $\{S_t\}_{t\in\mathbf{T}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. If $\mathbf{T} = \{t_1, \cdots, t_n\}$ has finitely many elements, then the stochastic process can be described completely by the joint distribution of the random variables $S_{t_1}, \cdots, S_{t_n}$. In other words, it is completely characterized by the probability distribution function of the vector valued random variable

$$Z := (S_{t_1}, \cdots, S_{t_n}) : \Omega \to \mathbb{R}^n.$$

For each $\omega \in \Omega$, the map $S(\omega) : \mathbf{T} \to \mathbb{R}$ defined by $t \to S_t(\omega)$ is a function from $\mathbf{T}$ to $\mathbb{R}$, which we call a **sample path**. Let's denote by $\mathrm{Map}(\mathbf{T}; \mathbb{R})$ the collection of all functions from $\mathbf{T}$ to $\mathbb{R}$. Thus, a stochastic process can be viewed as a function S from Ω to $\mathrm{Map}(\mathbf{T}; \mathbb{R})$ defined by $S : \Omega \to S(\omega) \in \mathrm{Map}(\mathbf{T}; \mathbb{R})$. Here $S(\omega) : t \to \mathbb{R}$ is defined by $S(\omega)(t) = S_t(\omega)$. In many applications, one simply take Ω as a subset of Map($\mathbf{T};\mathbb{R}$). In such a case, any $\omega \in \Omega$ is a function in Map($\mathbf{T};\mathbb{R}$) and hence, the function S from $\Omega \to$ Map($\mathbf{T};\mathbb{R}$) is realized through the default inclusion

$$S(\omega) := \omega(\cdot) \qquad \forall\,\omega \in \Omega \subset \mathrm{Map}(\mathbf{T};\mathbb{R}),$$
$$S(\omega)(t) = S_t(\omega) = \omega(t) \qquad \forall\,t \in \mathbf{T}\,\forall\omega \in \Omega \subset \mathrm{Map}(\mathbf{T};\mathbb{R}),$$

Note that the probability is needed to build upon a subset Ω of the space Map($\mathbf{T};\mathbb{R}$) of functions.

Exercise 3.5. *Assume that $f : \mathbb{R} \to \mathbb{R}$ is a simple function. Prove the law of unconscious statistician.*

Exercise 3.6. *Let $\Omega = \{1, 2, 3, 4\}$. Let $\mathcal{F}$ be the smallest σ-algebra that contains $\{1\}$ and $\{2\}$.*
(i) List all the elements in $\mathcal{F}$;
(ii) Is the event $\{1, 2, 3\}$ observable?

Exercise 3.7. *For every random variable X, show that*

$$\mathbf{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

Exercise 3.8. *Suppose X is $N(\mu, \sigma^2)$ distributed. Find the following:*

$$\mathbb{E}[X], \quad \mathbf{Var}[X], \quad \mathbb{E}[(X - \mathbb{E}[x])^3], \quad \mathbb{E}[(X - \mathbb{E}[x])^4], \quad \mathbb{E}[e^{\mathbf{i}\lambda X}]\ (\lambda \in \mathbb{C}).$$

Exercise 3.9. *Let $f, g : \mathbb{R} \to \mathbb{R}$ be continuous functions. Show that there exists a smallest σ-algebra $\mathcal{F}$ on $\mathbb{R}$ such that both f and g are $\mathcal{F}$ measurable. Also show that each element in $\mathcal{F}$ is a Borel set.*

Suppose $f = \mathbf{1}_{[1,\infty)}$ and $g = \mathbf{1}_{(-\infty,0]}$. Describe $\mathcal{F}$. Also show that if $h : \mathbb{R} \to \mathbb{R}$ is $\mathcal{F}$ measurable, then there exist constants c_1, c_2, c_3 such that $h = c_1\mathbf{1}_{(-\infty,0]} + c_2\mathbf{1}_{(0,1)} + c_3\mathbf{1}_{[1,\infty)}$. Consequently, $h = c_2 + (c_1 - c_2)g + (c_3 - c_2)f$.

3.5 Random Walk

Brownian motion is one of the most important building block for stochastic process. To study Brownian motion, we begin with a random walk, a discretized version of the Brownian motion. Roughly speaking,

> A **random walk** is the motion of a particle on a line, which walks in each unit time step a unit space step in one direction or the opposite with probability 1/2 each.

3.5.1 Description

Mathematically, we describe a random walk on the real line by the following steps.

1. Let $X_1, X_2, X_3, \cdots$ be a sequence of **independent** binomial random variables taking values $+1$ and -1 with equal probability:

$$\text{Probability}(X_i = 1) = \frac{1}{2}, \qquad \text{Probability}(X_i = -1) = \frac{1}{2}.$$

Such a sequence can be obtained, for example, by tossing a fair coin, head for 1 and tail for -1.

2. Let Δt be the unit time step and Δx be the unit space step. Let $W_{i\Delta t}$ be the position of the particle at time $t_i := i\Delta t$. It is a random variable, and can be defined as

$$\begin{aligned} W_0 &:= 0, \\ W_{i\Delta t} &:= \sum_{j=1}^{i} X_j\,\Delta x = W_{[i-1]\Delta t} + X_i\,\Delta x, \quad i = 1, 2, \cdots . \end{aligned}$$

Here the last equation $W_{i\Delta t} = W_{[i-1]\Delta t} + X_i \Delta x$ means that the particle moves from the position $W_{[i-1]\Delta t}$ at time $t_{i-1} = [i-1]\Delta t$ to a new position $W_{i\Delta t}$ at time $t_i = i\Delta t$ by walking one unit space step Δx, either to the left or to the right, depending on the choice of X_i being -1 or $+1$.

3. Though not necessary, it is sometimes convenient to define the position W_t of the particle at an arbitrary time $t \geqslant 0$. There are jump version and continuous versions. Here we use a constant speed version by a linear interpolation:

$$W_t = \left([i+1] - \frac{t}{\Delta t}\right) W_{i\Delta t} + \left(\frac{t}{\Delta t} - i\right) W_{[i+1]\Delta t} \quad \forall\, t \in (i\Delta t, [i+1]\Delta t),\ i = 0, 1, \cdots . \tag{3.2}$$

We call $\{W_t\}_{t\geqslant 0}$ the process of a **random walk** with time step Δt and space step Δx.

3.5.2 Characteristic Properties of a Random Walk

The random walk is described by the stochastic process $\{W_t\}_{t\geqslant 0}$ defined earlier. Here the probability space for the process is determined by the probability space associated with the random variables $\{X_1\}_{i=1}^{\infty}$.

A standard probability space $(\Omega, \mathcal{F}, \mathbb{P})$ for a sequence $\{X_i\}_{i=1}^{\infty}$ of binary random variables can be obtained as follows.

1. First we set

$$\Omega = \{-1, 1\}^{\mathbb{N}} = \{(x_1, x_2, \cdots) \mid x_i \in \{1, -1\} \forall\, i \in \mathbb{N}\}.$$

Each $\omega = (x_1, x_2, \cdots) \in \Omega$ can be regarded as the record of a sequences of coin tossing where x_i records the outcome of the ith toss, $x_i = +1$ for head and $x_i = -1$ for tail.

2. We define random variables X_i for each $i \in \mathbb{N}$ by

$$X_i(\omega) = x_i \qquad \forall\, \omega = (x_1, x_2, \cdots) \in \Omega.$$

Thus, $X_i(\omega)$ is the i-th outcome of the event $\omega \in \Omega$. Note that X_i takes only two values, $+1$ and -1. We denote

$$\begin{aligned} \Omega_i^{+1} &:= \{\omega \in \Omega \mid X_i(\omega) = +1\} = \{(x_1, x_2, \cdots) \in \Omega \mid x_i = +1\}, \\ \Omega_i^{-1} &:= \{\omega \in \Omega \mid X_i(\omega) = -1\} = \{(x_1, x_2, \cdots) \in \Omega \mid x_i = -1\}. \end{aligned}$$

3. The σ algebra $\mathcal{F}$ on Ω is the smallest σ-algebra on Ω such that all $X_1, X_2, \cdots$ are measurable. Clearly, it is necessary and sufficient to define $\mathcal{F}$ as the σ-algebra generated by the the countable boxes

$$\Omega_1^{+1},\ \Omega_1^{-1},\ \Omega_2^{+1},\ \Omega_2^{-1},\ \cdots .$$

Under this σ-algebra, each X_i is $(\Omega, \mathcal{F})$ measurable.

For each $(x_1, \cdots, x_n) \in \{-1, 1\}^n$, the cylindrical box $c(x_1, \cdots, x_n)$ is defined by

$$\begin{aligned} c(x_1, \cdots, x_n) &:= (x_1, \cdots, x_n) \times \prod_{j=n+1}^{\infty} \{-1, 1\} \\ &= \{(x_1, \cdots, x_n, y_{n+1}, y_{n+2}, \cdots) \mid y_j \in \{-1, 1\} \forall j \geq n+1\}. \end{aligned}$$

Note that each cylindrical box belongs to $\mathcal{F}$ since

$$c(x_1, \cdots, x_n) = \bigcap_{i=1}^{n} \{\omega \in \Omega \mid X_i(\omega) = x_i\}.$$

4. To define $\mathbb{P}$ such that both $X_i = 1$ and $X_i = -1$ has probability $1/2$, we first define $\mathbb{P}$ on the each cylindrical box by

$$\mathbb{P}(c(x_1, \cdots, x_n)) = 2^{-n} \qquad \forall n \in \mathbb{N}, (x_1, \cdots, x_n) \in \{-1, 1\}^n.$$

One can show that such defined $\mathbb{P}$ on all cylindrical boxes can be extended onto $\mathcal{F}$ to become a probability measure. Hence, we have a standard probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

5. One can show that under the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $\{X_i\}_{i=1}^{\infty}$ is a sequence of i.i.d. random variables having the property that the probability of $X_i = \pm 1$ is $1/2$.

The stochastic process $\{W_t\}_{t \geqslant 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ bas the following properties:

Properties of Random Walk

1. $W_0(\omega) \equiv 0$ for every $\omega \in \Omega$;
2. $\mathbb{E}[W(t)] = 0$ and $\mathbf{Var}[W(t)] = \sigma t$ for every $t \in \mathbf{T}$, where
$$\mathbf{T} = \{i \Delta t\}_{i=0}^{\infty}, \qquad \sigma = \frac{(\Delta x)^2}{\Delta t}.$$
3. For every $t_0, t_1, t_2, \cdots, t_n \in \mathbf{T}$ with $0 = t_0 < t_1 < \cdots < t_n$, the following random variables are independent:
$$W(t_n) - W(t_{n-1}), \quad W(t_{n-1} - t_{n-2}), \quad \cdots, \quad W(t_1) - W(t_0).$$
4. For every $\omega \in \Omega$, the function $t \in [0, \infty) \to W_t(\omega) \in \mathbb{R}$ is a continuous function.

3.5.3 Probabilities Related To Random Walk

In the rest of this section, we assume that $\{W_t\}_{t \geqslant}$ is a random walk with $\Delta x = 1$ and $\Delta t = 1$.

Example 3.2. *Find the probability that $W_3 \geqslant 0$.*

Solution. If we use "u" for up and "d" for down, any three-step random walk can be registered (denoted) by a three letter word consisting of only two alphabets "u" and "d". The set $\{\omega \mid W_3 \leqslant 0\}$ consists of random walks registered as "uuu", "uud", "udu", or "duu". Since there are a total of 8 three-step random walks, each of which has probability 1/8. Hence

$$\text{Prob}(W_3 \geqslant 0) = \frac{4}{8} = \frac{1}{2}.$$

Example 3.3. *Find the probability that $W_3 \geqslant 0$ and $W_4 \leqslant 0$.*

Solution. Among all four step random walks, those satisfying $W_3 \geqslant 0$ and $W_4 \leqslant 0$ are registered as "uudd", "udud", "duud". Since there are a total of 16 four-step random walks,

$$\text{Prob}(W_3 \leqslant 0, W_4 \geqslant 0) = \frac{3}{16}.$$

Example 3.4. *Find $\mathbb{E}[\tau]$ and $\mathbb{E}[W_\tau]$ where τ is the random variable defined by*

$$\tau = \min\{t \geqslant 0 \mid \text{either } t = 4 \text{ or } W_t \geqslant 1\}. \tag{3.3}$$

Solution. Note that if $\tau < 4$, then τ is the first time t such that $W_t = 1$. Hence τ is integer valued. We need only working on four-step random walks. We calculate

$$\text{Prob}(\tau = 0) = 0, \quad \text{Prob}(\tau = 1) = 1/2, \quad \text{Prob}(\tau = 2) = 0, \quad \text{Prob}(\tau = 3) = 1/8, \quad \text{Prob}(\tau = 4) = 3/8.$$

Hence,

$$\mathbb{E}[\tau] = \frac{1}{2} + 3 * \frac{1}{8} + 4 * \frac{3}{8} = \frac{19}{8}.$$

Note that when $\tau < 4$, we have $W_\tau = 1$. Also, the set $\tau = 4$ consists of four-step random walks registered as "dddd", "dddu", "ddud", "dduu", "dudd", "dudu". Hence,

$$\mathbb{E}[W_\tau] = \int_{\tau<1} W_\tau P(d\omega) + \int_{\tau=4} W_\tau P(d\omega) = \frac{5}{8} + \frac{1}{16}\Big\{ -4 - 2 - 2 - 0 - 2 - 0\Big\} = 0.$$

Example 3.5. *Consider the game that one wins one dollar if the outcome of a coin tossed is head, and lost one dollar otherwise. Consider the following strategy: Leave the game as soon as one wins a total of one dollar; otherwise, leave the game after four betting. What is the expected wining of such strategy?*

Solution. Denote by τ the time that the gambler leaves the game. It is given by (3.3). For each event $\omega \in \Omega$, she has $W_{\tau(\omega)}(\omega)$ amount of money when her leaves the game. Hence, for such a strategy, the expected wining is $\mathbf{E}[W_\tau] = 0$.

Exercise 3.10. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be defined as in §3.5.2. Show that*

$$\mathbb{P}(\{\omega \in \Omega \mid X_i(\omega) = 1) = \tfrac{1}{2},$$
$$\text{Prob}(X_i \in A, X_j \in B) = \text{Prob}(X_i \in A)\,\text{Prob}(X_j \in B) \quad \forall i \neq j.$$

Exercise 3.11. *Show that the stochastic process $\{W_t\}$ constructed in this section has the four listed properties for random walk.*

Exercise 3.12. *Show that for every positive integer n, $W_{n\Delta t}$ has the range $\{k\Delta x\}_{k=-n}^{n}$ and*

$$\text{Prob}(W_{n\Delta t} = k\Delta x) = \frac{C_n^k}{2^n}, \qquad C_n^k := \frac{k!(n-k)!}{n!}.$$

Exercise 3.13. *Let $\{W_t\}$ be a random walk with $\Delta x = 1$ and $\Delta t = 1$. Calculate the following probabilities:*

$$\text{Prob}(W_3 = 0, W_6 = 0), \qquad \text{Prob}(W_3 \in [-1,0], W_6 \in [0,1]), \qquad \text{Prob}(\max_{0 \leqslant t \leqslant 6} W_t \leqslant 1).$$

Exercise 3.14. *Let n be a positive integer and set $\Delta t = 1/n$ and $\Delta x = 1/\sqrt{n}$. Let $\Omega_n = \{-1,1\}^n$, $\mathcal{F}_n = 2^{\Omega_n}$ (the collection of all subsets of Ω_n) and*

$$\mathbb{P}(\omega) = \frac{1}{2^n} \quad \forall \omega \in \Omega_n.$$

1. *Show that $(\Omega_n, \mathbb{F}_n, \mathbb{P}_n)$ is a probability space.*
2. *In time interval $[0,1]$, show that there are a total of 2^n samples random walks. We denote by $W^1, \cdots, W^{2^n}$ all the sample random walks.*
3. *Denote by $C([0,1]) = C([0,1];\mathbb{R})$ the space of all continuous functions from $[0,1]$ to $\mathbb{R}$. Denote by $\mathcal{B}$ the Borel algebra on $C([0,1])$ generated by all open sets under the norm*

$$\|x\| = \sup_{t\in[0,1]} |x(t)| \quad \forall\, x \in C([0,1]).$$

   For every subset B of $C([0,1])$, we define $\mathbb{P}(B)$ as the number of random walks in B divided by 2^n. Show that $\mathbb{P}$ is a probability measure on $(C([0,1]),\mathcal{B})$.

4. *Let $W : \Omega \to C([0,1])$ be defined by $W(\omega)$ being the function $t \in [0,1] \to W_t(\omega)$. Show that W lift the probability space $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ to the probability space $(C([0,1]),\mathcal{B},\mathbb{P})$; that is, prove the following:*

   (a) for every $A \in \mathcal{F}_n$, $W(A) \in \mathcal{B}$ and $\mathbb{P}_n(A) = \mathbb{P}(W(A))$;

   [Notice that A is a finite set so $W(A)$ is also a finite set, and hence is a closed set in $C([0,1])$ which is of course a Borel set.]

   (b) for every $B \in \mathcal{B}$, $W^{-1}(B) \in \mathcal{F}_n$ and $\mathbb{P}(B) = \mathbb{P}_n(W^{-1}(B))$.

   [Count how many random walks are in the set B.

3.6 A Model for Stock Prices

In this section, we present a model for stock prices, by taking the limit of discretized model.

1. Random Walk and Brownian Motion

Let $\{\epsilon_i\}_{i=0}^{\infty}$ be a sequence of independent, identically distributed real valued random variables with mean zero and variance one. For simplicity, we assume that

$$\text{Prob}(\epsilon_i = 1) = \text{Prob}(\epsilon_i = -1) = \tfrac{1}{2} \qquad \forall i = 0, 1, \cdots.$$

Fix any $\Delta t > 0$. Consider the following random variables $\{z^{\Delta t}(t)\}_{t\geq 0}$ defined by

$$z^{\Delta t}(0) \equiv 0, \qquad z^{\Delta t}(t) = \sqrt{\Delta t} \sum_{0\leqslant i<t} \epsilon_i \quad \forall\, t > 0.$$

A **random walk** is the motion of a particle whose position at t is given by $z^{\Delta t}(t)$ for all $t \geq 0$.

Now fix a positive time t. Let $n = n(\Delta t, t)$ be an integer such that $t \in ((n-1)\Delta t, n\Delta t]$. Then

$$z^{\Delta t}(t) \quad = \quad \sqrt{\Delta t}\sum_{i=0}^{n-1}\epsilon_i = \sqrt{t}\sqrt{\frac{n\Delta t}{t}}\,\frac{\sum_{i=0}^{n-1}\epsilon_i}{\sqrt{n}}.$$

According to the central limit theorem, as $\Delta t \to 0$, the right-hand side has a limit in distribution, and the limit is normally distributed. Let's denote this limit by B_t. We know B_t is a random variable in some measure space $(\Omega, \mathcal{F}, \mathrm{m})$. The measure space $(\Omega, \mathcal{F}, \mathrm{m})$ is extremely hard to construct so we shall not worry about it. What we care is the distribution of B_t, which we can derive without much difficulty.

(A) First of all, from the central limit theorem, B^t is normally distributed with variance t:

$$\mathrm{Prob}(B^t \in A) := N(0,t;A) \qquad \forall\, A \in \mathcal{B}$$

where $\mathcal{B}$ is the σ-algebra of Borel set on $\mathbb{R}$ and $N(\mu, \sigma; A)$ represents the normal distribution

$$N(\mu,\sigma;A) := \frac{1}{\sqrt{2\pi}\,\sigma}\int_A e^{-(x-\mu)^2/(2\sigma^2)}dx \qquad \forall\ A \in \mathcal{B}.$$

(B) Next, for any fixed t and τ satisfying $t > \tau \geq 0$, we investigate the random variable $B^t - B^\tau$. Assume that $t \in ((n-1)\Delta t, n\Delta]$ and $\tau \in (k-1)\Delta t, k\Delta t]$ we have

$$\frac{z^{\Delta t}(t) - z^{\Delta t}(\tau)}{\sqrt{t-\tau}} = \sqrt{\frac{(n-k)\Delta t}{t-\tau}}\,\frac{\sum_{i=k-1}^{n-1}\epsilon_i}{\sqrt{n-k}}.$$

Again, sending $\Delta t \to 0$ we see that $B_t - B_\tau$ is normally distributed, with mean zero and variance $t - \tau$.

(C) Finally, take any $\{t_i, \tau_i\}_{i=1}^m$ that satisfies $0 \leqslant \tau_1 < t_1 < \tau_2 < t_2 < \cdots < \tau_m < t_m$. When Δt is sufficiently small, we know that $z^{\Delta t}(t_1) - z^{\Delta t}(\tau_i)$, $i = 1, \cdots, m$, are independent. Hence in the limit, we know that $B_{t_i} - B_{\tau_i}$, $i = 1, \cdots, m$ are also independent, i.e.

$$\mathrm{Prob}(B_{t_i} - B_{\tau_i} \in A_i\ \forall\, i = 1,\cdots,m) = \Pi_{i=1}^m \mathrm{Prob}(B_{t_i} - B_{\tau_i} \in A_i) \qquad \forall A_1,\cdots,A_m \in \mathcal{B}.$$

These three properties characterize all needed properties of the Brownian motion, also known as Winner Process. We formalize it as follows.

A **stochastic process** is a collection $\{\xi_t\}_{t\geqslant 0}$ of random variables in certain measure space $(\Omega, \mathcal{F}, \mathrm{m})$.

A **Brownian motion** or **Wiener process** is a stochastic process $\{B_t\}_{t\geq 0}$ satisfying the following:

1. $B_0 \equiv 0$;
2. for any $t > \tau \geqslant 0$, $B_t - B_\tau$ are normally distributed with mean zero and variance $t - \tau$;
3. for any $0 \leqslant t_1 < t_2 < \cdots < t_m$, the following increment are independent:

$$B_{t_m} - B_{t_{m-1}}, \quad B_{t_{m-1}} - B_{t_{m-2}}, \cdots, B_{t_2} - B_{t_1}.$$

We know that to model n independent real valued random variables, we need to use a sample space as big as $(\mathbb{R}^n, \mathcal{B}^n)$. If we are going to describe a sequence of i.i.d. random variables, we need a space something like $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}^{\mathbb{N}})$. Now to model the Brownian motion, we could use the space $\mathbb{R}^{[0,\infty)}$. However, this space is enormously big and we can hardly define a σ-algebra and meaningful measure on it. Deep mathematical analysis shows that Brownian motion can be realized on the space of all continuous functions from $[0,\infty) \to \mathbb{R}$. That is, one can take $\Omega = C([0,\infty);\mathbb{R})$, on it build a σ-algebra and define a measure. As a result, for every $\omega \in \Omega$, $B_t(\omega)$ is a continuous function from $t \to B_t(\omega)$.

2. Generalized Winner Process and Ito Process

Note that in discretized approximation of the Brownian motion, we have

$$z^{\Delta t}(t+\Delta t) - z^{\Delta t} = \epsilon(t)\sqrt{\Delta t}$$

where $\epsilon(t)$ is a random variable with mean zero and variance one. Hence, **symbolically** we can write

$$dB_t = \epsilon(t)\sqrt{dt}$$

where $\epsilon(t)$ is normally distributed having mean zero and variance 1. We have to say that the Brownian process is nowhere differentiable since

$$\mathbf{E}\Big(\Big[\frac{B_t - B_\tau}{t-\tau}\Big]^2\Big) = \frac{t-\tau}{(t-\tau)^2} = \frac{1}{t-\tau} \to \infty \quad \text{as } \tau \searrow t.$$

In engineering field, people use dB_t/dt symbolically to describe **white noise**.

The general Wiener process can be written as

$$W_t = \nu t + \sigma B_t \qquad \text{or} \quad dW_t = \nu dt + \sigma dB_t.$$

An **Ito process** is a solution to the stochastic differential equation

$$dx_t = a(x_t,t)dt + b(x_t,t)dB_t$$

Since with probability one the sample path $t \to B_t(\omega)$ is not differentiable, spacial tools, e.g. stochastic calculus are need. The following Ito's lemma [12] is no doubt one of the most important results.

Lemma 3.1 (**Ito Lemma**). *Suppose $\{x_t\}$ is a stochastic process satisfying $dx = a(x,t)dt + b(x,t)dB_t$. Let f be a smooth function : $\mathbb{R} \times [0,\infty) \to \mathbb{R}$. Then*

$$df(x_t,t) = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial t}\,dt + \frac{b^2}{2}\frac{\partial^2 f}{\partial x^2}\,dt.$$

If we use the ordinary differentials, the rule can be memorized as

$$(dt)^2 = 0, \qquad dtdB_t = 0, \qquad (dB_t)^2 = dt.$$

3. The Lognormal Process for Stock Prices.

A basic assumption in the Black–Scholes model is the **geometric Brownian motion** for the stock price S^t. In the form of stochastic differential equation, it reads

$$\frac{dS^t}{S^t} = \Big(\nu + \frac{\sigma^2}{2}\Big)dt + \sigma\, dB^t.$$

Here B^t is the standard Brownian motion also called Wiener process.

In the discretized case, we can write it as

$$\frac{S^{t+\Delta t} - S^t}{S^t} = \Big(\nu + \frac{\sigma^2}{2}\Big)\Delta t + \sigma\sqrt{\Delta t}\ \ \epsilon(t)$$

where $\text{Prob}(\epsilon(t) = 1) = \text{Prob}(\epsilon(t) = -1) = 1/2$. It then follows that

$$\begin{aligned}
\ln S^{t+\Delta t} - \ln S^t &= \ln\Big\{1 + \frac{S^{t+\Delta t} - S^t}{S^t}\Big\} \\
&= \frac{S^{t+\Delta t} - S^t}{S^t} - \tfrac{1}{2}\Big(\frac{S^{t+\Delta} - S^t}{S^t}\Big)^3 + O(1)\Big(\frac{S^{t+\Delta t} - S^t}{S^t}\Big)^3 \\
&= \Big(\nu + \frac{\sigma^2}{2}\Big)\Delta t + \sigma\sqrt{\Delta t}\,\epsilon(t) - \frac{1}{2}\sigma^2\Delta t\,(\epsilon(t))^2 + O(\Delta t^{3/2}) \\
&= \nu\Delta t + \sigma\sqrt{\Delta t} + ((\Delta t^{3/2})
\end{aligned}$$

here we use the assumption $\text{Prob}(\epsilon^2(t) = 1) = 1$. Hence, in the limit, we should have, symbolically

$$d\ln S^t = \nu\, dt + \sigma dB_t.$$

Hence, a geometric Brownian motion is also called a **lognormal process**. Here lognormal means log is normal not log of normal. To make everything rigorous, one needs first define stochastic calculus. For this we omit here.

Finally, since $d(\ln S^t - \mu t - B^t) = 0$, the solution is given by

$$\ln S^t = \ln S^0 + \nu t + B_t \qquad \text{or} \qquad S^t = S^0 e^{\nu t + \sigma B_t}.$$

Exercise 3.15. *A stock price is governed by* $S^0 \equiv 1$ *and* $d\ln S^t = \nu dt + \sigma dB_t$. *Find the following:*

$$\mathbf{E}[\ln S^t], \qquad \mathbf{Var}(\ln S^t), \qquad \ln\mathbf{E}(S^t), \qquad \mathbf{Var}(S^t).$$

Exercise 3.16. *If* $R_1, R_2, \cdots, r_n$ *are return rates of a stock in each of* n *periods. The arithmetic mean* R_A *and geometric mean* R_G *return rates are defined by*

$$R_A = \frac{1}{n}\sum_{i=1}^{n} R_i, \qquad R_G = \left(\prod_{i=1}^{n}(1 + R_i)\right)^{\frac{1}{n}} - 1.$$

Suppose \$40 is invested. During the first it increases to \$60 and the second year it decreases to \$48. What is the arithmetic mean and geometric mean?

When is it appropriate to use these means to describe investment performance?

Exercise 3.17. *The following is a list of stock price in 12 weeks:*

10.00, 10.08, 10.01, 9.59, 9.89, 10.55, 10.96, 11.25, 10.86, 11.01, 11.79, 11.74.

(1) From these data, find appropriate ν *and* σ^2 *in the geometric Brownian motion process for the stock price, take unit by year.*

(2) Suppose R *is the annual rate of return of the stock. Find approximation for* $\mathbf{E}(R)$ *and* $\mathbf{Var}(R)$.

3.7 Continuous Model As Limit of Discrete Model

In this section, we shall derive, in an indirect way, the Black–Scholes equation for pricing derivative securities. In the sequel, we shall assume that ν_0, ν, σ, T are all fixed constants, where $T > 0, \sigma > 0$.

We start with the price formula from the finite state model. We wish to derive the limit of the price, as $\Delta t \to 0$, so that we can obtain the continuous limit.

We shall use Taylor's expansion

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + O(x^4), \qquad \ln(1+x) = x - \frac{x^2}{2} + O(x^3) \qquad \text{as } x \to 0\,.$$

We can expand the risk neutral probability p by

$$\begin{aligned} p &= \frac{e^{(\nu_0-\nu)\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}}. \\ &= \frac{\sigma\sqrt{\Delta t} + [\nu_0 - \nu - \sigma^2/2]\Delta t + \sigma^3\Delta t^{3/2}/6 + O(\Delta t^2)}{2\sigma\sqrt{\Delta t} + 2\sigma^3(\Delta t)^{3/2}/6 + O(\Delta t^{5/2})} \\ &= \frac{1}{2}\Big\{1 + \frac{\nu_0 - \nu - \frac{\sigma^2}{2}}{\sigma}\sqrt{\Delta t} + O(\Delta t^{3/2})\Big\}. \end{aligned}$$

It then follows that for any integer $k \in [0, n]$,

$$\begin{aligned} &\ln[p^k(1-p)^{n-k}] = k\ln p + (n-k)\ln(1-p) \\ = \;& k\ln\Big\{\frac{1}{2}\Big[1 + \frac{\nu_0-\nu-\frac{1}{2}\sigma^2}{\sigma}\sqrt{\Delta t} + O(\Delta t^{3/2})\Big]\Big\} + (n-k)\ln\Big\{\frac{1}{2}\Big[1 - \frac{\nu_0-\nu-\frac{1}{2}\sigma^2}{\sigma}\sqrt{\Delta t} + O(\Delta t^{3/2})\Big]\Big\} \\ = \;& n\ln\frac{1}{2} + (2k-n)\frac{(\nu_0-\nu)-\frac{\sigma^2}{2}}{\sigma}\sqrt{\Delta t} - \frac{n\Delta t}{2}\Big\{\frac{\nu_0-\nu-\frac{\sigma^2}{2}}{\sigma}\Big\}^2 + nO(\Delta t^{3/2}) \\ = \;& -n\ln 2 + \frac{(2k-n)}{\sqrt{n}}\;\frac{(\nu_0-\nu-\sigma^2/2)\sqrt{T}}{\sigma} - \frac{1}{2}\Big(\frac{(\nu+\sigma^2/2-\nu_0)\sqrt{T}}{\sigma}\Big)^2 + O\Big(\frac{1}{\sqrt{n}}\Big). \end{aligned}$$

Now we define

$$x_k = \frac{2k-n}{\sqrt{n}} \quad \forall\, k \in \mathbb{Z}, \qquad \Delta x = x_{k+1} - x_k = \frac{2}{\sqrt{n}}.$$

Then

$$k = \frac{n}{2}\Big\{1 + \frac{x_k}{\sqrt{n}}\Big\}, \qquad n-k = \frac{n}{2}\Big\{1 - \frac{x_k}{\sqrt{n}}\Big\}.$$

For the factories, we use the Stirling's formula

$$k! = \sqrt{2\pi}\exp\Big((k+1/2)\ln k - k + \frac{\theta_k}{12k}\Big) \qquad \forall k \geq 1, 0 < \theta_k < 1.$$

It follows that when $k = 1, \cdots, n-1$,

$$\frac{1}{\Delta x}\frac{n!}{k!(n-k)!} = \frac{1}{\sqrt{2\pi}}\exp\Big([n+1]\ln n - \ln 2 - [k+1/2]\ln k - ([-k+1/2]\ln[n-k] + \xi_k\Big)$$

where

$$\xi_k = \frac{\theta_n}{12n} - \frac{\theta_k}{12k} - \frac{\theta_{n-k}}{12(n-k)} \qquad 0 < \theta_k, \theta_{n-k}, \theta_n < 1.$$

Therefore, substituting $k = n[1 + x_k/\sqrt{n}]/2$ and $n - k = n[1 - x_k/\sqrt{n}]/2$ we have

$$\begin{aligned} & \ln\Big(\frac{\sqrt{2\pi}}{\Delta x}\frac{n!}{k!(n-k)!}\Big) - \xi_k = [n+1]\ln n - \ln 2 - [k+1/2]\ln k - [n-k+1/2]\ln[n-k] \\ = \;& n\ln 2 - \frac{n}{2}\Big(1+\frac{1}{n}+\frac{x_k}{\sqrt{n}}\Big)\ln\Big(1+\frac{x_k}{\sqrt{n}}\Big) - \frac{n}{2}\Big(1+\frac{1}{n}-\frac{x_k}{\sqrt{n}}\Big)\ln\Big(1-\frac{x_k}{\sqrt{n}}\Big) \\ = \;& n\ln 2 - \frac{n+1}{2}\ln\Big(1-\frac{x_k^2}{n}\Big) - \frac{\sqrt{n}x_k}{2}\ln\frac{1+x_k/\sqrt{n}}{1-x_k/\sqrt{n}}. \end{aligned}$$

When $|x_k| < 2n^{1/4}$, we have

$$\begin{aligned} & \frac{n+1}{2}\ln\Big(1-\frac{x_k^2}{n}\Big) + \frac{\sqrt{n}x_k}{2}\ln\frac{1+x_k/\sqrt{n}}{1-x_k/\sqrt{n}}. \\ = \;& \frac{n+1}{2}\Big\{-\frac{x_k^2}{n}+\frac{O(1)x_k^4}{n^2}\Big\} + \frac{\sqrt{n}\,x_k}{2}\Big\{\frac{2x_k}{\sqrt{n}}+\frac{O(1)x_k^3}{n^{3/2}}\Big\} \\ = \;& \tfrac12 x_k^2 + \frac{O(1)x_k^4}{n}. \end{aligned}$$

Combining all these together, we than obtain when $|x_k| \le 2n^{1/4}$,

$$\begin{aligned} & \frac{1}{\Delta x}\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \\ = \;& \frac{1}{\sqrt{2\pi}}\exp\Big\{-\tfrac12 x_k^2 + x_k\frac{(\nu_0-\nu-\sigma^2/2)\sqrt{T}}{\sigma} - \tfrac12\Big(\frac{(\nu+\sigma^2/2-\nu_0)\sqrt{T}}{\sigma}\Big)^2 + \frac{O(1)x_k^4}{n}\Big\} \\ = \;& \frac{1}{\sqrt{2\pi}}\exp\Big(-\tfrac12\Big[x_k+\frac{(\nu+\sigma^2/2-\nu_0)\sqrt{T}}{\sigma}\Big]^2 + \frac{(O(1)x_k^4}{n}\Big). \end{aligned}$$

When $|x_k| \ge 2n^{1/4}$, one can verify that

$$\rho_k := \frac{1}{\Delta x}\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} \le O(1)e^{-\sqrt{n}}$$

since when $2n^{1/4} \le |x_k| \le 2n^{1/4}+1, \rho_k < e^{-\sqrt{n}}$ and

$$\frac{\rho_{k+1}}{\rho_k} = \frac{(n-k)p}{(k+1)(1-q)} > 1 \qquad \text{when} \;\; x_k < -n^{1/4},$$
$$\frac{\rho_{k+1}}{\rho_k} = \frac{(n-k)p}{(k+1)(1-q)} < 1 \qquad \text{when} \;\; x_k > n^{1/4}.$$

Hence, assume that f is continuous and bounded, we have

$$\begin{aligned} P_{\Delta t}(S,T) \;:=\; & \sum_{k=0}^{n}\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}e^{-\nu_0 T}f(S\,e^{\nu T+\sigma\sqrt{\Delta t}(2k-n)}) \\ =\; & \sum_{|x_k|<2n^{-1/4}}\frac{e^{-\nu_0 T}}{\sqrt{2\pi}}e^{-\frac12[x_k+\frac{(\nu+\sigma^2/2-\nu_0)\sqrt{T}}{\sigma}]^2+O(1)x_k^4/n)}f(Se^{\nu T+\sigma\sqrt{T}x_k})\Delta x + O(1)ne^{-\sqrt{n}}. \end{aligned}$$

Sending $\Delta t \to \infty$ (i.e. $n \to \infty$) we then obtain

$$\begin{aligned} \lim_{\Delta t\to 0} P_{\Delta t}(S,T) \;=\; & \int_{\mathbb{R}}\frac{e^{-\nu_0 T}}{\sqrt{2\pi}}e^{-\frac12(x+\frac{(\nu+\sigma^2/2-\nu_0)\sqrt{T}}{\sigma})^2}e^{-rT}f(Se^{\nu T+\sigma\sqrt{T}x})dx \\ =\; & \frac{e^{-\nu_0 T}}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{-z^2/2}f(e^{\ln S+\sigma z\sqrt{T}+(\nu_0-\sigma^2/2)T})\,dz. \end{aligned}$$

We summarize our calculation as follows.

Theorem 3.3 *Suppose the risk-free interest rate is a constant ν_0 and the unit price of the underlying stock is a* **geometric Brownian** *(or* **lognormal***) process*

$$\log S^t = \ln S + \nu t + \sigma B_t \quad \forall\, t \geqslant 0$$

where B_t is the standard Brownian motion process. Then a contingent claim at time $T > 0$ with payoff $f(S^T)$ has price given by the **Black-Scholes' pricing formula**

$$P(S,T) = \frac{e^{-\nu_0 T}}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-z^2/2} f(e^{\ln S + \sigma z\sqrt{T} + (\nu_0 - \sigma^2/2)T})\, dz.$$

In addition, at any time $t \in (0,T)$ and spot stock price s, the value of the contingent claim is

$$V(s,t) = \frac{e^{-\nu_0(T-t)}}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-z^2/2} f(e^{\ln s + \sigma z\sqrt{T-t} + (\nu_0 - \sigma^2/2)(T-t)})\, dz.$$

Furthermore, at any time $t \in (0,T)$ and spot price s, the portfolio replicating the contingent claim is given by $n_S(s,t)$ shares of stock and $n_{rf}(s,t)$ shares of risk-free asset (whose unit share price is $e^{\nu_0 t}$) where

$$n_S(s,t) = \frac{\partial V(s,t)}{\partial s}, \qquad n_{rf}(s,t) = e^{-\nu_0 t}\Big\{V(s,t) - s\, n_S(s,t)\Big\}.$$

We leave the derivation of formula for n_S and n_{rf} as an exercise.

Here we make a few observations:

(i) The parameter ν does not appear in the formula. Namely, the mean expected return of the stock is irrelevant to the price. This sounds very strange, but it explains the importance of Black-Scholes' work.

(ii) One notices that

$$V(s,t) = P(s, T-t).$$

That is, if the current stock price is s and there is $T-t$ time remaining toward to final time T, then the price of the contingent claim is $P(s,T-t)$, so is the value $V(s,t)$ of the portfolio.

(iii) Denote by

$$\Gamma(x,\tau) := \frac{1}{\sqrt{2\pi\tau}} e^{-x^2/(2\tau)} \quad \forall x \in \mathbb{R}, \tau > 0.$$

Then a change of variable $y = \ln S + \sigma z\sqrt{T} + (\nu_0 - \sigma^2/2)T$ we have

$$\begin{aligned} P(S,T) &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\sigma^2 T}} e^{-(y - \ln S - [\nu_0 - \sigma^2/2]T)^2/(2\sigma^2 T)} f(e^y) dy \\ &= e^{-rT} \int_{\mathbb{R}} \Gamma\Big(y - \ln S - [\nu_0 - \sigma^2/2]T, \sigma^2 T\Big) f(e^y) dy. \end{aligned}$$

Direct differentiation gives

$$\begin{aligned}
\frac{\partial P(S,T)}{\partial S} &= -\frac{e^{-\nu_0 T}}{S}\int_{\mathbb{R}} \Gamma_x f\, dy,\\
\frac{\partial^2 P(S,T)}{\partial S^2} &= -\frac{1}{S}\frac{\partial P}{\partial S} + \frac{e^{-\nu_0 T}}{S^2}\int_{\mathbb{R}} \Gamma_{xx} f\, dy,\\
\frac{\partial P(S,T)}{\partial T} &= -\nu_0 P - (\nu_0 - \sigma^2/2)e^{-\nu_0 T}\int_{\mathbb{R}} \Gamma_x\, f\, dy + \sigma^2 \int_{\mathbb{R}} \Gamma_\tau\, f\, dy,
\end{aligned}$$

Finally using $\Gamma_\tau = \frac{1}{2}\Gamma_{xx}$ we then obtain

$$\frac{\partial P}{\partial T} = -\nu_0 P + \Big(\nu_0 - \frac{\sigma^2}{2}\Big)S\frac{\partial P}{\partial S} + \frac{\sigma^2 S^2}{2}\Big\{\frac{\partial^2 P}{\partial S^2} + \frac{1}{S}\frac{\partial P}{\partial S}\Big\} = \frac{\sigma^2 S^2}{2}\frac{\partial^2 P}{\partial S^2} + \nu_0 S\frac{\partial P}{\partial S} - \nu_0 P.$$

Using $V(s,t) = P(s, T-t)$ we can derive an equation for V. We summarize our result as follows.

Theorem 3.4 *Suppose the risk-free interest rate is ν_0 and the price S^t of a security satisfies*

$$\ln S^t = \ln S + \nu t + \sigma B_t \qquad \forall\, t \geqslant 0$$

where B_t is the Brownian motion process. Consider a derivative security whose payoff occurs only at $t = T$ and equals $f(S^T)$. Then its price at $t = 0$ is $P(S,T)$ which, as function of $S > 0, T > 0$, satisfies the following **Black-Scholes' equation**

$$\frac{\partial P(S,T)}{\partial T} = \frac{\sigma^2 S^2}{2}\frac{\partial P}{\partial S^2} + \nu_0 S\frac{\partial P}{\partial S} - \nu_0 P \qquad \forall\, S > 0, T > 0. \tag{3.4}$$

Analogously, at any time $t \in [0,T]$ and spot price s of the security, the value $V(s,t)$ of the derivative security satisfies

$$\frac{\partial V(s,t)}{\partial t} + \frac{\sigma^2 s^2}{2}\frac{\partial V}{\partial s^2} + \nu_0\, s\,\frac{\partial V}{\partial s} = \nu_0 V \quad \forall t < T, s > 0. \tag{3.5}$$

Both $P(S,T)$ and $V(s,t)$ can be solved by supplying the respective initial conditions

$$P(S,0) = f(S) \quad \forall S > 0, \qquad V(s,T) = f(s) \quad \forall s > 0.$$

It is very important to know that ν plays no rule here. This is one of the Black–Scholes' most significant contribution towards the investment science.

That ν is irrelevant is due to the fact that only risk-neutral probability play roles here.

3.8 The Black–Scholes Equation

Here we provide a direct derivation for the Black-Scholes equation and a proof for the pricing formula.

Considered in the problem is a market system consisting of a risk-free bond and a risky stock. The price B^t of the bond and the spot price S^t of the stock obey the stochastic differential equations

$$\frac{dB^t}{B^t} = r\,dt, \qquad \frac{dS^t}{S^t} = \mu\,dt + \sigma\,dW_t$$

where $\{W_t\}_{t\geqslant 0}$ is the standard Wiener (Brownian motion) process. To be more general, we shall not assume that r, μ, σ are constants. In stead, we assume that $r = r(S,t), \mu = \mu(S,t)$ and $\sigma = \sigma(S,t)$ are given functions S and t. Of course, for the Black-Scholes equation to be well-posed (existence of a unique solution) we do need to assume that σ is positive and all r, μ, σ are bounded and continuous.

The problem here is to price, at time $t < T$ and spot stock price S^t, a derivative security (legal document) which will pay $f(S^T)$ at time T.

We shall carry out the task in two steps. In the first step, we show that if the derivative security can be replicated by a portfolio of stocks and bounds, then the value $V(S,t)$ of the portfolio at time t and spot price S must satisfy the Black-Scholes equation.

In the second step, we construct explicitly a self-financing portfolio whose value $V(S,t)$ is exactly the solution to the Black-Scholes equation. Thus, the price of the derivative is equal to $V(S,t)$; that is to say, the solution to the black-Scholes equation provides the price to the derivative security.

We have to say that step 1 is only a derivation of the equation. It is not part of the proof. If only a proof is needed, then step 1 is totally unnecessary. That is, only step 2 is the real proof that the price of derivative security satisfies the Black-Scholes equation. We present step 1 here is to let the reader see how Black-Scholes equation is first formally derived, and then shown to be the right one.

1. Assume that there is a replicating portfolio for the security. We denote by $n_s(S,t)$ the number of shares of stock and by $n_b(S,t)$ the number of shares of bond in the portfolio at time t and spot stock price S.

At time t and spot stock price S^t, the portfolio's value is

$$V^t = n_s(S^t,t)S^t + n_b(S^t,t)B^t.$$

At time $t + dt$ and spot stock price S^{t+dt}, the portfolio's value is

$$V^{t+dt} = n_s(S^t,t)S^{t+dt} + n_b(S^t,t)B^{t+dt}.$$

Thus, the change dV^t of the value of the portfolio due to change of prices of the stock and bond is

$$\begin{aligned} dV^t &= V^{t+dt} - V^t = n_s(S^t,t)\{S^{t+dt} - S^t\} + n_b(S^t,t)\{B^{t+dt} - B^t\} && (3.6)\\ &= n_s(S,t)dS + n_b(S,t)dB = n_s\{\mu S dt + \sigma S dW\} + n_b r B dt \\ &= \{\mu n_s S + n_b r B\}dt + \sigma n_s S dW^t && (3.7) \end{aligned}$$

after we plug in the assumed dynamics for prices of the stock and bond. Note that dV^t is a random variable, normally distributed.

Now assume that V^t can be written as $V(S^t,t)$ where $V(\cdot,\cdot)$ is a certain known function. Let's see how can we do this. Given a function $V(s,t)$ on $\mathbb{R} \times (-\infty, T]$, when we replace s by S^t, we obtain a random variable defined on the same space as that of the Brownian motion. By Ito's lemma, we know that $V(s,t)|_{s=S^t}$ relates the Brownian motion according to

$$\begin{aligned} dV(S^t,t) &= \frac{\partial V}{\partial t}dt + \frac{\partial V}{\partial S}dS + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}(dS)^2 \\ &= \Big\{\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{\sigma^2 S^2}{2ven}\frac{\partial^2 V}{\partial S^2}\Big\}dt + \sigma S\frac{\partial V}{\partial S}dW. && (3.8) \end{aligned}$$

Here we have used the fact that $(dt)^2 = 0, dtdW = 0, (dW)^2 = dt$ so $(dS)^2 = \sigma^2 S^2 dt$.

Hence, to have $V^t = V(s,t)|_{s=S^t}$, it is necessary and sufficient for coefficients of dt and dW in (3.7) and (3.8) to be exactly equal. Thus, we must have

$$\begin{aligned} \mu n_s S + n_b r B &= \frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2}, \\ \sigma n_s S &= \sigma S \frac{\partial V}{\partial S}. \end{aligned}$$

This is equivalent to require

$$n_s = \frac{\partial V}{\partial S}, \qquad n_b = \frac{1}{rB} \Big\{ \frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} \Big\}. \tag{3.9}$$

Therefor, the portfolio and the only portfolio that can replicate the derivative security is n_s shares of stock and n_b shares of bound, where n_s and n_b are as above. This is the only way that we can get rid of the randomness caused by Brownian motion process dW in the price change of stock.

We repeat a few more words about the randomness. Here S^{t+dt} is a random variable with normal distribution. The function $V(s, t + dt)$ itself is not a random variable, it becomes a random variable only when we replace s by S^{t+dt}. That the random variable $n_s(S^t,t)S^{t+dt} + n_b(S^t,t)B^{t+dt}$ is exactly the same as $V(S^{t+dt}, t + dt)$ is a very strong requirement. In the discrete model, we have learned of how to construct a replicating portfolio that matches exactly the required payment for contingent claim, regardless of which state the stock price lands on. Here is the same situation. The randomness is get rid of by matching the coefficients of two $dV's$. The former from the actual behavior of the stock price change, the other from Ito'e lemma and our hypothesis that $V^{t+dt} = V(s, t + dt)|_{s=S^{t+dt}}$, where $V(s, t + dt)$ is a function to be constructed without knowledge of the outcomes of actual price.

The value of the portfolio is $V = n_s S + n_b B$. Hence we need, in view of (3.9),

$$V = n_s S + n_b B = S \frac{\partial V}{\partial S} + \frac{B}{rB} \Big\{ \frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} \Big\}.$$

After simplification, this becomes

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2} \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} = rV,$$

which is exactly the famous Black–Scholes equation.

So far we have derived the Black–Scholes equation. We know that if a derivative security can be replicated by a portfolio, then its value or price must satisfy the Black-Scholes equation.

2. Now let V be the solution to the Black-Scholes equation with "initial" condition $V(s,T) = f(s)$ for all $s > 0$. Let n_s and n_b be defined as in (3.9). Consider the portfolio consisting of $n_s(S^t,t)$ shares of stock and $n_b(S^t,t)$ shares of bound at time $t + 0$ and spot stock price S^t.

First of all the value of the portfolio is

$$n_s S + n_b B = V(s,t)$$

by the definition of n_s, n_b and the differential equation for V.

Now we show that the portfolio is self-financing. For this, we calculate the capital needed to maintain such a portfolio.

At time $t + 0$, we have a portfolio of $n_s(S^t,t)$ shares of stock and $n_b(S^t,t)$ shares of bond. At time $t + dt$ before rebalancing, its value is $n_s(S^t,t)S^{t+dt} + n_b(S^t,t)B^{t+dt}$ which we wish to rebalance to

$n_s(S^{t+dt}, t+dt)$ shares of stock and $n_b(S^{t+dt}, t+dt)$ shares of bond. The capital δ needed to perform such a trade is

$$\begin{aligned}
\delta &:= \{n_b(S^{t+dt}, t+dt)S^{t+dt} + n_b(S^{t+dt}, t+dt)B^{t+dt}\} - \{n_s(S^t, t)S^{t+dt} + n_b(S^t, t)B^{t+dt}\} \quad (3.10) \\
&= \Big[\{n_b(S^{t+dt}, t+dt)S^{t+dt} + n_b(S^{t+dt}, t+dt)B^{t+dt}\} - \{n_b(S^t, t)S^t + n_b(S^t, t)B^t\}\Big] \\
&\quad + n_s(S^t, t)\Big[S^t - S^{t+dt}\Big] + n_b(S^t, t)\Big[B^t - B^{t+dt}\Big] \\
&= dV - n_s dS - n_b dB \\
&= dV - \frac{\partial V}{\partial S}\Big\{\mu\, Sdt + \sigma\, SdW\Big\} - \frac{1}{rB}\Big\{\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2}\Big\}\Big\{rB\, dt\Big\} \\
&= dV - \Big\{\frac{\partial V}{\partial t} + \mu S\frac{\partial V}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2}\Big\}dt - S\frac{\partial V}{\partial S}dW \\
&= 0.
\end{aligned}$$

Thus, the portfolio is self-financing. Since the outcome of the portfolio at time T is $V(S^T, t)$ which is equal exactly $f(S^T)$, regardless what S^T is, we see that the value of the contingent claim at any time t and spot stock price S^t had to be $V(S^t, t)$. Since if the the claim is sold for more, say at $\hat{V}(S^t, t) > V(S^t, t)$. Then we sell it at price $\hat{V}$, and form a portfolio of value $V(S^t, t)$, keep $\hat{V}(S^t, t) - V(S^t, t)$ in our pocket. Manage the portfolio in a self-financing way (prescribed by (n_s, n_b)) till the end of time T, at this time, the value of the portfolio exactly pays the claim. Thus there is no future payoff and we have a profit at time t. Similarly, if $\hat{V}(S^t, t) < V(S^t, t)$ one can do other way around. Since such arbitrage is excluded from the mathematical perfection, we conclude that the value of the contingent claim has to be $V(S^t, t)$.

Therefore, the price of the derivative security must be equal to V which is the unique solution to the Black-Scholes equation.

We summarize our result as follows.

Theorem 3.5 *Consider a system consisting of a risk-free asset and a risky asset whose price obeys a geometric Brownian motion process. Then any contingent claim with only a fixed one time payment can be uniquely replicated and therefore be priced, and the price can be calculated from the solution to the Black-Scholes equation.*

Exercise 3.18. *Complete the argument that if the price $\hat{V}(S^t, t)$ of the contingent claim is smaller than $V(S^t, t)$ at some time $t < T$ and some spot stock price S^t, then there is an arbitrage.*

Exercise 3.19. *Note that δ in (3.10) can be expresses as*

$$\begin{aligned}
\delta &= S^{t+dt}[n_s(S^{t+dt}, t+dt) - n_s(S^t, t)] + B^{t+dt}[n_b(S^{t+dt}, t+dt) - n_b^t(S^t, t)] \\
&= S^{t+dt}dn_s + B^{t+dt}dn_b = \{S + dS\}dn_s + (B + dB)dn_b.
\end{aligned}$$

Using Ito's lemma, the expression of n_s and n_b in (3.9), and the Black-Scholes equation for V show directly that $\delta = 0$.

Exercise 3.20. *Assume that r and σ are constants and $\sigma > 0$. Make the change of variables from (S, t, V) to (x, τ, v) by*

$$x = \ln S + \Big(r - \frac{\sigma^2}{2}\Big)(T - \tau), \qquad \tau = \frac{\sigma^2}{2}(T - t), \qquad V(s, t) = v(x, \tau)e^{-r(T-t)}$$

Show that the Black-Scholes equation for $V(S, t)$ becomes the following linear equation for v:

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} \quad \forall\, x \in \mathbb{R},\ \tau > 0, \qquad\qquad v(x, 0) = f(e^x) \quad \forall\, x \in \mathbb{R}.$$

Also show that the solution for v is given by the following formula:

$$v(x, t) = \frac{1}{\sqrt{4\pi\tau}} \int_{\mathbb{R}} e^{-(x-y)^2/(4\tau)} f(e^y) dy \qquad \forall\, x \in \mathbb{R}, \tau > 0.$$

Exercise 3.21. *Assume risk-free rate is r and volatility of a stock is $\sigma > 0$. Both r and σ are constants. Find the price for European put and call options, with duration time T and strick price K.*

Chapter 4

Optimal Portfolio Growth

In this chapter we consider multi-period investments. This leads to the study of dynamic portfolio in which one takes every opportunity to rebalance portfolio according to revelation of outcome of financial market.

To explain the idea, let's consider a simple situation. There is a stock whose price after each period is either double or reduce to half, with equal probability. If wealth is initially invested in the stock and unattended later on for quite many periods, the wealth will not change much. Therefore one may give up the idea of investment. However, the theory of dynamic portfolio tells us that there is indeed an excellent opportunity of investment. Suppose we start with \$200 and invest half of the wealth in to the stock. At the end of first period, if stock price doubles, we have a total wealth \$300, consisting of \$100 cash and \$200 worth of stock. We then happily cash in §50 stock, so the new balance is \$150 stock and \$150 cash. If unfortunately the stock price is halved, we have a total wealth of \$150, consisting of \$100 cash and \$50 worth of stock. Without hesitation we buy \$25 more stock, so the new balance is \$75 cash and \$75 worth of stock. Later on at the end of each period we always split the total wealth into half cash and half stock. In the long run, it is almost sure that the total wealth grows exponentially. The reason behind this is that we are following the dictum:

"sell high, buy low."

Conclusions of multi-period investment situations are not mere variations of single-period conclusions, rather they often *reverse* those earlier conclusions. This makes the subject exciting, both intellectually and in practice. Once subtleties of multi-period investment are understood, the reward in terms of enhanced investment performance can be substantial. This chapter shows how to design portfolios that have maximal growth.

4.1 Risk Aversion

1. Utility Function.

Consider the rationale of a person buying a lottery ticket with \$1, hoping to win a prize of \$1,000,000,000 with chance of 1 per 2,000,000,000. With \$1, there are two choices: \$1 risk-free if one keeps the money; a random payoff with expectation \$0.50 and great risk (uncertainty) if one buys the lottery ticket. In view of the mean-variance theory, what is the logic that one chooses investment with smaller expected return and larger "risk"?

To explain the rationale, we use von Neumann and Morgenstern's idea [20], introducing the following:

A utility function of an investor is a function $U : \mathbb{R} \to \mathbb{R}$ such that the decision to choose among investment plans with respect to random payments $X_1, \cdots, X_m$ is based on the maximum of the expectation $\mathbf{E}(U(X))$; namely, the investor with utility function U chooses a plan with return X_i satisfying

$$\mathbf{E}(U(X_i)) = \max_{1 \leqslant j \leqslant m} \mathbf{E}(U(X_j)).$$

The one general restriction placed on the form of the utility function is its monotonicity:

$$x > y \qquad \Longrightarrow \qquad U(x) > U(y).$$

Other than this restriction, the utility function can take any form.

Example 4.1. Suppose we use $U(x) = x^2/(100 + x)$ as our utility function.

1. Consider two choices: (a) a free lottery ticket having 1/2,000,000,000 chance of wining \$1,000,000,000, and (b) \$1.00 cash. What do we choose? We calculate which choice gives larger expected utility.
   For choice (a), $\mathbf{E}(U(X_a)) = 0 + (10^9)^2/(100 + 10^9)/(2 \times 10^9) \approx 0.50$.
   For choice (b), we have $\mathbf{E}(U(X_b)) = 1^2/(100 + 1) \approx 0.01$.
   Clearly, option (a) is chosen.

2. Consider two options: (c) 1,000 free lottery tickets, (b) \$1,000 cash. Which one do we choose?
   For choice (c), $\mathbf{E}(U(X_c)) \approx (10^9)^2/(100 + 10^9) * 1,000/(2 \times 10^9) \approx 500$.
   For choice (d), $\mathbf{E}(U(X_d)) = 1,000^2/(100 + 1000) \approx 909$.
   Hence, according the rank criterion, option (d) is chosen.

The simplest utility function is $U(x) = x$. An individual using this utility function ranks random payoffs by expected returns; that is, given choices $X_1, \cdots, X_m$ of payoffs, the payoff X_i is chosen if

$$\mathbf{E}(X_i) = \max_{1 \leqslant j \leqslant m} \mathbf{E}(X_j).$$

This utility function $U(x) = x$ is called **risk neutral** since it does not count for any risk being made.

From now on, we give another name to the standard deviation σ of a return—**volatility**. It is risk, but in another point of view, it is a chance. Later we shall see how volatility can be a good thing.

In practice, there are certain types of utility functions that are popular:

1. **Exponential**

   $$U(x) = -e^{-\alpha x}, \qquad \alpha > 0.$$

   This utility function has negative values, but it does not matter, as long as it is strictly increasing.

2. **Logarithmic**

   $$U(x) = \ln x, \qquad \forall x > 0, \quad U(x) = -\infty \quad \forall x \leqslant 0.$$

   Note that this function has a severe penalty for $x \approx 0$. Namely, if an investment has a chance that nothing will be payed, such a plan is out of consideration.

3. **Power**

$$U(x) = x^\alpha \qquad \alpha > 0, x \geqslant 0.$$

The case $\alpha = 1$ is the risk-neutral utility.

4. **Quadratic**

$$U(x) = 2Mx - x^2, \qquad x \leqslant M, M > 0.$$

In using this utility, it is assumed that the payoff has no chance of being larger than M.

5. **Rational**

$$U(x) = \frac{x^\alpha}{M + x}, \qquad x \geqslant 0,\ M > 0,\ \alpha \geqslant 1.$$

We have seen the use of this utility in the lottery example.

It is important to observe that if $k > 0$ and $b \in \mathbb{R}$, there is no difference between the utility function U and $V = kU + b$. Thus, in practice, we can scale a utility function conveniently.

2. The Quadratic Utility Function

Suppose we use the utility function $U(x) = 2Mx - x^2$ and have a number of assets $\mathsf{a}_1, \cdots, \mathsf{a}_m$ to invest upon for a total capital V_0. Assume the return of asset a_i is R_i. Then for a portfolio with weight $\mathbf{w} = (w_1, \cdots, w_m), \sum_{i=1}^m w_i = 1$, on assets $(\mathsf{a}_1, \cdots, \mathsf{a}_m)$, its final value is

$$\mathcal{V}[\mathbf{w}] = \sum_{i=1}^m (V_0 w_i)(1 + R_i) = V_0 + V_0 \sum_{i=1}^m w_i R_i = V_0 + V_0(\mathbf{w}, \mathbf{R})$$

where $\mathbf{R} := (R_1, \cdots, R_m)$ is a vector valued random variable. Hence, the rank of the portfolio $\mathbf{w}$ is made according to the value

$$\begin{aligned}\mathbf{E}(U(\mathcal{V}[\mathbf{w}])) &= 2MV_0^2 - V_0^2 + (2MV_0 - 2V_0^2)\mathbf{E}\Big((\mathbf{w}, \mathbf{R})\Big) - V_0^2 \mathbf{E}\Big((\mathbf{w}, \mathbf{R})^2\Big) \\ &= 2MV_0^2 - V_0^2 + 2V_0(M - V_0)\mu - V_0^2(\mu^2 + \sigma^2)\end{aligned}$$

where μ is the expected return and σ is the standard deviation of the return of portfolio:

$$\mu = \sum_{i=1}^m w_i \mu_i, \qquad \sigma^2 = \sum_{i,j=1}^m w_i w_i \sigma_{ij}, \qquad \mu_i = \mathbf{E}(R_i), \quad \sigma_{ij} := \mathbf{Cov}(R_i, R_j).$$

Taking the equivalent utility function $V(x) = [U(x) - 2MV_0^2 + V^0]/V_0^2$ and denoting $b = M/V_0 - 1$ we have

$$\mathbf{E}(\mathcal{V}[\mathbf{w}]) = b\, \mu - \mu^2 - \sigma^2.$$

It is easy to show that there is at least a maximizer, and in terms of the mean-variance Markowitz theory, the maximizer corresponds to a particular solution on the Markowitz frontier.

3. Risk Aversion[1]

[1] According to web dictionary, aversion: A fixed, intense dislike; repugnance.

Risk aversion is a penalty on the uncertainty (risk) of returns in ranking choices of investments. A utility function U is said to be **risk averse** on $[a, b]$ if it is concave on $[a, b]$, i.e.

$$U''(x) < 0 \quad \forall\, x \in [a, b].$$

If U is concave everywhere, it is said to be **risk averse**. The degree of risk aversion is formally defined by the **Arrow-Pratt absolute and relative risk aversion coefficients** defined as

$$a(x) := -\frac{U''(x)}{U'(x)}, \qquad r(x) = -\frac{xU''(x)}{U'(x)}.$$

Example 4.2. Suppose U is risk averse and we have two options: (a) flatly receive \$M; (b) based on the toss of a fair coin—head, we win \$10; tail, win nothing.

Now if we use a risk averse utility function U, we can make decision based on the expected utilities as follows:

Option (a): $\mathbf{E}(U(X_a)) = U(M)$.

Option (b): $\mathbf{E}(U(X_b)) = \frac{1}{2}\{U(0) + U(10)\}$.

(i) If $M = 5$, we see that $U(5) > \frac{1}{2}\{U(0) + U(10)\}$ since U is concave. Hence, option (a) is selected.

(ii) Suppose $U(x) = 25x - x^2$. Then $\mathbf{E}(U((X_a)) = M(25 - M)$ and $\mathbf{E}(U(X_b)) = 75 = \bar{M}(25 - \bar{M})$ where $\bar{M} = 3.49$. Hence, if $M > 3.49$, one prefers to receive \$M for sure instead of having a 50-50 chance of getting \$10 or 0.

From this example, we see that choosing a risk averse utility function lays penalty on uncertainties.

4. Certainty Equivalent.

The actual value of the expected utility of a random wealth variable is meaningless except in comparison with that of another alternative. There is a derived measure with units that do have intuitive meaning. This measure is certainty equivalent.

The **certainty equivalent** of a random variable X under utility U is the unique number c satisfying

$$U(c) = \mathbf{E}(U(X)).$$

In making a decision, one compares the certain equivalents of all possible payoffs and chooses the one having the highest certain equivalent.

Given a random variable X, denote its expectation $\mathbf{E}(X)$ by $\bar{X}$. Then by Taylor's expansion,

$$U(X) = U(\bar{X}) + U'(\bar{X})(X - \bar{X}) + (X - \bar{X})^2 \int_0^1 (1 - \theta)U''(\bar{X} + \theta[X - \bar{X}])d\theta.$$

Taking expectation on both sides we obtain

$$U(c) = U(\bar{X}) + \int_0^1 (1 - \theta)\mathbf{E}\Big(U''(\bar{X} + \theta[X - \bar{X}])(X - \bar{X})^2\Big)d\theta\ .$$

Hence, if $U'' \leqslant 0$, we have $U(c) \leqslant U(\bar{X})$ and also $c \leqslant \bar{X}$. From here we see that risk aversion puts penalty on risky payoffs in ranking investment plans.

The risk aversion characteristics of an individual depends on the individual's felling about risk, his or her current financial situation (such as net worth), the prospects for financial gains or requirements (such as college expenses) and individual's age. Financial planers can obtain such function by asking certain questions and based on answers to have values on certain parameters in a general formula, typically linear combinations of exponential functions.

Exercise 4.1. (certainty equivalent) *An investor has utility function $U(x) = x^{1/4}$. He has a new job offer which pays \$80,000 with a bonus being \$0, \$10,000, \$20,000, \$30,000, \$40,000, or \$50,000, each with equal probability. What is the certainty equivalent of this job offer.*

What are the corresponding certain equivalents when $U(x) = x,\ \ln x,\ x^2,\ -e^{-x}$, respectively?

Exercise 4.2. *Consider an investment of total capital V_0 among assets $\mathsf{a}_1, \cdots, \mathsf{a}_m$, each of which has a positive expected return. Assume that $\mathbf{C} = (\sigma_{ij})_{m\times m}$ is positive definite. Show that with a quadratic utility function $U(x) = 2Mx - x^2$, there is a unique optimal portfolio. Also find the certain equivalent of the optimal portfolio.*

Exercise 4.3. *Show that (i) the absolute risk aversion coefficient is a constant for exponential utility functions, and (ii) the relative risk aversion coefficient is constant for logarithmic and power utilities.*

Exercise 4.4. *Suppose X is a random wealth variable which has small $\mathbf{E}(|X - \mathbf{E}(X)|^3)$. Show that its certainty equivalent c can be approximated by*

$$c \approx \mathbf{E}(X) - \frac{1}{2}\, a(\mathbf{E}(X))\, \mathbf{Var}(X).$$

Exercise 4.5. *Why does a utility function have to be strictly increasing?*

Exercise 4.6. *For the lottery ticket example in Example 4.1, find the certainty equivalent $c(x)$ of x lottery tickets. When $c(x) > x$ and when $x > c(x)$?*

4.2 Portfolio Choice

In this section we focus on a single-period portfolio problem in which an investor uses the expected utility criterion to rank investment alternatives. Using the basic framework of Markowitz theory, we shall obtain conclusions more decisive than the original theory.

Suppose an invertor prefers a particular utility function U and has a total capital $V^0 > 0$ to invest among m assets $\mathsf{a}_1, \cdots, \mathsf{a}_m$. For $i = 1, \cdots, m$, the asset a_i has an initial unit share price $S_i^0 > 0$ and a final (end of period) price S_i, a non-negative bounded random variable with positive expectation $\mathbf{E}(S_j)$ on certain probability space $(\Omega, \mathcal{F}, P)$.

The investor wishes to form a portfolio to maximize the expected utility of final wealth. We denote a generic portfolio by $\mathbf{w} = (w_1, \cdots, w_m) \in \mathbb{R}^m$ with $\sum_{i=1}^n w_i = 1$, where w_i is the initial weight of total value of asset a_i in the portfolio. Since the initial unit share price for asset a_i is S_i^0 and a total of $V_0 w_i$ capital is invested in it, there are $V^0 w_i / S_i^0$ shares of asset a_i in the portfolio. Hence, the final value $\mathcal{V}[\mathbf{w}]$ of the portfolio is a random variable given by

$$\mathcal{V}[\mathbf{w}](\omega) := \sum_{i=1}^{m} \frac{V_0 w_i}{S_i^0}\, S_i(\omega) \qquad \forall\, \omega \in \Omega.$$

The investor's problem, or optimal portfolio problem, can be formulated as follows:

Investor's Problem: Find $\mathbf{w}^* \in W := \{\mathbf{w} \in \mathbb{R}^m \mid (\mathbf{w}, \mathbf{1}) = 1,\ \mathcal{V}[\mathbf{w}] \geqslant 0\}$ such that

$$\mathbf{E}\Big(U(\mathcal{V}[\mathbf{w}^*])\Big) = \max_{\mathbf{w}\in W} \mathbf{E}\Big(U(\mathcal{V}[\mathbf{w}])\Big). \tag{4.1}$$

We now show that this problem is connected to arbitrage.

Theorem 4.1 **(Portfolio Choice Theorem)** *Suppose U is continuous and increasing in $(0,\infty)$, $U(0) := \lim_{x\searrow 0} U(x)$ and $\lim_{x\to\infty} U(x) = \infty$. Also $\mathbf{E}(U(\mathcal{V}[\mathbf{w}])) > -\infty$ for some $\mathbf{w} \in W$.*

Then the optimal portfolio problem (4.1) has a solution if and only if there is no-arbitrage.

The solution, if exists, is unique if U is risk averse and all $a_1, \cdots, a_n$ are linearly independent.

Here by linearly dependent if means there exists a non-zero vector $\mathbf{w} \in \mathbb{R}^m$ such that $\mathcal{V}^T[\mathbf{w}] \equiv 0$. Therefore, linear independent means that

$$\mathcal{V}^t[\mathbf{w}] \equiv 0 \qquad \Longrightarrow \qquad \mathbf{w} = \mathbf{0}. \tag{4.2}$$

Theorem 4.2 **(Portfolio Pricing Theorem)** *Let $\mathbf{w}^* = (w_1^*, \cdots, w_n^*)$ be an optimal portfolio and V^* be the corresponding final payoff. Assume that $\mathbf{w}^*$ is in the interior of W. Then any derivative security of the underlying assets with a final payoff X has an initial price of*

$$P(X) = \frac{1}{1+\mu_0}\int_\Omega X(\omega)\mathbb{P}(d\omega), \qquad \mathbb{P}(A) := \frac{\int_A U'(V^*)\,P(d\omega)}{\int_\Omega U'(V^*)P(d\omega)} \qquad \forall\, A \in \mathcal{F},$$

where $\mu_0 > -1$ is a constant (the risk-free return rate) and $\mathbb{P}$ is called **risk-neutral probability**.

Proof of Theorem 4.2. Consider the Lagrangian

$$L(\mathbf{w},\lambda) = \mathbf{E}(U(\mathcal{V}[\mathbf{w}])) - \lambda\, V_0\{(\mathbf{w},\mathbf{1}) - 1\}.$$

Since $\mathbf{w}^*$ is a minimizer in the interior of W, according to a general theory from calculus, the first variation of L with respect to λ and $\mathbf{w}$ is zero. As $\mathcal{V}[\mathbf{w}] = \sum_i V^0 w_i S_i/S_i^0$, we then have

$$0 = \frac{\partial L(\mathbf{w}^*,\lambda)}{\partial w_i} = \mathbf{E}\Big(U'(V^*)S_i\Big)\frac{V_0}{S_i^0} - \lambda V_0 \quad \forall\, i = 1, \cdots, m. \tag{4.3}$$

Since $U' > 0$ and $\mathbf{E}(S_i) > 0$, we see that $\lambda > 0$. Hence, define

$$\mu_0 = \frac{\lambda}{\mathbf{E}(U'(V^*))} - 1$$

we have $\lambda = (1+\mu_0)\mathbf{E}(U'(V^*))$ so that

$$S_i^0 = \frac{1}{\lambda}\,\mathbf{E}(U'(V^*)S_i) = \frac{1}{1+\mu_0}\,\frac{\mathbf{E}(U'(V^*)S_i)}{\mathbf{E}(U'(V^*))} = \frac{1}{1+\mu_0}\int_\Omega S_i(\omega)\mathbb{P}(d\omega)) \qquad \forall\, i = 1, \cdots, m.$$

This is the price formula for each individual assets. In addition, if there is a risk-free asset having return rate $\hat{\mu}_0$, its final payoff S must be $S(\cdot) \equiv (1+\hat{\mu}_0)S^0$. Substituting this into the the pricing formula we obtain $\hat{\mu}_0 = \mu_0$.

Suppose X is the final payoff of a derivative security of the underlying assets. Then in the one-period case X is a linear combination of $S_1, \cdots, S_m$. Thus writing $X = \sum_{i=1}^m x_i S_i$ we have, by no arbitrage assumption, its initial price has to be

$$P(X) = \sum_{i=1}^m x_i S_i^0 = \frac{1}{1+\mu_0}\int_\Omega \sum_{i=1}^m x_i S_i \mathbb{P}(d\omega) = \frac{1}{1+\mu_0}\int_\Omega X(\omega)\,\mathbb{P}(d\omega).$$

This completes the proof. □

The pricing equation tells us that the initial unit share price S_i^0 of asset a_i is the discounted expectation of its payoff S_i under the risk-neutral probability measure.

If the number of states in Ω is finite and the utility function satisfies $U(0) = -\infty$ and $U' > 0$ on $(0, \infty)$, then any optimal portfolio $\mathbf{w}^*$ is in the interior of W. This gives an alternative proof for the positive state prices theorem of the finite state model. Clearly, the result here is more general and deeper than that of the positive state prices theorem in the finite state model.

We remark that the resulting risk-neutral probability measure may depend on $V^0 > 0$ and on the choice of U. Nevertheless, all resulting formulas provide the same price for every derivative security.

Proof of Theorem 4.1. We divide the proof into two parts.

(a)(i) Suppose there is a type B arbitrage. Then there is an investment $\hat{\mathbf{w}} = (\hat{w}_1, \cdots, \hat{w}_m)$ such that $(\hat{\mathbf{w}}, \mathbf{1}) = 0$, $\mathcal{V}[\hat{\mathbf{w}}] \geqslant 0$ and $\mathbf{E}(\mathcal{V}[\hat{\mathbf{w}}]) > 0$. Let $\mathbf{w} = (1, 0, \cdots, 0)$. Then for any $\theta \geqslant 0$, $\mathbf{w} + \theta\hat{\mathbf{w}} \in W$. However, since $\mathcal{V}[\theta\hat{\mathbf{w}} + \mathbf{w}] = \theta\mathcal{V}[\hat{\mathbf{w}}] + \mathcal{V}[\mathbf{w}] \geqslant \theta\mathcal{V}[\hat{\mathbf{w}}]$, we have

$$\lim_{\theta\to\infty} \mathbf{E}\Big(U(\mathcal{V}[\mathbf{w}+\theta\hat{\mathbf{w}}])\Big) = \infty.$$

Namely, the investor's problem (4.1) does not have a solution.

(ii) Similarly, one can show that if there is a type A arbitrage, (4.1) also does not have s solution.

Hence, the existence of arbitrage implies the non-existence of a solution to (4.1).

(b) Suppose there is no arbitrage. We show that the investor's problem has a solution.

(i) First of all, we delete one by one those assets which are linear combinations of remaining ones. After finitely many steps, the remaining assets will be linearly independent. By no arbitrage A assumption, all portfolios can be constructed from the remaining ones. Hence, we can assume, without loss of generality, that $\mathsf{a}_1, \cdots, \mathsf{a}_m$ are linearly independent, i.e. (4.2) holds.

(ii) Let $\{\mathbf{w}^j\}_{j=1}^\infty$ be a maximizing sequence, i.e. $\mathbf{w}_j \in W$ for all $j \in \mathbb{N}$ and

$$\lim_{j\to\infty} \mathbf{E}(U(\mathcal{V}[\mathbf{w}^j])) = \sup_{\mathbf{w}\in W} \mathbf{E}(U(\mathcal{V}[\mathbf{w}])) \in (-\infty, \infty].$$

.

Denote by $\|\mathbf{w}\|$ the Euclidean $\mathbb{R}^m$ norm of $\mathbf{w}$. There are two possibilities:

$$\text{(a)}\quad \sup_{j\geqslant 1}\|\mathbf{w}_j\| < \infty, \qquad \text{(b)}\quad \sup_{j\geqslant 1}\|\mathbf{w}_j\| = \infty.$$

Consider case (a) $\sup_j \|\mathbf{w}_j\| < \infty$. In this case, we can select a subsequence, which we still denote by $\{\mathbf{w}_j\}$, such that for some $\mathbf{w}^* \in \mathbb{R}^m$, $\|\mathbf{w}_j - \mathbf{w}^*\| \to 0$ as $j \to \infty$. Then by continuity, $(\mathbf{w}^*, \mathbf{1}) = 1$,

$\mathcal{V}[\mathbf{w}^*] \geqslant 0$; i.e. $\mathbf{w}^* \in W$. In addition, let $M = \sup_{j\geq 1,\omega\in\Omega} \mathcal{V}[\mathbf{w}_j](\omega)$. Then $U(M) - U(\mathcal{V}[\mathbf{w}_j]) \geqslant 0$, so by Fatou's lemma,

$$\begin{aligned}
\mathbf{E}\Big(U(M)) - U(\mathcal{V}[\mathbf{w}^*])\Big) &= \mathbf{E}\Big(\lim_{j\to\infty}[U(M) - U(\mathcal{V}[\mathbf{w}_j])]\Big) \\
&\leqslant \lim_{j\to\infty} \mathbf{E}\Big(U(M) - U(\mathcal{V}[\mathbf{w}_j]))\Big) \\
&= U(M) - \sup_{\mathbf{w}\in W} \mathbf{E}(U(\mathcal{V}[\mathbf{w}])).
\end{aligned}$$

Thus, $\mathbf{E}(U(\mathcal{V}[\mathbf{w}^*])) \geq \sup_{\mathbf{w}\in W} \mathbf{E}(U(\mathcal{V}(\mathbf{w})))$, i.e. $\mathbf{w}^*$ is a maximizer so (4.1) has a solution.

Next consider case (b) $\sup_j \|\mathbf{w}_j\| = \infty$. Let

$$\mathbf{v}_j = \mathbf{w}_j/\|\mathbf{w}_j\| \qquad \forall\, j \in \mathbb{N}.$$

By selecting a subsequence if necessary, we can assume that $\lim_{j\to\infty} \|\mathbf{w}_j\| = \infty$ and for some $\mathbf{v}^* \in \mathbb{R}^m$, $\|\mathbf{v}_j - \mathbf{v}^*\| \to 0$ as $j \to \infty$. Clearly, we have $\|\mathbf{v}^*\| = 1$. Also since $\mathcal{V}[\mathbf{v}_j] = \mathcal{V}[\mathbf{w}_j]/\|\mathbf{w}_j\| \geqslant 0$, $\mathcal{V}[\mathbf{v}^*] \geqslant 0$. Finally, $(\mathbf{v}^*, \mathbf{1}) = \lim_{j\to\infty}(\mathbf{w}_j, \mathbf{1})/\|\mathbf{w}_j\| = \lim_{j\to\infty} 1/\|\mathbf{w}_j\| = 0$. As there is no type B arbitrage, we must have $\mathcal{V}[\mathbf{v}^*] \equiv 0$. In view of (4.2), we conclude that $\mathbf{v}^* = \mathbf{0}$, contradicting to the earlier conclusion that $\|\mathbf{v}^*\| = 1$. This contradiction shows that case (b) does not happen.

In conclusion, the invest's problem has a solution if and only if there is no arbitrage. ,

The proof for the second assertion of the theorem is left as an exercise. This completes the proof. □

Exercise 4.7. *(1) Make a mathematical definition for the terminology "arbitrage-free". You have to state the environment that the definition is to be used. For example, it could be as follows:*

"A **state of economy** is a set $\{S_1^0, \cdots, S_m^0\}$ of positive real numbers and a set $\{S_1, \cdots, S_m\}$ of real random variables on a probability space $(\Omega, \mathcal{F}, P)$. The state of economy is called **arbitrage-free** if (the following holds).... "

(2) Consider the following assets, with an initial payment \$100, it pays as follows:

a_0*: \$105, for sure;*

a_1*: \$95, \$100, or \$ 130, with probability 0.3, 0.4, 0.3, respectively.*

a_2*: \$90, \$100, or \$130, with probability 0.3, 0.4, 0.3, respectively.*

Are there arbitrage in the system consisting of only these three assets?

(3) Suppose the events of the first asset return \$95,100,130 exactly correspond to that of the second asset return \$130,100,90. Find the Markowitz efficient frontier (CAPM's capital market line). Also using $U(x) = \ln x$ *find the log-optimal portfolio.*

Exercise 4.8. *(a) Provide details on the parts (a)(ii) and (b)(i) in the proof of Theorem 4.1.*

(b) Prove the second assertion of Theorem 4.1. *[Hint: Suppose* $\mathbf{w}_1$ *and* $\mathbf{w}_2$ *are two solutions. Consider the weight* $\frac{1}{2}(\mathbf{w}_1 + \mathbf{w}_2)$*.]*

Exercise 4.9. *Suppose there are two investment opportunities:*

a_0*: earn a 20% risk-free interest;*

a_1*: earn a return of* $200\%, 0\%$*, or* -100% *with probability* $0.3, 0.4, 0.3$ *respectively.*

1. *Use* $U(x) = \ln x$ *solving (numerically) the investor's problem with* $V_0 = 10,000$*.*

   Also, find the expected return and certainty equivalent.

2. *Use* $U(x) = \sqrt{x}$ *solving the investor's problem with* $V_0 = 20,000$*.*

3. *Use* $U(x) = -e^{-x}$ *solving the investor's problem with* $V_0 = 1$ *and* $V_0 = 10$ *respectively.*

4. *Use* $U(x) = x^2$ *solving the investor's problem with* $V_0 = 30,000$*.*

Find the risk-neutral probability measures from solutions in part (1), (2) and (3) respectively. Does (4) provide a risk-neutral probability measure?

Finally, explain the four portfolio choices in terms the Markowitz or CAPM model.

Exercise 4.10. *Suppose the utility* U *has the following properties:*

$$U \in C^\infty(\mathbb{R}), \quad U' > 0 \ \text{ on } \ \mathbb{R}, \quad \lim_{x\to\infty} U(x) = \infty, \quad \lim_{x\to\infty} \frac{U(x)}{U(-\theta x)} = 0 \quad \forall\, \theta > 0.$$

Show that the following problem

$$\text{maximize } \ \mathbf{E}\Big(\mathcal{V}[\mathbf{w}]\Big) \quad \text{in} \quad \{\mathbf{w} \in \mathbb{R}^m \mid (\mathbf{w}, \mathbf{1}) = 1\}$$

has a solution if and only there is no-arbitrage.

Use this result prove the Theorem of positive state prices of the finite state model.

4.3 The Log-Optimal Strategy

From now on we investigate multi-period investments. The key difference between single period and multi-period is that the latter needs management, i.e., updating the portfolio at each trading time. We assume that one can buy and (short) sell for any quantity as wish, without any transaction cost.

1. An Investment Wheel Understanding portfolio growth requires that one adopt a long term viewpoint. To highlight the importance of such a viewpoint, we consider an investment wheel shown below. You are able to place a bet on any of the three sectors, named A, B and C respectively. In fact, you may invest different amounts on each of sectors independently. The numbers in sectors denote the winnings (multiplicative factor to your bet) for that sector after the wheel is spun. For example, if the wheel stops with the pointer at the top sector A after a spin, you will receive \$3 for every \$1 you invested on that sector (which means a net profit of \$2); all bets on other sectors are lost.

▼

3

A

B C

6

2

An Investment Wheel

The odds of three sectors are $1/2, 1/3, 1/6$, respectively. We can calculate the expected returns on each bet.

Invest on sector A: $\mathbf{E}(R_A) = 3 * 1/2 - 1 = 50\%$.

Invest on sector B: $\mathbf{E}(R_B) = 2 * 1/3 - 1 = -33\%$;

Invest on sector C: $\mathbf{E}(R_C) = 6 * 1/6 - 1 = 0\%$.

Now suppose we start with an initial capital V_0, say \$100. What is the best strategy to bet so that (in certain statistical sense) after n'th betting, V_n is the largest?

(i) Since sector A is the most attractive, we may intend to invest all money on sector A in each spin. However, we immediately realize that we go broke very quickly and cannot continue the game.

(ii) A second more conservative strategy would be to invest, say, only half of the money on sector A, and holding the other half. That way, if an unfavorable outcome occurs, we are not out of the game entirely. But it is not clear if this is the best that can be done.

2. Analysis

To begin a systematic search for a good strategy, let us limit our investigation to **fixed-proportion** strategies. These are strategies that prescribe proportions to each sector of the wheel, these proportions being used to apportion current wealth among the sectors as bets at each spin.

Let's use $\mathbf{w} = (w_1, w_2, w_3)$ as the proportions of money put on sectors A,B,C respectively. Of course, we need $w_1 \geqslant 0, w_2 \geqslant 0, w_3 \geqslant 0, w_0 := 1 - w_1 - w_2 - w_3 \geqslant 0$.

We denote by V_n the wealth after nth spin. It is a random variable depending on the outcome of the wheel, as well as the betting strategy. Fix a strategy $\mathbf{w}$, the wealth at time $t = n$ is given by

$$V_n = V_{n-1} e^{r_n[\mathbf{w}]}$$

where, denoting $\Omega = \{A, B, C\}$ the occurrence of A,B and C respectively,

$$\begin{aligned} \text{Prob}(A) &= 1/2, & r_n[\mathbf{w}](A) &= \ln(1 + 3w_1 - w_1 - w_2 - w_3), \\ \text{Prob}(B) &= 1/3, & r_n\mathbf{w}](B) &= \ln(1 + 2w_2 - w_1 - w_2 - w_3), \\ \text{Prob}(C) &= 1/6, & r_n[\mathbf{w}](C) &= \ln(1 + 6w_3 - w_1 - w_2 - w_3). \end{aligned}$$

It then follows that

$$V_n = V_0 e^{\sum_{i=1}^n r_i[\mathbf{w}]}, \qquad \ln\left(\frac{V_n}{V_0}\right)^{1/n} = \frac{1}{n}\sum_{i=1}^n r_n[\mathbf{w}].$$

Now sine $r_1, \cdots, r_n$ are i.i.d random variables, the law of large numbers therefore states that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n r_n[\mathbf{w}] = \nu := \mathbf{E}(r_1[\mathbf{w}]) \qquad \text{almost surely.}$$

We can summarize our calculation as follows:

Theorem 4.3 **(Logarithmic Performance)** *If $\{V_j\}_{j=0}^\infty$ is a random sequence of capital values generated by the process $V_k = V_{k-1}e^{r_k}$ where $r_1, r_2, \cdots$ are i.i.d. random variables, then in distribution,*

$$\ln\left(\frac{V_n}{V_0}\right)^{1/n} \quad \longrightarrow \quad \nu := \mathbf{E}(r_1) \qquad \text{as } n \to \infty.$$

Then, formally, we find that

$$V_n \approx V_0 e^{n\nu}.$$

In other words, for large n, the capital grows (roughly) exponentially with n at rate ν.

3. The Log-Optimal Strategy

The foregoing analysis reveals the importance of the number ν. If governs the rate of growth of the investment over a long period of repeated trials. It seems appropriate therefore to select the strategy that leads to the largest value of ν. We see that

$$\nu(\mathbf{w}) := \mathbf{E}(r[\mathbf{w}]) = \mathbf{E}(\ln V_1) - \ln V_0.$$

Hence, if we define our utility function as $U(x) = \ln x$, the problem of maximizing the growth rate ν is equivalent to maximizing the expected utility $\mathbf{E}(U(V_i))$ and using this strategy in every trial. In other words, by using the logarithm as utility function, we can treat the problem as if it were a single-period problem! We find the optimal growth strategy by finding the best thing to do on the first trial with the expected logarithm as our criterion. This single-step view guarantees the maximum growth rate in the long run. Note that this argument is based on the fact that all spins are independent. We summarize our discussion as follows:

The **log-optimal strategy**: *Given the opportunity to invest repeatedly in a series of similar prospects, it is wise to compare possible investment strategy relative to their long-term effects on capital. For this purpose, one useful measure is the expected rate of capital growth. If the opportunities have identical probabilistic properties, then this measure is equivalent to the expected logarithm of a single return, e.g. taking a logarithmic function as the utility function. In other words, long-term expected rate of capital growth can be maximized by selecting a single strategy that maximizes the expected logarithm of return at each trial.*

Although the log-optimal strategy maximizes the expected growth rate, the short run growth rate may differ. We can, however, make some quite impressive statement about the log-optimal strategy.

Theorem 4.4 (Characteristic Property of the Log-optimal Strategy) *Suppose two people start with the same initial capital; one uses the log-optimal strategy and the other does not. Denote the resulting capital streams by $\{V_k^A\}$ and $\{V_k^B\}$, respectively, for the periods $k = 1, 2, \cdots$. Then*

$$\mathbf{E}\left(\frac{V_k^A}{V_k^B}\right) \leqslant 1 \qquad \forall\, k = 1, 2, \cdots .$$

We leave the proof as an exercise.

4. Solution to the Investment Wheel Problem

Let's agree that the log-optimal strategy is used (otherwise what are we going to do?).

We compute the full optimal strategy for the investment wheel problem. Given a strategy $\mathbf{w}$, the logarithm of the expected growth rate is

$$\nu(\mathbf{w}) = \frac{1}{2}\ln(1+2w_1-w_2-w_3) + \frac{1}{3}\ln[1+w_2-w_1-w_3] + \frac{1}{6}\ln(1+5w_3-w_1-w_2).$$

Suppose $\mathbf{w}^*$ is an optimal strategy. First assume that $w_1^* > 0, w_2^* > 0, w_3^* > 0$. Then the partial derivatives of $\nu(\mathbf{w})$ with respect to w_1, w_2, w_3 are zero at $\mathbf{w}^*$. This leads to a system of three equations with three unknowns. One solution to the system is

$$w_1^* = \tfrac{1}{2} \quad , w_2^* = \tfrac{1}{3}, \quad w_3^* = \tfrac{1}{6}.$$

One can check that this is one of the optimal solution. Hence, the logarithm of the expected optimal growth rate is

$$\nu^* = \tfrac{1}{2}\ln\tfrac{3}{2} + \tfrac{1}{3}\ln\tfrac{2}{3} + \tfrac{1}{6}\ln 1 = \tfrac{1}{6}\ln\tfrac{3}{2} = \ln 1.0699$$

Thus, the average growth rate is approximately 107% per period.

Notice that the optimal strategy requires an investment on the unfavorable sector B which pays a negative expected return. This investment serves as a hedge for other sectors—it wins precisely when the others do not[2]. It is like fire insurance on your home, paying when other thins goes wrong.

Exercise 4.11. **(The Kelly Rule of Betting)** *Suppose you have the opportunity of investing in a prospect that will either double your investment, with probability p, or return nothing. Show that the log-optimal strategy is the following* **Kelly rule** [14]*:*

If $p > 1/2$, you should bet a fraction of $2p - 1$ of your wealth; otherwise, bet nothing.

Exercise 4.12. **(Volatility Pumping)** *Suppose there are two alternatives of investment available: (a) A stock that in each period either double or reduce by half, each has 50% chance; (b) hide money under mattress. Show that if one always invest half capital into the stock, then an expected 105.6% growth rate per period can be achieved.*

Here the gain is achieved by the volatility of the stock in a pumping action. If stock goes up in certain period, some of the proceeds are put aside (under the mattress). If on the other hand the stock goes down, additional capital is invested in it. This strategy follows automatically the dictum:

buy low, sell high.

Exercise 4.13. *(a) Prove Theorem 4.4. Hint: Use $e^x \geqslant 1 + x$ for all $x \in \mathbb{R}$.*

(b) In the same setting as Theorem 4.4, show that in measure

$$\liminf_{k\to\infty} \left(\frac{V_k^A}{V_k^B}\right)^{1/k} \geqslant 1.$$

Exercise 4.14. *Consider the investment wheel problem discussed in this section.*

(1) Find all log-optimal strategies;

(2) Use a random number generator simulating the investment wheel and compare graphically portfolio values of different strategies during the first 100 spins;

(3) Suppose the payoff factor for sector B *is changed from* 2 *to* 2.5 *whereas everything else is unchanged. Find all log-optimal strategies.*

[2]The algebraic system for (w_1, w_2, w_3) is actually degenerate. There is a whole family of optimal solutions. An alternative solution is $w_1 = 5/18, w_2 = 0, w_3 = 1/18$. In this solution, nothing is invested on the unfavorable sector; instead, one bets only 1/3 of total wealth, holding the remaining 2/3.

Exercise 4.15. *Let three assets (investment instruments) be the corresponding bets on sector A,C and B of the investment wheel, respectively, and denote by R_1, R_2, R_3 the corresponding return rate of one period (spin).*

1. *Show that the system is arbitrage-free.*
2. *Calculate the statistic parameters: the mean $\mu_i = \mathbf{E}(R_i)$; variance $\sigma_i = \mathbf{Var}(R_i)$; covariance $\sigma_{ij} = \mathbf{Cov}(R_i, R_j)$; correlation $\rho_{ij} = \sigma_{ij}/(\sigma_i\sigma_j)$. Show that the matrix $(\sigma_{ij})_{3\times 3}$ is degenerate.*
3. *Construct a risk-free asset a_0, and eliminate asset a_3 (investment on B sector) from the system.*
4. *Use the Markowitz theory (for a_1 and a_2) and CAMP theory (for $\mathsf{a}_0, \mathsf{a}_1, \mathsf{a}_2$) find the Markowitz efficient frontier, the capital market line, security market line, and market portfolio. Also discuss these mean-variance theories for the case where short selling is forbidden.*
5. *With utility function $U^\alpha(x) = \frac{x^\alpha - 1}{\alpha}$, $\alpha \geqslant 0$ (note $\lim_{\alpha\to 0} U^\alpha(x) = \ln x$). Find (using FindRoot software) the optimal investment strategy. Plot the corresponding μ-σ on the same plane as in (4).*
6. *State whatever your opinion on the investment portfolio after (4) and (5).*

4.4 Log-Optimal Portfolio—Discrete-Time

Now consider the management of a portfolio using the log-optimal strategy, over a time interval $[0, T]$ which is divided into a number of periods of duration Δt. We use notation

$$K = T/\Delta t, \qquad t_k = k\Delta t, \quad \forall\, k = 0, 1, \cdots, K, \qquad \mathbf{T} = \{t_i\}_{i=0}^K.$$

1. Asset's Performance

Suppose there are m assets $\mathsf{a}_1, \cdots, \mathsf{a}_m$ available for investment. We assume that the unit share price S_i^t of asset a_i at time $t \in \mathbf{T}$ obeys

$$R_i^t := \frac{\Delta S_i^t}{S_i^t} := \frac{S_i^{t+\Delta t} - S_i^t}{S_i^t} = \mu_i \Delta t + \Delta z_i^t \tag{4.4}$$

where $\mathbf{u} = (\mu_1, \cdots, \mu_m)$ is a constant vector, $\Delta\mathbf{z}^t := (\Delta z_1^t, \cdots, \Delta z_m^t)$ is a vector valued random variable satisfying

$$\mathbf{E}(\Delta z_i^t) = 0, \qquad \mathbf{Cov}(\Delta z_i^t, \Delta z_j^t) = \sigma_{ij}\Delta t \quad \forall\, i, j = 1, \cdots, m.$$

Here for simplicity, we assume $\mathbf{C} = (\sigma_{ij})_{m\times m}$ is a positive definite constant matrix. Also, all $\Delta\mathbf{z}^{t_0}, \cdots, \Delta\mathbf{z}^{t_K}$ are i. i. d. random variables. We use $\sigma_i = \sqrt{\sigma_{ii}} > 0$ to denote the standard deviation of the return R_i^t.

To make sure the prices are always positive, we assume for simplicity that

$$\left|\mu_i \Delta t + \Delta z_i^t\right| \leqslant \frac{1}{2}.$$

For convenience, we use vector notation $\mathbf{S}^t = (S_1^t, \ldots, S_m^t), \mathbf{R}^t = (R_1^t, \cdots, R_m^t)$.

2. The Log-Optimal Strategy

Fix any $t \in \mathbf{T}$. Suppose the current outcome is $\mathbf{S}^t$ and the portfolio value is V^t (here we ignore the martingale formalism). We consider optimal portfolio that maximize the expected utility. For a weight $\mathbf{w} = (w_1, \cdots, w_m) \in \mathbb{R}^m$ with $\sum_{i=1}^m w_i = 1$, the value of the corresponding portfolio at time $t + \Delta t$ is the random variable

$$\mathcal{V}^{t+\Delta t}[\mathbf{w}] = \sum_{i=1}^m \frac{w_i V^t}{S_i^t} S_i^{t+\Delta t} = V^t \sum_{i=1}^m w_i [1 + R_i^t] = V^t[1 + (\mathbf{w}, \mathbf{R}^t)].$$

It then follows that

$$\mathbf{E}\Big(\ln \frac{\mathcal{V}^{t+\Delta t}[\mathbf{w}]}{V^t}\Big) = \mathbf{E}\Big(\ln[1 + (\mathbf{w}, \mathbf{R}^t)]\Big).$$

For the time period $[t, t + \Delta t)$, using the utility function $U(x) = \ln x$, we can derive the system of equations for the optimal weight $\mathbf{w}^* = (w_1^*, \cdots, w_m^*)$ and corresponding Lagrangian multiplier λ as

$$\sum_{j=1}^m w_i = 1, \qquad \mathbf{E}\Big(\frac{R_i^t}{1 + (\mathbf{w}, \mathbf{R}^t)}\Big) = \lambda \quad \forall\, i = 1, \cdots, m. \tag{4.5}$$

This system has a unique solution since $U(x) = \ln(x)$ is strictly convex and all $R_1^t, \cdots, R_m^t$ are linearly independent (recall $\mathbf{C} = (\sigma_{ij})_{m\times m}$ is assumed to be positive definite). In addition, *the optimal weight* $\mathbf{w}$ *is time-independent* since all $\mathbf{R}^{t_0}, \cdots, \mathbf{R}^{t_k}$ are i.i.d. random variables. We use $\mathbf{R} = (R_1, \cdots, R_m)$ to denote a random variable having the same distribution as each of $\mathbf{R}^{t_0}, \cdots, \mathbf{R}^{t_K}$.

3. Asymptotic Expansion of the Solution

We try to solve the algebraic system (4.5), at least approximately. For this, we assume that Δt is small and Δz_i^t is not too large: for some positive constant M,

$$\mathbf{E}(|\Delta z_i^t|^3) \leqslant M \Delta t^{3/2} \quad \forall\, t, i.$$

The optimal weight $\mathbf{w}$ can be computed approximately as follows. By Taylor's expansion,

$$\frac{R_i}{1 + (\mathbf{w}, \mathbf{R})} = R_i\Big\{1 - (\mathbf{w}, \mathbf{R}) + O(1)(\mathbf{R}, \mathbf{w})^2\Big\}.$$

Then using the definition of $R_i = \mu_i \Delta t + \Delta z_i$ we have

$$\begin{aligned} \lambda &= \mathbf{E}\Big(\frac{R_i}{1 + (\mathbf{w}, \mathbf{R})}\Big) = \mathbf{E}(R_i) - \mathbf{E}(R_i(\mathbf{w}, \mathbf{R})) + O(1)\|\mathbf{w}\|^2 \mathbf{E}(|\mathbf{R}|^3) \\ &= \mu_i \Delta t - \sum_{j=1}^m w_j \sigma_{ij} \Delta t + O(\Delta t^{3/2})[1 + \|\mathbf{w}\|^2]. \end{aligned}$$

Hence, denoting $\mathbf{u} = (\mu_1, \cdots, \mu_m)$, we obtain

$$\mathbf{w} = \mathbf{u}\mathbf{C}^{-1} + \alpha \mathbf{1}\mathbf{C}^{-1} + O(\sqrt{\Delta t}), \quad \alpha := \frac{1 - (\mathbf{u}\mathbf{C}^{-1}, \mathbf{1})}{(\mathbf{1}\mathbf{C}^{-1}, \mathbf{1})}.$$

Thus, the log-optimal strategy is to redistribute the wealth according to the fixed weight $\mathbf{w}$ among assets $\mathsf{a}_1, \cdots, \mathsf{a}_n$ at each trading time $t_0, t_1, t_2, \cdots, t_k$. Denote the corresponding value of the portfolio using the log-optimal strategy at time t by V^t. We then have

$$V^{t_k} = V^{t_{k-1}}[1 + (\mathbf{w}, \mathbf{R}^{t_{k-1}})] = V^0 \prod_{i=0}^{k-1} [1 + (\mathbf{w}, \mathbf{R}^{t_i})].$$

Consequently,

$$\begin{aligned}
\mathbf{E}\Big(\ln \frac{V^T}{V^0}\Big) &= \sum_{i=0}^{K-1} \mathbf{E}\Big(\ln[1+(\mathbf{w},\mathbf{R}^{t_i})]\Big) = \frac{T}{\Delta t}\ \mathbf{E}\Big(\ln[1+(\mathbf{w},\mathbf{R})]\Big) \\
&= \frac{T}{\Delta t}\mathbf{E}\Big((\mathbf{w},\mathbf{R}) - \tfrac{1}{2}(\mathbf{w},\mathbf{R})^2 + O(1)|(\mathbf{w},\mathbf{R})|^3\Big) \\
&= \frac{T}{\Delta t}\Big\{(\mathbf{w},\mathbf{u})\Delta t - \tfrac{1}{2}(\mathbf{w}\mathbf{C},\mathbf{w})\Delta t + O(1)\Delta t^{3/2}\Big\} \\
&= \Big\{(\mathbf{w},\mathbf{u}) - \tfrac{1}{2}(\mathbf{w}\mathbf{C},\mathbf{w}) + O(\sqrt{\Delta t})\ \Big\}T.
\end{aligned}$$

We summarize our calculation as follows.

Theorem 4.5 **(Optimal Growth Rate Theorem)** *When the log-optimal portfolio rebalancing strategy is applied to an investment among m asserts in every trade of period Δt, the portfolio attains its maximum possible expected growth rate among all possible trading strategies. The maximum growth rate is*

$$\nu = \frac{1}{T}\,\mathbf{E}\Big(\ln\frac{V^T}{V^0}\Big) = \frac{1}{\Delta t}\max_{(\mathbf{w},\mathbf{1})=1}\mathbf{E}\Big(\ln[1+(\mathbf{w},\mathbf{R})]\Big) = \nu_{opt} + O(\sqrt{\Delta t})$$

where

$$\begin{aligned}
\nu_{opt} &= \max_{(\mathbf{w},\mathbf{1})=1}\Big\{(\mathbf{w},\mathbf{u}) - \tfrac{1}{2}(\mathbf{w}\mathbf{C},\mathbf{w})\Big\} = (\mathbf{w}_{opt},\mathbf{u}) - \tfrac{1}{2}(\mathbf{w}_{opt}\mathbf{C},\mathbf{w}_{opt}) \\
\mathbf{w}_{opt} &= \mathbf{u}\mathbf{C}^{-1} + \alpha\,\mathbf{1}\mathbf{C}^{-1}, \quad \alpha := \frac{1-(\mathbf{u}\mathbf{C}^{-1},\mathbf{1})}{(\mathbf{1}\mathbf{C}^{-1},\mathbf{1})}.
\end{aligned}$$

Exercise 4.16. *The following are return rates of two stocks in 10 periods. Start with \$100. Using the log-optimal rebalancing strategy find the value of the portfolio at the end of last period.*

R_1	*0.00*	*0.40*	*0.40*	*0.80*	*0.00*	*0.00*	*0.00*	*-0..40*	*0.40*	*-0.40*
R_2	*0.05*	*0.25*	*0.25*	*-0.35*	*0.05*	*0.05*	*0.05*	*0.65*	*0.25*	*0.65*

(Pretend that the statistics generated from above data using appropriate parameter estimators are the ones we got from history.)

Also using the same parameters calculate the growth of a portfolio without any rebalancing.

Exercise 4.17. *Consider two stocks with single period ($\Delta t = 0.49$) return rates R_1 and R_2 respectively. Assume that*

$\text{Prob}(R_1 = 0.2, R_2 = 0.2) = 1/4,\quad \text{Prob}(R_1 = 0, R_2 = 0) = 1/4\quad \text{Prob}(R_1 = 0.05, R_2 = 0.25) = 1/6,$
$\text{Prob}(R_1 = 0.25, R_2 = -0.05) = 1/6\quad \text{Prob}(R_1 = 0.1, R_2 = 0.1) = 1/6$

Find $\mu_1, \mu_2, \sigma_{11}, \sigma_{12}, \sigma_{22}$. Also find the log-optimal portfolio and optimal growth rate.

Exercise 4.18 (**rates**). *The following are distributions of return R from an investment after unit period:*

return R	-.20	-.10	0.00	0.10	0.20	0.30	0.40
probability	0.10	0.10	0.20	0.20	0.20	0.10	0.10

1. *Find the* **expected return** $\mu := \mathbf{E}(R)$ *and risk* σ *where* $\sigma^2 = \mathbf{Var}(R)$;
2. *Find the* **instantaneous return rate** $\hat{\mu} := \ln \mathbf{E}(R+1)$; *show that* $1 + \mu = e^{\hat{\mu}}$.
3. *Find the* **growth rate** $\nu := \mathbf{E}(\ln[1+R])$ *and volatility* $\hat{\sigma}$ *where* $\hat{\sigma}^2 = \mathbf{Var}(\ln[1+R])$.
4. *For* $x = [\nu + \frac{1}{2}\sigma^2]/\mu$, $[\nu + \frac{1}{2}\hat{\sigma}^2]/\hat{\mu}$, $\hat{\mu}/\mu$, *and* $\hat{\sigma}/\sigma$, *find* $x + 1/x - 2$. *Are they all small?*

4.5 Log-Optimal Portfolio—Continuous-Time

Optimal portfolio growth can be applied with any rebalancing period—a year, a month, a week, or a day. In the limit of very short time periods we consider continuous rebalancing, by taking the limit, as $\Delta t \to 0$ of the time discrete case. In fact, there is a compelling reason to consider the limiting situation: the resulting equations for optimal strategies turn out to be much simpler, and as a consequence it is much easier to computer optimal solutions. Hence, even if rebalancing is to be carried out, say weekly, it is convenient to use the continuous-time formulation to do the calculation. The continuous-time version also provides important insight; for example, it reveals very clearly how volatility pumping works.

1. Dynamics of Multiple-Assets

We first extend the discrete-time asset price model to the case of continuous-time model.

From a stochastic point of view, the limit as $\Delta t \to 0$ of the asset price dynamics (4.4) becomes the following version

$$\frac{dS_i}{S_i} = \mu_i \, dt + dB_i, \qquad i = 1, \cdots, m, \quad t > 0$$

where $\mathbf{B} := (B_1, \cdots, B_m)$ is a vector valued **Winner process** (**Brownian motion process**) satisfying

$$\mathbf{E}(dB_i) = 0, \qquad \mathbf{Cov}(dB_i, dB_j) = \sigma_{ij} dt, \quad \sigma_i = \sqrt{\sigma_{ii}}.$$

Note that by the Ito's lemma, the growth rate and its invariance of asset a_i can be calculated as follows: omitting all the indexes i,

$$\begin{aligned}
d(\ln S) &= \tfrac{1}{S}\, dS - \tfrac{1}{2S^2}\, (dS)^2 = \nu \, dt + dB, \\
\nu &:= \mu - \tfrac{1}{2}\sigma^2, \\
\ln S(t) &= \ln S(0) + \nu t + B(t), \\
S(t) &= S(0) e^{\nu\, t + B(t)}, \\
\mathbf{E}\Big(\ln \frac{S(t)}{S(0)}\Big) &= \nu\, t, \\
\mathbf{Var}\Big(\ln \frac{S(t)}{S(0)}\Big) &= \sigma^2\, t, \\
\mathbf{E}\Big(\frac{S(t)}{S(0)}\Big) &= \mathbf{E}(e^{\nu t + B(t)}) = e^{\nu_i t} \int_{\mathbb{R}} \frac{e^x}{\sqrt{2\pi t \sigma^2}} e^{-x^2/(2\sigma^2 t)} dx = e^{\mu\, t}, \\
\mathbf{Var}\Big(\frac{S(t)}{S(0)}\Big) &= \mathbf{E}\Big(\frac{S^2(t)}{S^2(0)}\Big) - \Big(\mathbf{E}\Big(\frac{S(t)}{S(0)}\Big)\Big)^2 = e^{2\mu\, t}(e^{\sigma^2 t} - 1).
\end{aligned}$$

We call ν the (long term) **growth rate** and μ the (instantaneous, or short term) **return rate**.

2. Equation of Dynamic Portfolio

Denote by $\mathbf{w} = (w_1, \cdots, w_m)$ the weight and V the value of a portfolio. It is crucial here to observe the following: The weight change does not affect the portfolio's value since it is only rebalancing, i.e. redistributing the wealth among investment instruments—assets; the value change of the portfolio is due to the unit price change of assets. In the time-discrete version, we have

$$\begin{aligned}\frac{\Delta V}{V} &:= \frac{V^{t+\Delta t} - V^t}{V^t} \\ &= \frac{\sum_{i=1}^m \frac{V^t w_i}{S_i^t} S_i^{t+\Delta t} - V^t}{V^t} \\ &= \sum_{i=1}^m w_i \frac{S_i^{t+\Delta t} - S_i^t}{S_i^t} = \sum_{i=1}^m w_i \frac{\Delta S}{S}.\end{aligned}$$

Hence, in the continuous-time limit, we have the **equation of dynamics portfolio**:

$$\frac{dV}{V} = \sum_{i=1}^m w_i \frac{dS_i}{S_i} = \sum_{i=1}^m \Big\{ w_i \mu_i \, dt + w_i \, dB_i \Big\}.$$

Consequently, by Ito's lemma,

$$d(\ln V) \;=\; \tfrac{1}{V} dV - \tfrac{1}{2V^2}(dV)^2 = \Big\{ (\mathbf{u}, \mathbf{w}) - \tfrac{1}{2}(\mathbf{wC}, \mathbf{w}) \Big\} dt + (\mathbf{w}, d\mathbf{B}).$$

3. Solution

Taking expectation and using the fact that $\mathbf{w}$ and $d\mathbf{B}$ are independent[3], we obtain

$$d\Big(\mathbf{E}(\ln V(t))\Big) = \mathbf{E}\Big((\mathbf{w}, \mathbf{u}) - \tfrac{1}{2}(\mathbf{wC}, \mathbf{w})\Big) dt.$$

Thus, the log-optimal strategy is to

$$\text{maximize} \qquad \mathbf{E}\Big((\mathbf{w}, \mathbf{u}) - \tfrac{1}{2}(\mathbf{wC}, \mathbf{w})\Big).$$

In sophisticated models, all $\mathbf{u}, \mathbf{C}$ are functions of asset's prices $\mathbf{s}$ and time t. Hence, expectation is needed. Nevertheless, in this situation $\mathbf{w}$ is also a function of $\mathbf{s}$ and t, hence to maximize the expectation, we need only to maximize the function inside the expectation. Thus, the above problem is equivalent to

$$\text{maximize} \qquad (\mathbf{w}, \mathbf{u}) - \tfrac{1}{2}(\mathbf{wC}, \mathbf{w}) \qquad \text{in } \{ \mathbf{w} \in \mathbb{R}^m \mid (\mathbf{w}, \mathbf{1}) = 1 \}.$$

If $\mathbf{u} = (\mu_1, \cdots, \mu_m)$ and $\mathbf{C} = (\sigma_{ij})_{m \times m}$ are functions of $\mathbf{s}$ and t, the solution is also a functions of $\mathbf{s}$ and t; namely, the resulting strategy depends on the spot prices $\mathbf{s} = \mathbf{S}^t$ of the assets and time t.

Here we assume that $\mathbf{u} = (\mu_1, \cdots, \mu_m)$ is a constant vector and $\mathbf{C} = (\sigma_{ij})_{m \times m}$ is a constant positive definite matrix, then it is easy to see that the maximum is obtained at a constant vector $\mathbf{w} = \mathbf{w}_{opt}$ where $\mathbf{w}_{opt}$ solves the following

[3] In the probability space where $d\mathbf{B}$ is defined, $\mathbf{w}$ is a constant function. This can be made rigorous if martingale or conditional probability are introduced.

The Log-Optimal Strategy Problem: Find $\mathbf{w}_{opt}$ such that $(\mathbf{w}_{opt}, \mathbf{1}) = 1$ and

$$(\mathbf{w}_{opt}, \mathbf{u}) - \tfrac{1}{2}(\mathbf{w}_{opt}\mathbf{C}, \mathbf{w}_{opt}) = \nu_{opt} := \max_{(\mathbf{w},\mathbf{1})=1} \Big\{(\mathbf{w}, \mathbf{u}) - \tfrac{1}{2}(\mathbf{w}\mathbf{C}, \mathbf{w})\Big\}.$$

At this moment, we see that the maximal growth of $\mathbf{E}(\ln V)$ is attained when we use the constant weight $\mathbf{w} = \mathbf{w}_{opt}$, and the optimal growth rate of $\mathbf{E}(\ln V)$ is ν_{opt}:

$$\mathbf{E}\Big(\ln \frac{V_{opt}(t)}{V(0)}\Big) = \nu_{opt}\, t \quad \forall t > 0.$$

We can calculate the volatility of the portfolio with weight $\mathbf{w}$:

$$\mathbf{Var}\Big(\ln \frac{V(t)}{V(0)}\Big) = \mathbf{E}\Big((\mathbf{w}, \mathbf{B})^2\Big) = (\mathbf{w}\mathbf{C}, \mathbf{w})\, t.$$

In many applications, $\mathbf{u}$ and $\mathbf{C}$ are not constants, in such cases we do not have a closed form for the solution; nevertheless, we have set-up a framework which enable us to find solution, at least numerically.

4. Examples

In the follows we provide a few examples.

Example 4.3. (One asset). Suppose we invest only in one asset whose price obeys $dS = S(\mu dt + \sigma dB_t)$ where B_t is the Standard Brownian motion process. Then $V(t) = V(0)e^{\nu\, t + \sigma B_t}$ and we find the growth rate of the expected logarithmic utility for a single asset investment as

$$\frac{1}{t}\,\mathbf{E}\Big(\ln \frac{V(t)}{V(0)}\Big) = \nu := \mu - \tfrac{1}{2}\sigma^2.$$

Example 4.4. (Multiple Uncorrelated Identical Assets). Suppose for simplicity that we have n assets whose unit share price obeys $dS_i = S_i(\mu\, dt + \sigma dB_i)$ where $\mathbf{Cov}(dB_i, dB_j) = \delta_{ij} dt$. Namely, all these asserts are uncorrelated and have the same probabilistic characteristics. Then the optimal portfolio problem can be easily solved:

$$\mathbf{w}_{opt} = \tfrac{1}{m}\mathbf{1}, \quad \nu_{opt} = \mu - \tfrac{1}{2m}\sigma^2.$$

From the expression of ν_{opt}, one clearly see the **volatility pumping effect**. By investing m assets, the growth rate has increased from ν of a single asset investment to ν_{opt}, a net increase of

$$\nu_{opt} - \nu = \tfrac{1}{2}(1 - \tfrac{1}{m})\sigma^2.$$

The pumping effect is obviously most dramatic when the original variance is high. After being convinced of this, you will be likely to enjoy volatility, seeking it out for your investment rather than shunning it, as you may have after studying the single-period theory.

Volatility is not the same as risk. Volatility is opportunity!

Example 4.5. (Volatility in Action) Suppose a stock has an expected growth rate of $\nu = 15\%$ and a volatility of $\sigma = 20\%$. These are fairly typical values. This means $\mu = 17\%$. By combining 10 such stocks in equal proportions (and assuming they are uncorrelated), we obtain an overall growth rate improvement of $\frac{1}{2}(1 - \frac{1}{10}) * 0.2^2 = 1.8\%$—nice, but not dramatic.

If instead the individual volatilities were $\sigma = 40\%$. The improvement in growth would be 7.1%. At volatilities of 60% the improvement would be 16.2%, which is truly impressive. Unfortunately, it is hard to find 10 uncorrelated stocks with this kind of volatility, so in practice one must settle for more modest gains. Of course, we must temper our enthusiasm with an accounting of the commissions associated with frequent trading.

5. Inclusion of a Risk-Free Asset

Suppose that there is a risk-free asset. We denote it by a_0 and its (continuously compounded) interest rate by $\mu_0 = \nu_0$. Then the unit price of the asset is $S_0(t) = e^{\nu_0 t}$ (assuming $S_0(0) = 1$ for simplicity). This can be put in the differential form

$$dS_0 = \mu_0 \, S_0 \, dt \, .$$

Now we can write a weight as $\hat{\mathbf{w}} = (w_0, \mathbf{w})$ where $w_0 = 1 - (\mathbf{w}, \mathbf{1})$ and $\mathbf{w} \in \mathbb{R}^m$ is arbitrary. Then the log-optimal problem becomes

$$\text{maximize} \quad [1 - (\mathbf{w}, \mathbf{1})]\mu_0 + (\mathbf{u}, \mathbf{w}) - \tfrac{1}{2}(\mathbf{wC}, \mathbf{w}) \qquad \text{in} \ \ \mathbb{R}^m \, .$$

Setting the derivative with respect to w_i, $i = 1, \cdots, m$, equal to zero we obtain a system of equations for the log-optimal portfolio, which we highlight:

Theorem 4.6 (log-optimal portfolio theorem) *When there is a risk-free asset, the log-optimal portfolio has weights for the risky assets that satisfy*

$$\mathbf{wC} = \mathbf{u} - \mu_0 \mathbf{1} \quad \textit{or} \quad \sum_{j=1}^{m} \sigma_{ij} w_i = \mu_i - \mu_0 \quad \forall\, i = 1, \cdots, m.$$

Example 4.6. (A single risky asset and a risk-free asset) Suppose there is a single stock with price S and riskless bond with price B. These prices are governed by

$$dS = S(\mu \, dt + \sigma \, dW_t), \qquad dB = \nu_0 B \, dt \quad (\mu_0 = \nu_0)$$

where W is the standard Winner process. The log-optimal strategy will have a weight on the risky asset $w = (\mu - \mu_0)/\sigma^2$. The corresponding optimal growth and its corresponding variance is then

$$\nu_{opt} = \nu_0 + \frac{(\mu_0 - \mu)}{2\sigma^2}, \qquad \sigma_{opt} = \frac{|\mu - \mu_0|}{\sigma}.$$

Let's plug some numbers in it. Assume that $\mu_0 = \nu_0 = 10\%$, $\mu = 17\%$, $\sigma = 20\%$. Then we find $w = 1.75$ which means we must borrow the risk-free asset to leverage the stock holding. We also find the optimal growth rate $\nu_{opt} = 0.10 + (0.7)^2/(2 * 0.2^2) = 16.125\%$. This is only a slight improvement over the $\nu = \mu - \frac{1}{2}\sigma^2 = 15\%$ expected growth rate of the stock alone. The new standard deviation is $\sigma_{opt} = 0.7/0.20 = 35\%$, which is much worse that that of the stock $\sigma = 20\%$.

From this example, we see that the log-optimal strategy does not give much improvement in the expected value and it worsens the variance (risk) significantly. This shows that the log-optimal approach is not too helpful unless there is opportunity to pump between various stocks with high volatility, in which case, there can be dramatic improvement.

Exercise 4.19. *Suppose there are three stocks and one risk free assets, with following parameters*

Asset	μ_i	σ_{ij}		
a_0	0.10	a_1	a_1	a_3
a_1	0.24	0.09	0.02	0.01
a_2	0.20	0.02	0.07	-0.01
a_3	0.15	0.01	-0.01	0.03

Find the log-optimal portfolio $\mathbf{w}_{opt}$*, its growth rate* ν_{opt}*, and its standard deviation* σ_{opt}*.*

Exercise 4.20. *Suppose there are* m *stocks. Each of them has a price that is governed by geometric Brownian motion. Each has* $\nu_i = 15\%$ *and* $\sigma_i = 40\%$*. However, these stocks are correlated and for simplicity assume that* $\sigma_{ij} = 0.08$ *for all* $i \neq j$*. What is the value of* ν *and* σ *for a portfolio having equal portions invested in each of the stocks?*

4.6 Log-Optimal Pricing Formula (LOPF)

The log-optimal strategy has an important role as a universal pricing asset and the pricing formula, first presented by Long [16], is remarkably easy to derive.

1. The Basic Assumption

Here we summarize what we have studied for the continuous log-optimal model.

We assume that there are m risky assets with prices each governed by geometric Brownian motion, also known as log-normal process, as

$$d\ln S_i = \nu_i\,dt + dB_i \quad \forall\, i = 1, \cdots, m,$$
$$\mathbf{E}(dB_i) = 0, \qquad \mathbf{Cov}(dB_i, dB_j) = \sigma_{ij}dt, \quad \sigma_i = \sqrt{\sigma_{ii}}.$$

There is also a risk-free assert with interest rate $\nu_0 = \mu_0$. This can put in the same form as above by

$$d\ln S_0 = \nu_0\,dt.$$

We call ν_i the **expected (long-term) growth rate** since, omitting indexes,

$$S(t) = S(0)e^{\nu t + B(t)}, \qquad \mathbf{E}\Big(\ln\frac{V(t)}{V(0)}\Big) = \nu\,t.$$

The stochastic equation is equivalent to

$$\frac{dS_i}{S_i} = \mu_i\,dt + dB_i, \qquad \mu_i = \nu_i + \tfrac{1}{2}\sigma_i^2.$$

From this equation, we see that $\mu = \nu + \frac{1}{2}\sigma^2$ is the **(instantaneous) expected return rate**. Accumulatively, we have

$$\mathbf{E}(S(t)) = S(0)e^{\mu t}.$$

A set of weight $\mathbf{w} = (w_1, \cdots, w_m)$ defines a dynamic portfolio in the usual way[4]. Different from the mean-variance theory where weight only refers to initial time, here the weight refers to all time, so rebalancing of portfolios are the key in dynamic portfolio management.

For a dynamic portfolio with constant weight $\mathbf{w}$ on risky assets and weight $1 - (\mathbf{w}, \mathbf{1})$ on risk-free asset, its value V is governed by the geometric Brownian motion

$$d \ln V = \Big(\nu_0[1 - (\mathbf{w}, \mathbf{1})] + (\mathbf{w}, \mathbf{u}) - \tfrac{1}{2}(\mathbf{w}\mathbf{C}, \mathbf{w})\Big)dt + (\mathbf{w}, d\mathbf{B})$$

or equivalently, setting $\mu_0 = \nu_0$,

$$\frac{dV}{V} = \sum_{i=0}^{m} w_i \frac{dS_i}{S_i} = \Big(\mu_0[1 - (\mathbf{w}, \mathbf{1})] + (\mathbf{w}, \mathbf{u})\Big)dt + (\mathbf{w}, d\mathbf{B})$$

where $\mathbf{B} = (B_1, \cdots, B_m)$. When $\mathbf{u}$ and $\mathbf{C}$ are constants, its solution is given by

$$V(T) \quad = \quad V(0) \exp\Big(\{\nu_0[1 - (\mathbf{w}, \mathbf{1})] + (\mathbf{w}, \mathbf{u}) - \tfrac{1}{2}(\mathbf{w}\mathbf{C}, \mathbf{w})\}t + (\mathbf{w}, \mathbf{B})\Big).$$

It has the property that

$$\begin{aligned} \frac{1}{t}\mathbf{E}\Big(\ln \frac{V(t)}{V(0)}\Big) &= \nu := \nu_0[1 - (\mathbf{w}, \mathbf{1})] + (\mathbf{w}, \mathbf{u}) - \tfrac{1}{2}(\mathbf{u}\mathbf{C}, \mathbf{w}), \\ \frac{1}{t}\mathbf{Var}\Big(\ln \frac{V(t)}{V(0)}\Big) &= \sigma^2 := (\mathbf{w}\mathbf{C}, \mathbf{w}). \end{aligned}$$

The log-optimal portfolio is constructed according to the maximization of expected logarithmic utility function. The choice of the logarithmic utility is by nature. The logarithmic utility maximizes the expected overall growth rate. As a result, when $\mathbf{u} = (\mu_1, \cdots, \mu_n)$ is a constant vector and $\mathbf{C} = (\sigma_{ij})$ is a constant matrix, the log-optimal portfolio corresponds to a constant weight $\mathbf{w}_{opt}$ trading strategy. The weight is obtained by solving the system

$$\sum_{j=1}^{n} \sigma_{ij} w_{j,opt} = \mu_i - \mu_0 \quad \forall i = 1, \cdots, m.$$

Denote by V_{opt} the value of the log-optimal portfolio. For each asset a_i we can define their correlation coefficient $\sigma_{i,opt}$ by

$$\sigma_{i,opt} = \mathbf{Cov}\Big(\frac{dV_{opt}}{V_{opt}}, , \frac{dS_i}{S_i}\Big)\Big/ dt\ .$$

Also, we can define

$$\beta_{i,opt} = \frac{\sigma_{i,opt}}{\sigma^2_{opt}}$$

as the best **linear predicator** in **linear regression** of S_i by V_{opt}.

Under these settings, we have the following.

[4]If $\mathbf{u} = (\mu_1, \cdots, \mu_n)$ and $\mathbf{C} = (\sigma_{ij})_{m\times m}$ are functions of t and the spot price $\mathbf{s}$, then $\mathbf{w} = \mathbf{w}(\mathbf{s}, t) : \mathbf{R}^m \times [0, \infty) \to \mathbb{R}^m$. This corresponds to the following trading strategy: At time t, find out the stock price $\mathbf{s} = \mathbf{S}^t$ and rebalance the portfolio according to the weight $\mathbf{w}(\mathbf{s}, t)$. The function $\mathbf{w}(\mathbf{s}, t)$ is calculated at time $t = 0$, based on the geometric motion assumptions.

Theorem 4.7 **(Log-Optimal Pricing Formula (LOPF))** *With $\mu_0 = \nu_0$, there holds*

$$\mu_{opt} - \mu_0 = \sigma^2_{opt}, \qquad \mu_i - \mu_0 = \sigma_{i,opt} = \beta_{i,opt}\sigma^2_{opt} = \beta_{i,opt}(\mu_{opt} - \mu_0),$$

$$\nu_{opt} = \mu_{opt} - \tfrac{1}{2}\sigma^2_{opt}, \qquad \nu_i - \nu_0 = \sigma_{i,opt} - \tfrac{1}{2}\sigma^2_i = \beta_{i,opt}\sigma^2_{opt} - \tfrac{1}{2}\sigma^2_i.$$

Proof. The result follows from the equation for log-optimal portfolio:

$$\mu_i - \mu_0 = \sum_{j=1}^{m} \sigma_{ij} w_j.$$

By definition,

$$\begin{aligned} \sigma_{i,opt} &= \mathbf{Cov}\Big(\frac{dV_{opt}}{V}, \frac{dS_i}{S_i}\Big)\Big/dt \\ &= \sum_{i=1}^{m} w_j \mathbf{Cov}(dB_j, dB_i)\Big/dt = \sum_{j=1}^{m} w_j\sigma_{ij} = \mu_i - \mu_0. \end{aligned}$$

Apply this to the optimal portfolio we also have $\mu_{opt} - \nu_0 = \sigma_{opt,opt} = \sigma^2_{opt}$. The rest equations are derived by playing around these identities. This completes the proof. □

We make a few remarks.

(1) Many indexes such as The Dow Jones Industrial Average can be indeed served as the log-optimal portfolios since they are computed according to the very idea of constant weights rebalancing strategy and the capitalization weight. Hence, we at least have a very reliable reference to look for representatives of log-optimal portfolios.

(2) According to these formulas, the covariance $\sigma_{i,opt}$ of the asset a_i with the log-optimal portfolio completely determines the instantaneous expected return rate μ_i via $\mu_i = \mu_0 + \sigma_{i,opt}(\mu_{opt} - \mu_0)$.

Typically the overall growth rate ν_i is of primary concern in dynamic portfolio management. The pricing formula shows that $\nu_i = \nu_0 + \beta_{i,opt}\sigma^2_{opt} - \frac{1}{2}\sigma^2_i$. The second term $\beta_{i,opt}\sigma^2_{opt}$ is parallel to the CAMP model. However, for large volatility the last term $-\frac{1}{2}\sigma^2_i$ comes to the play and decreases ν.

(3) If we speculate that volatility σ of an asset in the system is proposition to its beta value, i.e. $\sigma = \beta\sqrt{2\kappa}$. Then we see a quadratic relation between the beta value and its overall growth rate ν via

$$\nu = \nu_0 + \sigma_{opt}\beta - \kappa\beta^2. \tag{4.6}$$

Thus is a parabola open downwards; in particular the overall growth rate ν has a ceiling. This is completely different from the capital market line where return rate has no ceiling as long as risks are large enough.

(4) If we were to look at a family of many real stocks, we would not expect the corresponding (ν, β) pair to fall on a single parabola described by (4.6). However, according to the theory discussed, we would expect a scatter diagram of all stocks to fall roughly along such a parabolic curve. Indeed a famous comprehensive study by Fama and French [9] for market returns for decades of data seem to confirm such a statement. This study has been used to argue that the traditional relation predicated by CAMP does not hold, since the return is clearly not linear in β.

(5) Finally, we emphasize that LOPF is independent of how investors behave. It is a mathematical identity. All that matters is whether stock prices really are lognormal (log is normal or geometric Brownian motion) precess as assumed by the model. Since returns are indeed close to being lognormal the log-optimal pricing model must closely hold as well.

2. LOPF and Black-Scholes Equation

The log-optimal pricing can be applied to derivative securities, and the resulting formula is precisely the Black–Scholes equation. Hence we obtain a new interpretation of the important Black–Scholes result and see power of the LOPF. The lop-optimal pricing equation is more general than the Black–Scholes equation since log-optimal pricing applies more generally—not just to derivative assets.

Now we use the LOPF to derive the Black–Scholes equation. For this, it is assumed that there is an underlying system consisting of a risk-free asset a_0 with interest rate $\nu_0 = \mu_0$ and a risky asset a_1 whose price is governed by the geometric Brownian motion process

$$dS = \mu\, S\, dt + \sigma\, S\, dB_t$$

where B_t is the standard Wiener process. Let $F(S,t)$ be the price of an asset a_2 that is a derivative of the underlying asset a_0 and a_1.

First of all, by Ito's lemma, the value F of the derivative security a_2 satisfied the geometric motion

$$\frac{dF}{F} = \frac{1}{F}\Big(\frac{\partial F}{\partial t} + \mu S\frac{\partial F}{\partial S} + \frac{1}{2S^2}\frac{\partial F^2}{\partial S^2}\Big)dt \;+\; \frac{\sigma S}{F}\frac{\partial F}{\partial S}\,dB_t.$$

Thus this asset a_2 has instantaneous return rate

$$\mu_2 := \frac{1}{F}\Big(\frac{\partial F}{\partial t} + \mu S\frac{\partial F}{\partial S} + \frac{1}{2S^2}\frac{\partial F^2}{\partial S^2}\Big).$$

Now consider the system consists of three assets: $\mathsf{a}_0, \mathsf{a}_1, \mathsf{a}_2$, and the corresponding log-optimal portfolio a_{opt}. Since a_2 is a derivative of a_0 and a_1, it cannot enhance the return of a_{opt}. Hence, the log-optimal portfolio is a combination of the asset a_0 and a_1 with weight calculated in Example 4.6. Specifically, the weight is $w = (\mu - \mu_0)/\sigma^2$. That is

$$\frac{dV_{opt}}{V_{opt}} = [1-w]\mu_0 dt + w\frac{dS}{S} = \Big\{[1-w]\mu_0 + w\mu\Big\}dt + \frac{\mu-\mu_0}{\sigma}dB_t.$$

It then follows that

$$\sigma_{2,opt} = \mathbf{Cov}\Big(\frac{dF}{F}, \frac{dV_{opt}}{V_{opt}}\Big)\Big/dt = \frac{(\mu-\mu_0)S}{F}\frac{\partial F}{\partial S}\;.$$

Hence, by the pricing formula, $\mu_2 - \mu_0 = \sigma_{2,opt}$ we obtain

$$\frac{1}{F}\Big(\frac{\partial F}{\partial t} + \mu S\frac{\partial F}{\partial S} + \frac{1}{2S^2}\frac{\partial F^2}{\partial S^2}\Big) - \mu_0 = \frac{(\mu-\mu_0)S}{F}\frac{\partial F}{\partial S}.$$

After simplification, we then obtain the Black-Scholes equation

$$\frac{\partial F}{\partial t} + \mu_0 S\frac{\partial F}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 F}{\partial S^2} = \mu_0 F.$$

We now have three different interpretations of the famous Black-Scholes equation. The first is a no-arbitrage interpretation, based on the observation that a combination of two risky assets can reproduce a risk-free asset and its rate of return must be identical to the risk-free asset. The second is a backward solution process of the risk-neutral pricing formula. The third is that the Black–Scholes equation is a special case of the log-optimal pricing formula.

The power of the log-optimal pricing formula (LOPF) is made clear by the fact that the Block-Scholes equation can be directly derived from LOPF. However, the LOPF is not limited to the the pricing of derivatives—it is a general result.

Exercise 4.21. *Calculate the betas and $\sigma_{i,opt}$ for three stock problem in Exercise 4.19.*

Exercise 4.22. *We know the growth rate ν and volatility σ of a dynamic portfolio with fixed weight $\mathbf{w}$ is given by*

$$\nu(\mathbf{w}) := (\mathbf{w}, \mathbf{u}) - \tfrac{1}{2}\sigma^2(\mathbf{w}), \quad \sigma(\mathbf{w}) = \sqrt{(\mathbf{wC}, \mathbf{w})}.$$

Mimic the mean-variance single period portfolio theory, perform the following:

1. *Describe on the ν-σ plane the* **feasible region** *defined by*

$$D := \{(\sigma(\mathbf{w}), \nu(\mathbf{w})) \mid \mathbf{w} \in \mathbb{R}^m, (\mathbf{w}, \mathbf{1}) = 1\}$$

2. *Find the* **minimum log-variance**

$$\sigma_*^2 = \min_{(\mathbf{w},\mathbf{1})=1} \sigma^2(\mathbf{w}).$$

3. *Find the* **efficient frontier**

$$\nu_{\max}(s) := \max_{\sigma(\mathbf{w})=s} \nu(\mathbf{w}) \qquad \forall s \geq \sigma_*.$$

4. *Prove the* **Two Fund Theorem***:*

   Any point on the efficient frontier can be achieved as a mixture of any two points on that frontier. In addition, the minimum-log-variance portfolio and the log-optimal portfolio can be used.

5. *Now assume that a risk-free asset is included so that*

$$\nu(\mathbf{w}) = [1 - (\mathbf{w}, \mathbf{1})]\nu_0 + (\mathbf{u}, \mathbf{w}) - \tfrac{1}{2}\sigma^2(\mathbf{w}), \qquad \sigma(\mathbf{w}) = \sqrt{(\mathbf{w}C, \mathbf{w})}.$$

   Performing the same analysis as above show the following **One Fund Theorem:**

   Any efficient portfolio can be achieved by a mixture of risk-free and log-optimal portfolio.

   Also show that the Markowitz portfolio lies strictly inside the feasible region.

Exercise 4.23. *A stock price is governed by $dS = \mu S\, dt + \sigma S\, dB_t$ where B_t is the standard Wiener process. Risk-free interest rate is μ_0. Consider the utility $U(x) = x^\alpha$ ($0 < \alpha < 1$). Let w, a constant, be the proposition of wealth invested in stock in an constantly rebalanced portfolio. Show that*

$$\mathbf{E}\Big(U(V(t))\Big) = U(V(0)) \,\exp\Big([\mu_0 + w(\mu - \mu_0) + \tfrac{1}{2}(\alpha - 1)w^2\sigma^2]\alpha t\Big).$$

An invertor, using $U(x) = x^\alpha$ as her utility function, wants to construct a constantly rebalanced portfolio of these assets (stock and risk-free bond) that maximizes the expected value of her power utility at all time $t \geqslant 0$. Show that the proportion w of wealth invested in the stock is a constant given by

$$w = \frac{\mu - \mu_0}{(1 - \alpha)\sigma^2}.$$

Bibliography

[1] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy **81** (1973), 637–654.

[2] N.F. Chen, R. Roll, & S.A. Ross, *Ecpnomic forces and stock market*, Journal of Bussiness, **59** (1986), 383–403.

[3] J.C. Cox, S.A. Ross, and M. Rubinstein, *Option pricing: a simplified approach*, J. Financial Economics, **7** (1979), 229–263.

[4] William G. Cochran, Annals Math. Stats **23** (1952), 315–245.

[5] J. Dubin, Regional Conf. Series on Applied Math. **9** (1973), SIAM.

[6] D. Duffie, Dynamic Asset Pricing Theory, 2nd Ed. Princeton University Press, NJ, 1996.

[7] A. Etheridga, A course in Financial Calculus, Cambridge, University Press, 2002, -521-89077-2.

[8] L. Fisher & J. McDonald, Fixed Effects Analysis of Variance, Academic Press, new York, 1978.

[9] E.F. Fama & K.R. French, *The cross-section of expected stock returns*, Journal of Finance, **47** (1992), 427–465.

[10] W.C. Guenther, Analysis of Variance, Prentice-Hall, New jessy, 1964.

[11] P. R. Holmes, Measure Theorm, Springer-Verlag, New York, 1974.

[12] K. Ito, *On a formula concerning stochastic diffrentials*, Nagoya Mathematics Journal, **3** (1951), 55–65.

[13] I. Karatzas & S. Shreve, Methods of Mathematical Finance, Springer-Verlag, New York, 1998.

[14] J. L. Jr. Kelly, *A new interpretation of information rate*, Bell System Technical Journal, **35** (1956), 917–926.

[15] Donald E. Knuth, The Art of Computer Programming, 2nd Ed. Vol. **2**, Califfornia, Addison-Wesley, 1981.

[16] J.B. Jr. Long, *The Numeraire portfolio,* Journal of Financial Economics, **26** (1990), 29–69.

[17] R.C. Merton, Continuous–Time Finance, Blackwell, Cambridge, MA, 1990.

[18] R. C. Merton, *Theory of rational option pricing*, Bell Journal of Economics and Management Science, **4**(1973), 141–183.

[19] Merrill Lynch, Pierce, Fenner & Smith, Inc., SECURITY RISK EVALUATION, 1994.

[20] J. von Neumann & O. Morgenstern, THEORY OF GAMES AND ECONOMICS BEHAVIOR, Princeton University Prss, NJ, 1944.

[21] B. Oksendal, STOCHSATIC DIFFERENTIAL EQUATIONS,5th Ed. Springer, 2002.

[22] Karl Pearson, Philosophical Magazine, Series 5, **5** (1900), 157–175.

[23] R.J.Jr. Rendleman and B.J. Bartter, *Two–state option pricing*, Journal of Finance, **34** (1979), 193–1110.

[24] Steven Roman, INTRODUCTION TO THE MATHEMATICS OF FINANCE: FROM RISK MANAGEMENT TO OPTION PRICING, Springer-Verlag, New York, 2004.

[25] S.A. Ross, *The arbitrage theory of capital asset pricing*, Journal of Economic Theory **13** (1976), 341–360.

[26] A. M. Ross, INTRODUCTION TO PROBABILITY MODELS, 7th ed. Academic Press, New York, 2000.

[27] Walter Rudin, REAL AND COMPLEX ANALYSIS McGraw-Hill Science/Engineering/Math; 3 edition, 1986.

[28] W. F. Sharpe, INVESTMENT, Prentice Hall, Englewood Clliffs, NJ, 1978.

[29] A.N. Shiryayev, PROBABILITY, Springer-Verlag, New York, 1984.

Index