2014

Keeping Your Options Open: An Introduction to Pricing Options

Ryan F. Snyder
The College of Wooster, rsnyder14@wooster.edu

Keeping Your Options Open: An Introduction to Pricing Options

Independent Study Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Bachelor of Arts in the Department of Mathematics and Computer Science at The College of Wooster

by
Ryan Snyder

The College of Wooster
2014

Advised by:

Dr. Robert Wooster

Abstract

An option is a contract which gives the holder of the option the right, but not the obligation, to buy or sell a given security at a given price, which is called the strike price. For example, suppose Yahoo stock is currently trading at $10 per share. A person could buy an option that gives him or her the ability to purchase shares of Yahoo stock for $12 in one year. If the price of Yahoo stock is greater than $12 in one year, the holder of the option will make money. However, he or she will not use the option if the stock price is less than $\mathbb{1}$2 because it will not be profitable. This situation illustrates that there is a financial advantage to owning options. Thus, options are not handed out for free. This Independent Study introduces a model called the binomial asset pricing model that can be used to price options.

Also, it explores certain mathematical properties necessary to the pricing process, such as sigma-algebras, measurability, conditional expectation, and martingales. The final chapter compares a real-world option price with the price given by the binomial model as well as applying the model in a completely different context—determining whether or not selected players from the 2003 NBA draft were worth their rookie salaries.

Acknowledgements

First, I would like to express my appreciation to my advisor, Dr. Robert Wooster. I do not believe that this Independent Study would have been possible without his knowledge and guidance. His advice on both my research and my career have been invaluable, and I truly consider him a friend. Further, I would also like to thank the rest of the math department. I could not ask for a better group of mentors throughout my development as a math major. Finally, I would like to thank my family, Tony, Kathy, and Anthony Snyder, for supporting me throughout my life.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Important Definitions

A **security** is a financial instrument that represents an ownership or creditor position in a publicly traded corporation. Two examples of securities are stocks (owner) and bonds (creditor). An **option** is a contract which gives the holder of the option the right, but not the obligation, to buy or sell a given security at a given price, which is called the **strike price**. The strike price will be denoted by K in future chapters. There are two main types of options: call options and put options. An option in which the holder is given the right to *buy* a given security at the strike price is a **call option**. A **put option** gives the holder of the option the right to *sell* a given security at the strike price. At this point, some people may wonder what the advantage is to purchasing an option rather than simply buying shares of stock directly. One reason is that the option can significantly increase the holder's profit, depending on the type of option and the behavior of the stock price. This can be illustrated with an example, which is similar to an example in Williams. [11].

Consider a **European call option** in which the holder of the option has the right to buy a certain number of shares for a given price on a given date. The stipulation of European options is that they can only be exercised (used) on the date specified in the contract. On September 26, 2013, a European call option on Buffalo Wild Wings (BWLD) stock has a price of $20. The option expires on October 19, 2013, and the strike price is $90.00. The price of one share of stock in BWLD on September 26 is $110. Imagine Peyton bought the option on 100 shares of BWLD, which would cost $2000. On October 19, Peyton would have the right to purchase 100 shares of BWLD for $9000 (the strike price of $90 per share).

Now, the stock price can either increase or decrease; we will assume in this example that it can go up or down by 40%. First, suppose the stock price of BWLD increases to $154. Peyton will exercise the option, so his net profit per share will be $\$154 - \$90 - \$20 = \$44$. This is calculated by taking the stock price on October 19 and subtracting the strike price and the price of the option per share. Therefore, Peyton's net profit would be $4400, which is a $\frac{4400}{20} = 220\%$ profit on the initial investment in the option. To see the advantage, if Peyton would have invested the $2000 directly into BWLD stock instead of purchasing the option, he could have bought 18 whole shares. Thus, his profit would have been $\$44 \times \$18 = \$792$ on an investment of $\$110 \times \$18 = \$1980$. This is only a $\frac{792}{1980} \times 100 = 40\%$ profit, which is significantly less than the profit made by purchasing the option.

Second, suppose the stock price of BWLD decreases to $66 per share. Peyton will not exercise the option, so he would take a $2000 loss (the price he paid for the option). This time, if Peyton would have used the \$2000 to purchase 18 whole shares of BWLD, his loss would have been $\$44 \times \$18 = \$792$. In other words, this is a 40% loss on his original investment.

This example illustrates that call options can be very risky because both the payoff as well as the loss can be substantially higher, depending on the stock price. Clearly, there must be some other uses of options other than simply trying to make a profit. One important use of options is called hedging. **Hedging** is a risk management strategy used in limiting or offsetting the probability of a loss from fluctuations in the prices of commodities, currencies, or securities. Essentially, it allows risk to be transfered or reduced without the purchase of insurance. For example, a put option allows its holder to sell a stock at a particular price. This reduces the risk of the stock price dropping very rapidly. There are more complicated hedging strategies, but this provides a basic understanding.

Two other important ideas in mathematical finance are **short and long positions** in a security. At face value, these may sound like they refer to the amount of time that a stock is held, but this is not the case at all. In essence, these terms define whether or not an investor borrowed money in order to buy shares of stock (long position) or sold shares of stock that he or she did not own (short position). A short position in a stock may require further explanation. Suppose Ricky short sells 100 shares of stock. The 100 shares of stock are sold just as if Ricky had possessed them himself, and he receives the proceeds from the sale. The shares of stock can come from a broker's own inventory of stock in various companies or from one of the other shareholders of the company. The catch is that by short selling, Ricky has made a promise to deliver those shares back to the lender at some point in time. In other words, Ricky is “betting” that the stock price will decrease, so he can replace the 100 shares of stock at the lower price and make a profit. However, if the stock price increases, he still has to eventually replace the 100 shares, even if he loses money in the process.

1.2 History

As Section 1.1 illustrated, there is a certain financial advantage associated with options because the holder of the option has the *right but not the obligation* to engage in a future action. This means that when it would be beneficial, he or she will engage, but the holder will choose not to exercise the option under disadvantageous circumstances. With this in mind, it seems as though the holder should have to pay to own this type of advantage, and, indeed, this is true. However, the underlying question throughout the history of finance is: how much should a given option cost? This section explores the history of the answer to that question. It gathers its information from Boyle [2] and Korn [8].

We begin in 1900 with the contributions made by a French mathematician, Louis Bachelier. The first ideas about pricing options involved the ability to effectively model the future movements of stock prices. As is the case with any intro probability course, there are two ways to view time: discrete and continuous. Likewise, stock prices can be modeled in either context. Bachelier was the first to attack discrete time modeling by using a **random walk**. Informally, a random walk is a path that consists of a sequence of random steps. For example, Bachelier modeled stock prices by using coin tosses at each discrete time interval. The movement of the stock price (up or down) is determined by whether heads or tails was flipped. Generalizing to continuous time involves increasing the frequency of the time intervals, or, intuitively, constantly flipping the coin. Bachelier also laid the foundation for this generalization when he showed that as the frequency of the time intervals increases, the random walk starts to behave like **Brownian motion**. Brownian motion is the continuous time analog to a random walk, but we will not explore it here. Unfortunately, Bachelier's contributions to option pricing were not recognized during his lifetime as no one paid much attention to his thesis until the 1950's.

Up until the 1950's, it had been assumed that asset-price movements followed a normal distribution. However, one drawback to this assumption is that the normal distribution allows for negative values. Even though stocks can become worthless, stock prices cannot be negative. Paul Samuelson, an American economist, was interested in option pricing, and he was further intrigued when he came across Bachelier's unknown book, "rotting in the library at the University of Paris" [2]. Samuelson made two main contributions to the field. First, he assumed that stock returns follow a *lognormal distribution*, which solved the problem of negative stock prices because the lognormal distribution does not allow negative numbers. Second, he derived a formula for the price of an option, which involved several variables including the expected return on the stock and the expected return on the option. However, he could not figure out a way to solve for or estimate these variables. If he had, he would have solved the option pricing problem.

As time and research progressed, it was learned that Samuelson did not have the entire formula correct. He accurately identified the expected return on the stock as an important variable, but the expected return on the option was replaced with a discount rate (also unknown). In 1967, Ed Thorp realized that he could set both the expected return on the stock as well as the discount rate equal to the riskless interest rate (explained in Chapter 3). In fact, the resulting formula is still used today and is known as the **Black-Scholes Formula**. However, Thorp was unable to mathematically prove that his formula was correct. Instead, he used his new formula to trade approximately $\$100,000$ in the options market, and, in the end, broke even. He decided that the formula had proven itself in action.

This leaves us questioning why this famous formula is not called the *Thorp Formula*? The reason is because the final touches to the option pricing problem were made by Fischer Black and Myron Scholes. Black began working on the problem in 1965 and had made steady progress, eventually deriving a differential equation, and its solution would be the option price. He took some time off, as the frustration surrounding the equation began to build. Then, Black and Scholes joined forces in a final attempt to solve the equation. Finally, in 1969, they solved it, and unlike Thorp, were able to prove that this was indeed the solution. After nearly 70 years, the problem originally worked on by Bachelier had been solved. Interestingly, Fischer Black, who helped derive the most significant formula in finance history, had never taken a formal economics or finance course in his life.

The work of Black and Scholes was eventually published in 1973, and it was then that Thorp first saw the resemblance. When asked why he did not go public with his formula 6 years earlier, he said he was planning on setting up a hedge fund and using the formula as a competitive advantage. Today, most (if not all) of the credit is attributed to Black and Scholes, but how does Thorp feel about that? He wrote: "Black-Scholes was a watershed-it was only after seeing their proof that I was certain that this was the formula-and they justifiably get all the credit. They did two things that are required: They proved the formula (I didn't) and they published it (I didn't)" [2].

The Black-Scholes formula was a shocking but complex development in the world of finance. It involves very complicated mathematical ideas that are very difficult to grasp. The next development in option pricing came when a professor, John Cox, wanted to teach his students about option pricing. He believed that the math involved in the Black-Scholes Formula was too complicated. So, in 1979, Cox, Stephen Ross, and Mark Rubenstein converted the continuous time concepts to discrete time, eliminating the calculus from the formula. They assumed that, in discrete time, the stock price movements follow a binomial distribution. As many intro probability courses point out, as the number of trials of a binomial distribution increases, it can be approximated by a normal distribution. This fact links the discrete time binomial model to the continuous time Black-Scholes model.

1.3 Outline of the Independent Study

The remainder of this Independent Study (I.S.) will read in the following way. Chapter two will outline some important probability concepts that must be understood before introducing the pricing model. The main concepts include sigma-algebras, measurability, and conditional expectation given a sigma-algebra. Chapter three develops the binomial asset pricing model, also called the Cox-Ross-Rubinstein model. First, it introduces the notation and assumptions associated with the model. Then, it provides a formal example about why there is only one "efficient" price for any given option, demonstrating why any other price will not work. Finally, Chapter 3 identifies two distinct probability measures and explains their significances. Chapter 4 gives a detailed explanation of an important property contained in the binomial model—martingales. There are certain properties associated with martingales that are essential in the field of financial mathematics, and this chapter explores them. Lastly, Chapter 5 develops two real-world applications of the binomial model. First, it compares the theoretical price (given by the binomial model) to the actual real-world price of two options—Netflix and Johnson and Johnson. Second, it uses the binomial model as a tool to decide whether or not the top ten picks in the 2003 NBA draft were worth their rookie salaries. This I.S. will conclude at the end of Chapter 5 with a discussion of the pros and cons of the binomial asset pricing model.

Chapter 2

Probability

2.1 Sigma-Algebras

Before we begin our study of mathematical finance, we must understand a few important probability concepts, which can be further explored in [1]. The first is the idea of a σ-algebra (sigma-algebra). Recall from probability that we begin with a sample space, denoted by Ω (Omega), which is the set of all possible outcomes of a random experiment.

Definition A collection $\mathcal{F}$ of subsets of Ω is called a σ-algebra if the following four properties hold:

(1) $\emptyset \in \mathcal{F}$

(2) $\Omega \in \mathcal{F}$

(3) For any event A, $A \in \mathcal{F} \implies A^c \in \mathcal{F}$

(4) $A_1, A_2, ... \in \mathcal{F} \implies \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ and $\bigcap_{i=1}^{\infty} A_i \in \mathcal{F}$.

In words, this definition is often stated as: *σ-algebras are closed under compliments, countable unions, and countable intersections*. However, one point to be made is that the second part of (4) need not be included in the definition because it follows from complements, countable unions, and DeMorgan's Law.

In introductory probability, the most common σ-algebra is the power set of the sample space because it contains all combinations of the events (all combinations of the subsets of Ω). Therefore, the power set of the sample space is necessarily always a σ-algebra. However, it is not the case that a collection of subsets of the sample space *must* be the power set in order to be a σ-algebra. An example will make this clearer.

Example 1. *Let $\Omega = \{a, b, c\}$. What is the power set of Ω? Next, consider the event $B = \{a\}$. What is the smallest σ-algebra containing B?*

The largest σ-algebra of Ω is its power set:
$\mathcal{F} = \{\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}, \emptyset\}$. It is clear that this is a σ-algebra. The power set of Ω contains B, but it also contains many unnecessary events when attempting to name the smallest σ-algebra containing B. The smallest σ-algebra containing B is: $\mathcal{F}_B = \{\{a\}, \{b, c\}, \{a, b, c\}, \emptyset\}$; this is called the σ-algebra *generated* by B. In order to generate $\mathcal{F}_B$, we start with B. Then, the complement of B ($B^c = \{b, c\}$) must also be contained in $\mathcal{F}_B$. Next, we must union those two events, which produces $\Omega = \{a, b, c\}$, and complement the resulting set, which is $\emptyset$. Now, $\mathcal{F}_B$ satisfies all parts of the definition, so it is a σ-algebra that is smaller than the power set of Ω.

At this point, one might question how this is applicable to financial mathematics. We utilize σ-algebras often in the later chapters, but we will attempt to get an intuitive understanding here. The binomial model, introduced in Chapter 3, uses sequences of coin tosses to depict periods of time, just like the work of Bachelier. In other words, a coin is flipped at the beginning of each new time period, and, depending on the result of the toss, the value of a stock either increases or decreases. Therefore, it is useful to predict the value of a stock in the next period using the information that we have gathered through the current period. Suppose it is currently period 1, and we are attempting to predict the price of a given stock in period 2 using the information we have collected through period 1. In order to mathematically illustrate that we "know" all the information through time 1, we use a σ-algebra: $\mathcal{F} = \{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}\}$. In this case, our sample space Ω is all of the possible outcomes of two successive coin tosses. For example, if we know which events in $\mathcal{F}$ happened, we would know whether or not the event $\{HH, HT\}$ happened. In other words, we would know whether or not the first toss was heads. Likewise, knowing whether or not the event $\{TH, TT\}$ happened tells us if the first toss was tails. However, neither of these events give us any information on the second coin toss, so this σ-algebra only gives us information through the first coin toss (first time period).

2.2 Measurability

The next important probability concept is the idea of measurability. Recall the definition of a random variable from introductory probability. A **random variable** is a function X from the sample space to the real numbers. This is illustrated in Figure 2.1.

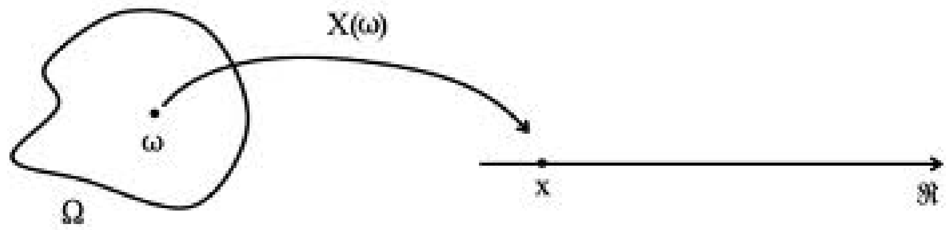

Figure 2.1: A random variable is a function from the sample space to the real numbers [5].

However, one piece of the definition of a random variable that is often left out is the fact that a random variable must also be measurable.

Definition A random variable X is **measurable** with respect to a σ-algebra $\mathcal{F}$ if for all real numbers a, $\{\omega : X(\omega \geq a\} \in \mathcal{F}$. In this case, we would say X is $\mathcal{F}$-measurable.

This definition can be difficult to understand. To start, let us consider an example.

Example 2. *Suppose the sample space is the real numbers from zero to one:* $\Omega = [0, 1]$. *Also, suppose we have the simplest possible* σ*-algebra:* $\mathcal{F} = \{\emptyset, \Omega\}$, *and the probability measure* $\mathcal{P}$ *is the euclidian measure of length. Finally, assume that* X *is a random variable such that* $X(\omega) = \omega$.

Clearly, the length of the interval [0,1] is 1, so this probability measure meets the requirement that $\mathcal{P}(\Omega) = 1$. Now, we choose $a = \frac{1}{2}$ (this could be any real number). In order for X to be $\mathcal{F}$-measurable, it must be the case that any event (subset of Ω) that X maps to a real number greater than $\frac{1}{2}$ is an element of $\mathcal{F}$.

In this example, subintervals are events because Ω is a closed interval. Since X is the identity function, the interval $\left[\frac{1}{2}, 1\right]$ gets mapped to real numbers greater than or equal to $\frac{1}{2}$. However, this event is not an element of $\mathcal{F}$ because $\mathcal{F}$ only has two elements. In mathematical notation:

$$X^{-1}\left(\left[\frac{1}{2}, \infty\right)\right) = \left\{\omega \in \Omega : X(\omega) \geq \frac{1}{2}\right\} = \left[\frac{1}{2}, 1\right] \notin \mathcal{F}.$$

Therefore, X is not $\mathcal{F}$-measurable.

Example 3. *Now, we will slightly alter Example 2 by changing the random variable. Let X be a random variable such that $X(\omega) = c \in \mathbb{R}$.*

In this example, we must consider two cases. First, assume that $a \leq c$. Since, every real number in the interval [0,1] gets mapped to c, it is the case that all of Ω satisfies the quality of being mapped to a real number greater than or equal to a. In mathematical notation:

$$X^{-1}([a, \infty)) = \{\omega \in \Omega : X(\omega) \geq a\} = [0, 1] = \Omega \in \mathcal{F},$$

where X^{-1} refers to the inverse image. Second, assume that $a > c$. Since every real number in the interval [0,1] gets mapped to c, there are no subsets of Ω that satisfy the quality of being greater than or equal to a. In other words, $\emptyset$ satisfies the condition. In mathematical notation:

$$X^{-1}([a, \infty)) = \{\omega \in \Omega : X(\omega) \geq a\} = \emptyset \in \mathcal{F}.$$

Since the definition of measurable holds for all values of a, X is $\mathcal{F}$-measurable. In fact, since $\mathcal{F}$ is the trivial σ-algebra, the only $\mathcal{F}$-measurable functions are constants.

Next, let us consider an example using the coin toss space that we have used previously.

Example 4. *In this example, let X be a random variable that gives the number of tails in the first two coin tosses. Also, let the sample space Ω be all possible results of three coin tosses. Finally, let*

$$\mathcal{F}_1 = \{\emptyset, \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}, \{HHH, HHT, HTH, HTT\}, \\ \{THH, THT, TTH, TTT\}\}$$

and we want $\mathcal{F}_2$ to be the σ-algebra generated by the following sets:

$$\emptyset, \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}, \{HHT, HHH\}, \\ \{HTH, HTT\}, \{THH, THT\}, \{TTH, TTT\}.$$

In this example, $\mathcal{F}_1$ is the σ-algebra that gives the information known through one coin flip and $\mathcal{F}_2$ reveals the information after two tosses of the coin. In other words, if we know which events in $\mathcal{F}_2$ have occurred, we know the results of the first two coin tosses. Intuitively, we know that X is not $\mathcal{F}_1$-measurable because $\mathcal{F}_1$ only tells us the result of the first coin flip, and X is the number of tails after *two* coin tosses. How can we see this by using the definition of measurable? Let us give a name to the set used in the definition: $A_a = \{\omega \in \Omega : X(\omega) \geq a\}$. If $a = 1.5$, then $A_{1.5} = \{\omega \in \Omega : X(\omega) \geq 1.5\}$. In words, $A_{1.5}$ is the set of all events that the random variable maps to a real number greater than or equal to 1.5. These events include the sequences of coin tosses that resulted in at least two tails. However, none of the subsets of $\mathcal{F}_1$ give this information. In other words, knowing that one of the events in $\mathcal{F}_1$ occurred does not tell whether or not $A_{1.5}$ occurred. In mathematical notation:

$$X^{-1}([1.5, \infty)) = \{TTT, TTH\} \notin \mathcal{F}_1.$$

Thus, X is not F_1-measurable.

On the other hand, suppose we know which events in $\mathcal{F}_2$ happened. This would tell us whether or not $A_{1.5}$ happened, so X is $\mathcal{F}_2$-measurable. For example, suppose that we know that the event $\{HTH, HTT\}$ occurred. Since X gives the number of tails in two coin tosses, the result of the third coin toss is irrelevant. In this event, the first two coin tosses are the same, namely HT, so we know $X = 1$. Intuitively, the notion of measurability in the discrete setting can be stated as follows: *For each event in a σ-algebra, suppose we know whether or not that event occurred. Then, if a random variable is measurable with respect to that σ-algebra, we can calculate the value of that random variable.*

2.3 Conditional Expectation

The final and most important probability concept that we will discuss is conditional expectation. Conditional probability is a topic covered in introductory probability, which allows us to calculate the probability that an event ocurrs given the fact that another event ocurred. In other words, conditional probability is a number that behaves exactly like any other probability. Expected value (expectation of a random variable) is also a core topic of introductory probability, which states the value that we would "expect" a random variable to take on if an infinite number of trials of the random experiment were performed. In other words, the expectation of a random variable is also a number. Conditional expectation, on the other hand, is a random variable that combines both of these ideas. Conditional expectation states the value we would "expect" a random variable to take on given some other information. However, it depends on the results of a random experiment. Before we can give the definition of conditional expectation, we first must introduce the idea of **restricted expectation**.

Definition Suppose that X is a random variable and B is an event. The expectation of X restricted to B is defined as:

$$E[X;B] = \sum_{\omega \in B} X(\omega)\mathbb{P}(\{\omega\}).$$

However, another way that this definition is commonly stated is:

$$E[X;B] = E[X\mathbb{1}_B], \tag{2.1}$$

where $\mathbb{1}_B$ is the **indicator function**. The indicator function of a subset B of Ω

$$\mathbb{1}_B : \Omega \to \{0,1\}$$

is defined as:

$$\mathbb{1}_B(\omega) = \begin{cases} 1 & \text{if } \omega \in B \\ 0 & \text{if } \omega \notin B. \end{cases}$$

In this context, this function indicates whether or not an element of the sample space is an element of a given subset of the sample space. All elements of Ω that are also elements of B take on the value 1, while all elements of Ω that are not elements of B take on the value 0. For example, suppose $\Omega = \{a, b, c, d, e\}$ and $B = \{a, d\}$. Then, $\mathbb{1}_B(a) = 1$, $\mathbb{1}_B(b) = 0$, $\mathbb{1}_B(c) = 0$, and so on. This paper will use (2.1) in future calculations involving restricted expectation.

Now, we are ready to give the definition of conditional expectation.

Definition Suppose there exist finitely (or countably) many sets $B_1, B_2, \ldots$, all having positive probability, such that they are pairwise disjoint, Ω is equal to their union, and $\mathcal{F}$ is the σ-algebra generated by all the B_i's. Then, the **conditional expectation** of a random variable X given $\mathcal{F}$ is:

$$E[X|\mathcal{F}](\omega) = \sum_i \frac{E[X; B_i]}{\mathbb{P}(B_i)} \mathbb{1}_{B_i}(\omega).$$

Now, we will look at an example that again involves a coin toss space, assuming that the coin is fair.

Example 5. *Suppose* $\Omega = \{HH, HT, TH, TT\}$, *and let* $\mathcal{F}_1 = \{\emptyset, \Omega, \{HH, HT\}, \{TH, TT\}\}$. *Also, let* S *be a random variable such that:*

$$S(\omega) = \begin{cases} 16 & \text{if } \omega = HH \\ 4 & \text{if } \omega = HT \text{ or } TH \\ 1 & \text{if } \omega = TT. \end{cases}$$

What is the conditional expectation of S *given* $\mathcal{F}_1$, *assuming that the coin is fair?*

In this example, $\mathcal{F}_1$ is the σ-algebra that gives all of the information up through one toss of the coin. What should the answer look like? The conditional expectation of S given $\mathcal{F}_1$ is a random variable that depends on the results of the coin tosses, so we will have one answer if the first toss was heads and a different answer if the first toss was tails. This is because events in $\mathcal{F}_1$ indicate the result of the first toss. Now, in order to begin the calculations, we first must determine B_i. In this example, the two sets whose union is Ω but also are pairwise disjoint are $\{HH, HT\}$ and $\{TH, TT\}$, so these will be B_1 and B_2, respectively. Also, since the coin is fair and there are only two events, the probability of each of them is $\mathbb{P}(B_i) = \frac{1}{2}$. The summation in the definition of conditional expectation can now be written as:

$$\begin{aligned} E[S|\mathcal{F}_1](\omega) &= \sum_{i=1}^{2} \frac{E[S;B_i]}{\mathbb{P}(B_i)} \mathbb{1}_{B_i}(\omega) \\ &= \frac{E[S;B_1]}{\frac{1}{2}} \mathbb{1}_{B_1}(\omega) + \frac{E[S;B_2]}{\frac{1}{2}} \mathbb{1}_{B_2}(\omega). \end{aligned} \tag{2.2}$$

The next thing to remember is the definition of restricted expectation, (2.1), which says that $E[S;B] = E[S\mathbb{1}_B]$. This means we must analyze the random variables $S\mathbb{1}_{B_1}$ and $S\mathbb{1}_{B_2}$. These random variables are easily computed because of the simplicity of the indicator function. The calculation yields

$$S\mathbb{1}_{B_1}(\omega) = \begin{cases} 16 & \text{if } \omega = HH \\ 4 & \text{if } \omega = HT \\ 0 & \text{if } \omega = TH \text{ or } TT, \end{cases} \qquad S\mathbb{1}_{B_2}(\omega) = \begin{cases} 0 & \text{if } \omega = HH \text{ or } HT \\ 4 & \text{if } \omega = TH \\ 1 & \text{if } \omega = TT. \end{cases}$$

Taking the expectation of these two random variables is also straightforward, remembering that we are flipping a fair coin. Thus,

$$E[S\mathbb{1}_{B_1}] = 16 \cdot \frac{1}{4} + 4 \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} = 5,$$
$$E[S\mathbb{1}_{B_2}] = 0 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} = \frac{5}{4}.$$

We can now substitute this answer into the earlier calculation of the conditional expectation, picking up where we left off from (2.2):

$$\begin{aligned} E[S|\mathcal{F}_1](\omega) &= \frac{E[S\mathbb{1}_{B_1}]}{\frac{1}{2}}\mathbb{1}_{B_1}(\omega) + \frac{E[S\mathbb{1}_{B_2}]}{\frac{1}{2}}\mathbb{1}_{B_2}(\omega) \\ &= 10\mathbb{1}_{B_1}(\omega) + \frac{5}{2}\mathbb{1}_{B_2}(\omega). \end{aligned}$$

Another way to write the asnwer that may help in the intuitive understading of conditional expectation is to write it as a piecewise function, shown below:

$$E[S|\mathcal{F}_1](\omega) = \begin{cases} 10 & \text{if } \omega \in B_1 \\ \frac{5}{2} & \text{if } \omega \in B_2. \end{cases}$$

By writing the solution as a piecewise function, it may be easier to realize that the conditional expectation is a random variable that depends on the value of ω. If $\omega \in B_1$ (which means the first toss was heads), the conditional expectation of S given $\mathcal{F}_1$ is 10, whereas if $\omega \in B_2$ (the first toss was tails), instead, the conditional expectation of S given $\mathcal{F}_1$ is $\frac{5}{2}$. This is exactly how we described the answer at the beginning of this example: we got one answer when the first coin toss was heads and a different answer when the first toss was tails. *To reiterate, the reason for this is because the σ-algebra $\mathcal{F}_1$ on which we are conditioning corresponds to knowing the result of the first toss, so the conditional expectation depends on the result of the first toss.*

The final things to cover in this section are a few important properties of conditional expectation, which we present in Theorem 1, but first, we must learn two more definitions.

Definition Two σ-algebras, $\mathcal{F}$ and $\mathcal{G}$, are **independent** if and only if for all $A \in \mathcal{F}$ and all $B \in \mathcal{G}$, $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

Definition Let X be a random variable, and let σ_X be the smallest σ-algebra for which X is measurable. X is **independent** of a σ-algebra $\mathcal{F}$ if and only if σ_X and $\mathcal{F}$ are independent.

Theorem 1. *Let X and Y be random variables. Then, the following properties hold:*

(i) ***Linearity of conditional expectations****: For all constants c_1 and c_2, we have*

$$E[c_1 X + c_2 Y|\mathcal{F}] = c_1 E[X|\mathcal{F}] + c_2 E[Y|\mathcal{F}].$$

(ii) ***Taking out what is known****: If X is $\mathcal{F}$-measurable, then*

$$E[XY|\mathcal{F}] = XE[Y|\mathcal{F}].$$

(iii) ***Iterated conditioning**: If $\mathcal{G} \subset \mathcal{F}$, then*

$$E[E[X|\mathcal{G}]|\mathcal{F}] = E[X|\mathcal{G}].$$

(iv) ***Independence**: If X is independent of $\mathcal{F}$, then*

$$E[X|\mathcal{F}] = E[X].$$

We will not take the time to prove this theorem, but we will try to gain an intuitive understanding. First, we consider an ordinary conditional expecation, say $E[X|\mathcal{F}]$. We can think of this expectation as the best prediction of X given some information conveyed by $\mathcal{F}$. Thinking of it in this way allows us to get a handle on Theorem 1. Part (*i*) is a common property that extends from the linearity of expectations. Intuitively, it says that the predicted value of $X + Y$ is the sum of the predicted values. Property (ii) states that if X is $\mathcal{F}$-measurable and we are given $\mathcal{F}$, then we know the value of X. In other words, since X is $\mathcal{F}$-measurable, our best predictor of X is itself. Another important fact contained in property (ii) is that for conditional expectations with respect to a $\mathcal{F}$, any $\mathcal{F}$-measurable random variables act like constants because they can be taken inside or outside the conditional expectation. The iterated conditioning property states that the average of the predicted value of X is the average value of X. In this property, we are essentially *estimating an estimate*. Since $\mathcal{G}$ is a subset of $\mathcal{F}$, $\mathcal{F}$ gives us more information than $\mathcal{G}$. In this property, we are first predicting X using some information in $\mathcal{G}$. Then, we are estimating that prediction based on the information in $\mathcal{F}$. According to (iii), that prediction is the same as simply predicting X using the information in $\mathcal{G}$. Finally, property (iv) is the independence portion of the theorem. If X is independent of $\mathcal{F}$, then knowing $\mathcal{F}$ does not give us any additional information about X. Thus, the best prediction of X using the information in $\mathcal{F}$ is simply the expected value of X without using any information. These properties will be used repeatedly in Chapter 4.

Chapter 3

The Binomial Model

3.1 Introducing the Model

The main purpose of the binomial asset pricing model is to identify the no-arbitrage price of an option. **Arbitrage** is formally defined as a trading strategy that begins with no money, has zero probability of losing money, but has a positive probability of making money. Essentially, an arbitrage means that there is no risk. In an arbitrage, there can be two outcomes: a person could make money with a certain probability or a person could break even with another probability. However, in an arbitrage, the person can implement the trading strategy without worrying about losing money.

This chapter begins by considering the one-period binomial model, which is regarded as the simplest. However, before we can begin the analysis of the mathematics, we must first discuss the variables, notation, and assumptions involved in the model. We use Shreve [10] to build the model in this chapter. The *one-period model* begins at time zero and ends at time one. At time zero, we are given the initial price of a stock, which is denoted S_0. At time one, the stock price can either increase to $S_1(H)$ or decrease to $S_1(T)$. The subscript in the notation for the stock prices stands for the period. In the one-period model, there will only be two distinct subscripts (time 0 or time 1), but in the multi-period model, there can theoretically be an unlimited number of distinct subscripts because more and more time periods can be considered. As mentioned earlier, a coin flip decides whether or not the price of the stock increases, but it is not necessarily a fair coin. For example, the stock may be more likely to increase than to decrease in value. Therefore, the probability of flipping a head (and the stock price increasing as a result) will be higher than the probability of flipping a tail. The only assumption made about the coin is that the probability of flipping a head, p, and the probability of flipping a tail, $q = 1 - p$, are both strictly positive. As can be expected, H and T in the notation for the time one stock price symbolize the outcome of heads and tails, respectively.

Naturally, the next question is: by how much does the initial stock price increase or decrease? This question is answered by two more model parameters: the up factor, u, and the down factor, d. The one important assumption about these parameters is that $u > d$, which is clear from the names *up* factor and *down* factor.

Now, we will consider an example of how the model works, so far. Suppose that $S_0 = 4$, $u = 2$, and $d = \frac{1}{2}$. This means that the stock price will increase to 8 with probability p (head is flipped), and will decrease to 2 with probability $1 - p$ (tail is flipped). This situation is depicted in Figure 3.1.

The interest rate in the model is r. Interest rates drive the idea that a

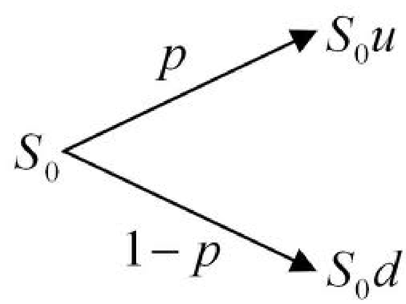

Figure 3.1: The one-period binomial model [3].

dollar today is worth more than a dollar in the future because it is possible to invest that dollar, earn interest, and, thus, have more than a dollar in the future. This is known as the time value of money or an investor's time preference. In the model, a dollar invested in the money market at time zero will yield $(1 + r)$ dollars at time one. Further, in the multi-period model, that dollar will earn $(1 + r)^n$ dollars at period n. In order to understand this concept, suppose $r = 0.10$ (10% interest). Then, we know that at time one, the dollar invested at time zero will earn an additional 10 cents, resulting in \$1.10 at time one. This can be written as $\$1.00 \times (1 + 0.10)^1$. Continuing, at time two, we will earn another 10% interest. However, in the model, we are dealing with **compound interest**, meaning that we earn interest on the interest we have earned in previous periods. Thus, at time two, we earn 10% interest on the money we currently have, \$1.10, which yields \$1.21. This can also be computed using the formula given earlier in this paragraph:

$$(1 + .10)^2 = \$1.21.$$

Similarly, borrowing a dollar from the money market at time zero will yield a debt of $(1 + r)^n$ dollars at time n. The underlying assumption here is that the interest rate for investing is the same as the interest rate for borrowing. Further, we also assume that $r > 0$.

The interest rate relates to the other two parameters in the model through the following inequality:

$$0 < d < 1 + r < u. \tag{3.1}$$

This inequality, Equation (3.1), is known as the *no-arbitrage condition*, which implies that if it does not hold, then there will be an arbitrage. Let us examine this more closely by considering each piece of the inequality. The first is obvious, $d > 0$, because we assumed earlier that the stock prices are positive. Therefore, we cannot multiply the initial stock price by a negative factor. Next, $d < 1 + r$ must hold because of a trading strategy that would result in no risk. Consider the situation where $d > 1 + r$, that is, the inequality does not hold. Then, a person could borrow money from the money market, and use that money to buy shares of stock (the number of shares bought at time zero will be denoted by Δ_0). If the stock price increases, it is clear that the person will make money because $u > 1$ (consequence of $r > 0$). If the stock price decreases by a factor of d, he or she will still be able to pay off the money borrowed as well as the interest, while still making a profit. The reason for this is because the return on the interest rate is less than the amount the stock price decreased. This creates an arbitrage.

Finally, suppose the last inequality does not hold, meaning that $1 + r > u$.

This situation will allow the person to sell shares short and invest the resulting income in the money market, where it will earn interest. If the stock price decreases, he or she will have no problem replacing the shares because they are worth less now at time one than when they were purchased at time zero. On the other hand, if the stock price increases, the stock is worth more than when it was purchased. However, since the return on the interest rate is greater than the amount that the stock price increased, he or she can replace the shares of stock and still pocket the remaining money as a profit. Again, there is an arbitrage because it was possible to make money without the possibility of losing money. We have now argued that if there is to be no arbitrage in the model, then the inequalities in Equation (3.1) must hold. However, this is a biconditional statement because it is also true that if the inequality holds, then there is no arbitrage. Before we can prove this form of the claim, we must introduce the wealth equation.

In this model, wealth has two components: the cash position and the position in the stock, both in dollars. The reason for this is we are only worried about pricing the option, so the only two things that will matter are investments in the stock and money market. First, the position in the stock is straightforward, given by $\Delta_0 S_1$. In words, the stock position at time one is found by multiplying the number of shares bought at time zero by the stock price at time one. The cash position is a little more complex. The assumption made in order to arrive at the formula for the cash position is that we invest the money that is left over after we purchase Δ_0 shares of stock at time 0. Thus, interest is earned on the remaining portion of the wealth from time zero. The formula for the cash position at time one is: $(1 + r)(X_0 - \Delta_0 S_0)$, where X_0 is the initial wealth at time zero. In words, we subtract the amount of money that we spent in order to purchase Δ_0 shares of stock at time zero (which cost S_0 per share) from our initial wealth X_0. This yields the amount that can be invested in the money market, thus, yielding more cash in the next period. The two components are then combined, resulting in the wealth equation: $X_1 = \Delta_0 S_1 + (1 + r)(X_0 - \Delta_0 S_0)$. However, the wealth equation is generalized for an n-period model by:

$$X_{n+1} = \Delta_n S_{n+1} + (1 + r)(X_n - \Delta_n S_n). \tag{3.2}$$

Now, we are able to return to the claim: if $0 < d < 1 + r < u$, then there is no arbitrage in the model. However, remember that there is a stronger claim represented, which we state in the next theorem.

Theorem 2. *Let d represent the down factor, r represent the interest rate, and u represent the up factor. Then, there is no arbitrage in the binomial model if and only if* $0 < d < 1 + r < u$.

Earlier in this section, we proved that if there is no arbitrage in the model, then Equation (3.1) must hold. Now, we want to show that if Equation (3.1) holds, then there is no arbitrage. Essentially, we want to show that if $X_0 = 0$ and X_1 is given by the wealth equation, then we cannot have $X_1 > 0$ with positive probability unless $X_1 < 0$ with positive probability, also. In words, if we start with no initial wealth at time zero, then we cannot have the possibility to make money (positive wealth) at time one unless there is a possibility that we lose money (negative wealth). It should be noted that this must be the case regardless of the number of shares of stock purchased. We will tackle this proof by manipulating the wealth equation.

Proof. Assume $0 < d < 1 + r < u$, $X_0 = 0$, and that $X_1 > 0$ with positive probability. Then, it must be the case that either $X_1(H) > 0$ or $X_1(T) > 0$. Without loss of generality, assume $X_1(H) > 0$. According to Equation (3.2), $X_1(H) = \Delta_0 S_1(H) + (1 + r)(X_0 - \Delta_0 S_0) > 0$. By substituting $S_1(H) = uS_0$ and $X_0 = 0$, we get: $\Delta_0 uS_0 + (1 + r)(-\Delta_0 S_0) > 0$. Factoring out $\Delta_0 S_0$ yields: $\Delta_0 S_0(u - (1 + r)) > 0$. Since an assumption of the model is that stock prices are always positive, $S_0 > 0$. Also, since $0 < d < 1 + r < u$, $(u - (1 + r)) > 0$. Thus, $\Delta_0 > 0$ in order for the inequality $\Delta_0 S_0(u - (1 + r)) > 0$ to be true.

Now, we analyze $X_1(T)$ in a similar fashion, knowing that $\Delta_0 > 0$. Again, we begin by using Equation (3.2) to define $X_1(T)$, continue by substituting $S_1(T) = dS_0$, and end by factoring. The steps are shown below:

$$\begin{aligned} X_1(T) &= \Delta_0 S_1(T) + (1 + r)(X_0 - \Delta_0 S_0) \\ &= \Delta_0 dS_0 + (1 + r)(-\Delta_0 S_0) \\ &= \Delta_0 S_0(d - (1 + r)). \end{aligned}$$

Again, $S_0 > 0$, but since $0 < d < 1 + r < u$, $(d - (1 + r)) < 0$. Thus, since $\Delta_0 > 0$ from the above argument, $\Delta_0 S_0(d - (1 + r)) < 0$. In other words, $X_1(T) < 0$. Therefore, we cannot have $X_1 > 0$ with positive probability unless $X_1 < 0$ with positive probability, also. Thus, there is no arbitrage. □

Finally, we will summarize a few more important assumptions in the following list:

1. The interest rate for investing is the same as the interest rate for borrowing.
2. We have unlimited short selling of stock.
3. There are no transactions costs (including bid-ask spreads) associated with the purchase of shares or investments made in the money market.
4. Our buying and selling is on a small enough scale that it does not affect the market.
5. At any time, the stock can only take on two possible values in the next period.

3.2 Example: Pricing an Option

Now, we have enough tools to understand how to price an option in the one-period binomial model. Consider a situation where one share of stock in company X is priced at \$4, and the interest rate in the money market is 25% (very unrealistic but effective for the example). After one period, experts predict that the stock price could either increase to \$8 or decrease to \$2. Also, the strike price of a European call option is \$6. Our goal is to combine activity in the stock and money market to allow our portfolio value to be exactly equal to the value of the option at time one. For example, at time one, if the stock price increased to \$8, the call option will allow us to purchase one share of stock in Company X for the lower price of \$6. Thus, the value of the option in this scenario is \$2 (\$8 − \$6). On the other hand, if the stock price decreases to \$2, the call option is worthless (\$0) because we could simply purchase one share of stock at the current price of \$2 rather than the \$6 price that the option allows. Therefore, by investing money in the money market and buying a certain number of shares of stock, we want our wealth (portfolio value) to equal \$2 if the stock price increases and \$0 if the stock price decreases. Specifically, this process is called **replicating the option**.

The first thing to do in this example is to convert the words into the parameters of the binomial model. In the previous paragraph, we were given the following information:

- $S_0 = \$4$,
- $r = \frac{1}{4}$,
- $S_1(H) = \$8$,
- $S_1(T) = \$2$,
- $K = 6$.

Recall the wealth equation for the one-period model, Equation (3.2): $X_1 = \Delta_0 S_1 + (1+r)(X_0 - \Delta_0 S_0)$. We want to substitute the known values into the wealth equation and solve for the unknown variables. It is important to remember that *wealth* is a random variable with two elements in its support. The value of the wealth random variable is determined by the coin flip, so this wealth equation is actually two-fold ($X_1(H)$ and $X_1(T)$). These two equations are presented below with the substitutions already made:

$$\begin{aligned}
X_1(H) &= \Delta_0 S_1(H) + (1+r)(X_0 - \Delta_0 S_0) = 2 \\
&= 8\Delta_0 + \frac{5}{4}X_0 - 5\Delta_0 = 2 \\
&= 3\Delta_0 + \frac{5}{4}X_0 = 2, \\
X_1(T) &= \Delta_0 S_1(T) + (1+r)(X_0 - \Delta_0 S_0) = 0 \\
&= 2\Delta_0 + \frac{5}{4}X_0 - 5\Delta_0 = 0 \\
&= -3\Delta_0 + \frac{5}{4}X_0 = 0.
\end{aligned}$$

Now, we have two equations and two unknowns, so we are able to solve the system of equations:

$$\begin{aligned}
X_1(H) &= 3\Delta_0 + \frac{5}{4}X_0 = 2, \\
X_1(T) &= -3\Delta_0 + \frac{5}{4}X_0 = 0, \\
\Delta_0 &= \frac{1}{3}, X_0 = \frac{4}{5} = 0.80.
\end{aligned}$$

The solution to the system of equations is $\frac{1}{3}$ shares of stock in Company X and initial wealth of \$0.80. These quantities replicate the option because no matter what happens to the stock, at time one, our portfolio value will equal the value of the option. Therefore, the no-arbitrage price of the option is \$0.80. At this point, it may be a bit unclear as to why this is the no-arbitrage price, so in order to fully understand this fact, we will consider the two situations where the price of the option is more than and less than the calculated fair price.

First, suppose the price of the option is \$1.00 (higher than the no-arbitrage price). The *seller* of the option, Sammy, would receive \$1.00 for the option from the *buyer*, Brian, and invest \$0.20 in the money market, which will be worth \$0.25 at time one. At time zero, Sammy has \$0.80, which is, not coincidentally, the initial wealth that we calculated in the system of equations earlier. Sammy now wants to purchase $\frac{1}{3}$ shares of stock in Company X, which will cost \$1.33 ($4 \times \frac{1}{3}$), so he must borrow \$0.53 from the money market. At time one, Sammy will owe \$0.67. If the stock price increases, his $\frac{1}{3}$ shares of stock will be worth \$2.67 ($8 \times \frac{1}{3}$). Selling his stake in Company X enables him to pay off his \$0.67 debt and still have \$2.00. Since the stock price increased, Brian will want to exercise the option, which gives him the right to purchase one share of stock for \$6.00 instead of \$8.00. However, Sammy is able to honor this deal because he still has \$2.00 after repaying his debt, so after receiving Brian's payment of \$6.00, Sammy is able to purchase one share of stock for \$8.00 to give to Brian and still break even. If the stock price decreases, his $\frac{1}{3}$ shares of stock will be worth \$0.67 ($2 \times \frac{1}{3}$), allowing him to pay off his debt in the money market. Since the stock price decreased, Brian will not exercise the option, so Sammy has no further obligation. However, Sammy still has an additional \$0.25 from his original money market investment. Therefore, he was able to start with no wealth and end with \$0.25 no matter what happens to the stock. This is an arbitrage opportunity, which exists because the price of the option is higher than the no-arbitrage price.

Second, suppose the price of the option is \$0.50 (lower than the no-arbitrage price). Brian should sell short $\frac{1}{3}$ shares of stock in Company X to generate \$1.33 of income. This allows him to buy the option for \$0.50, and still have \$0.83 left over. Next, Brian should invest \$0.53 in money market account A and invest the remaining \$0.30 in money market account B. It should be noted that both money market accounts earn the same interest rate. The only reason two accounts are used is for "book-keeping" purposes. If the stock price increases to \$8.00, Brian can exercise the option, allowing him to purchase one share of stock in Company X for \$6.00 instead of \$8.00. Then, he can use $\frac{1}{3}$ of those shares to replace the $\frac{1}{3}$ shares he sold short at time zero. At this point, he has −\$6.00 because he exercised the option, but has not made any money. However, he still has $\frac{2}{3}$ shares of stock remaining, so he can sell those at the current stock price of \$8.00 to receive \$5.33. Brian now has −\$0.67, but the money market investment of \$0.53 in money market account A at time zero has grown to \$0.67, so he is back to even. If the stock price decreases to \$2.00, the option is worthless, but Brian still needs \$0.67 to replace the $\frac{1}{3}$ shares of stock that he bought at time zero. Luckily, his investment in money market account A allows him to do this, so, again, Brian breaks even. However, in both scenarios, his investment of \$0.30 in money market account B has grown to \$0.37. Therefore, Brian was able to start with no wealth and end with \$0.37 regardless of what happens to the value of the stock. This is an arbitrage opportunity, which exists because the price of the option is lower than the no-arbitrage price.

3.3 Risk-Neutral vs. Actual Probabilities

The main thing that most investors consider when debating how to build a trading strategy is risk. If a stock is particularly "risky," investors are less likely to purchase shares in that company. There are several different ways to determine the risk associated with a particular stock. However, a very basic method involves analyzing the stock's average rate of growth with respect to a *risk-free* investment. For example, is it always a good investing decision to purchase shares of a stock if the stock price is expected to increase? The answer is no. It is possible that the stock price may increase, but the investor would have made more money if he or she had invested in the money market, instead, earning the interest rate. This would be the case if the rate of interest is higher than the growth rate of the stock price. In other words, we want to compare the stock price at time zero with the **discounted stock price** at time one. This means that we want to compare the stock price at time zero with the expected stock price at time one in time zero money. This idea relates back to the idea of an investor's time preference, described earlier.
In mathematical notation:

$$S_0 < \frac{1}{1+r}E[S_1] = \frac{1}{1+r}[pS_1(H) + qS_1(T)],$$

where p and q represent the actual probabilities that the stock price will increase or decrease, respectively. What does this inequality say? Essentially, it says that the average rate of growth of the stock should be strictly greater than the average rate of growth of a money market investment, which is necessary because otherwise an investor would never want to assume the risk of purchasing the stock. If, on the other hand, the inequality was flipped, we would be better off investing in the money market because its average rate of growth would be higher than that of the stock price. However, in this model, in order to price an option, we want investors to be neutral about risk. The reason for this is because our goal is to replicate the option. By using the risk-neutral probabilities, no matter how a man invests in the stock and money markets, he will receive the same average rate of return (the interest rate, r). Therefore, this allows the equality to hold in the inequality from above:

$$S_0 = \frac{1}{1+r}[\tilde{p}S_1(H) + \tilde{q}S_1(T)], \tag{3.3}$$

where $\tilde{p}$ and $\tilde{q}$ are the **risk-neutral probabilities** that make this equation true. Since $\tilde{p}$ and $\tilde{q}$ are probabilities, they sum to one, which means $\tilde{q} = 1 - \tilde{p}$. Thus, we can solve for $\tilde{p}$ directly.

$$\begin{aligned} S_0 &= \frac{1}{1+r}[\tilde{p}S_1(H) + \tilde{q}S_1(T)] \\ &= \frac{1}{1+r}[\tilde{p}uS_0 + \tilde{q}dS_0] \\ &= \frac{S_0}{1+r}[\tilde{p}u + \tilde{q}d] \\ &= \frac{S_0}{1+r}[\tilde{p}u + (1-\tilde{p})d] \\ &= \frac{S_0}{1+r}[\tilde{p}(u-d) + d]. \end{aligned}$$

Therefore, $\tilde{p} = \frac{1+r-d}{u-d}$, and $\tilde{q} = 1 - \tilde{p} = \frac{u-1-r}{u-d}$.

Now, let us back track to the example presented in section 3.2 about pricing the option. We reached a point where we needed to solve a system of two equations with two unknowns, which is a fairly simple calculation. Recall that the system says that our portfolio value at time one must equal the value of the option at time one, regardless of the result of the coin toss. Notice that we are again using the discounted prices because we want to compare prices in the present time period (time zero). Thus, the system of equations can be written as:

$$X_0 + \Delta_0\left(\frac{1}{1+r}S_1(H) - S_0\right) = \frac{1}{1+r}V_1(H),$$
$$X_0 + \Delta_0\left(\frac{1}{1+r}S_1(T) - S_0\right) = \frac{1}{1+r}V_1(T).$$

However, another way to solve this system of equations is to multiply the first by $\tilde{p}$ and the second by $\tilde{q} = 1 - \tilde{p}$ and then add them. This yields:

$$X_0 + \Delta_0\left(\frac{1}{1+r}[\tilde{p}S_1(H) + \tilde{q}S_1(T)] - S_0\right) = \frac{1}{1+r}[\tilde{p}V_1(H) + \tilde{q}V_1(T)].$$

As noted earlier, $\tilde{p}$ and $\tilde{q}$ were defined so that $S_0 = \frac{1}{1+r}[\tilde{p}S_1(H) + \tilde{q}S_1(T)]$. This means that the term multiplying Δ_0 is zero. Thus, we now have a much simpler formula for X_0:

$$X_0 = \frac{1}{1+r}[\tilde{p}V_1(H) + \tilde{q}V_1(T)].$$

We can also solve for Δ_0 by subtracting the two original equations:

$$X_0 + \Delta_0\left(\frac{1}{1+r}S_1(H) - S_0\right) = \frac{1}{1+r}V_1(H),$$
$$-[X_0 + \Delta_0\left(\frac{1}{1+r}S_1(T) - S_0\right) = \frac{1}{1+r}V_1(T)].$$

The only terms that differ in these two equations are those that depend on the coin toss, which means that this subtraction allows all other terms to drop out, yielding:

$$\Delta_0 = \frac{V_1(H) - V_1(T)}{S_1(H) - S_1(T)}.$$

Now, we can substitute the numbers from the example in section 3.2 to illustrate that we arrive at the same answer. First, $\tilde{p} = \frac{1+r-d}{u-d} = \frac{1}{2}$, so $\tilde{q} = \frac{1}{2}$, also.

$$\begin{aligned} X_0 &= \frac{1}{1+r}[\tilde{p}V_1(H) + \tilde{q}V_1(T)] \\ &= \frac{4}{5}\left[\frac{1}{2} \times 2 + \frac{1}{2} \times 0\right] \\ &= \frac{4}{5} = \$0.80, \\ \Delta_0 &= \frac{V_1(H) - V_1(T)}{S_1(H) - S_1(T)} \\ &= \frac{2-0}{8-2} = \frac{1}{3}. \end{aligned}$$

Finally, we can introduce the **risk-neutral pricing formula** for the one-period binomial model:

$$V_0 = \frac{1}{1+r}[\tilde{p}V_1(H) + \tilde{q}V_1(T)],$$

which is found simply by replacing X_0 with V_0 in the earlier formula because we want our portfolio to equal the value of the option in each period of time. This formula can be generalized to n-periods by

$$V_n = \frac{1}{1+r}[\tilde{p}V_{n+1}(H) + \tilde{q}V_{n+1}(T)]. \tag{3.4}$$

Essentially, this formula states that the price of the option in period n equals the discounted expected value of the value of the option in the next period under the risk-neutral probability measure, and as was shown earlier, any other price (higher or lower) leads to an arbitrage opportunity. This may be concerning to some readers because it seems as though the actual probabilities should be taken into account when pricing the option. However, this is not the case. The reason we use the risk-neutral probabilities is because we want the hedge to give us a portfolio that agrees with the value of the option, regardless of the result of the coin toss. In other words, given any possible stock price path, we want the hedge to work. In order to understand this, we must recall that we found the correct hedge amounts by solving a system of two equations with two unknowns, but there were no probabilities in the system. We introduced the risk-neutral probabilities into the system to allow some of the terms to cancel, but the actual probabilities would not allow such cancellations. Since this is an important fact about the model, we will summarize the preceeding discussion. *Since we want our hedge to work regardless of whether the stock price goes up or down, the actual probabilities of the up and down move are irrelevant. Rather, the size of the move is more important. In the binomial model, the price of the option depends on the set of possible stock price paths, but not on the probabilities that each path ocurrs.*

Chapter 4

Martingales

So far, we have studied some important probability concepts as well as introducing the binomial model. In this chapter, we will combine the previous two chapters to explore an important concept related to the binomial model—a martingale. Before giving the definition of a martingale, we first must understand the idea of an adapted stochastic process. An **adapted stochastic process** in discrete time is simply a sequence of random variables $X_0, X_1, X_2, \ldots$, with each X_n depending only on the first n coin tosses (and X_0 constant). For example, this sequence is said to be adapted to the information at time n. More generally, a sequence of random variables X_n is adapted if X_n is $\mathcal{F}_n$-measurable for all values of n. This chapter uses information from Bass [1] and Shreve [10].

4.1 The Martingale Property

Definition Suppose we have a sequence of σ-algebras $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \cdots$. A **martingale** is a sequence of random variables $M_0, M_1, \ldots, M_N$ such that:

(1) M_n is adapted to $\mathcal{F}_n$.

(2) $E[M_n^2] < \infty$ for all values of n.

(3) $M_n = E[M_{n+1}|\mathcal{F}_n]$ for $n = 0, 1, ..., N - 1$.

In words, a martingale has no tendency to rise or fall. Given the first n coin tosses (information up to time n is given by $\mathcal{F}_n$), the expected value of the martingale at time $n + 1$ is its value at time n. This is known as the "one-step-ahead" *martingale property*. However, there are other versions of the martingale property. Consider the following statement for $n \leq N - 2$. By conditioning on the first $n + 1$ coin tosses, the martingale property states:

$$M_{n+1} = E[M_{n+2}|\mathcal{F}_{n+1}]. \tag{4.1}$$

Then, we can further condition both sides by the first n coin tosses because we can always condition on more information. However, the right side of the equation can then be simplified further by the iterated conditioning property discussed earlier in Chapter 2. Now, equation (4.1) reads:

$$E[M_{n+1}|\mathcal{F}_n] = E[E[M_{n+2}|\mathcal{F}_{n+1}]|\mathcal{F}_n] = E[M_{n+2}|\mathcal{F}_n].$$

By analyzing the left-hand side of the equation, we can see that the martingale property states that this equals M_n directly. Thus, we are left with the "two-step ahead" version of the martingale property:

$$M_n = E[M_{n+2}|\mathcal{F}_n].$$

By continuing this conditioning and iterating process, we would arrive at the "multi-step ahead" version of the martingale property, which states:

$$M_n = E[M_m|\mathcal{F}_n].$$

whenever $0 \leq n \leq m \leq N$. This says that the expectation of a martingale in any future period m given the information up to the current period n is the value of the martingale in the current period n. Now, we can take the expectation of both sides of the equation and use the iterated conditioning property to get:

$$E[M_n] = E[E[M_m|\mathcal{F}_n]] = E[M_m].$$

This reinforces the fact that the expectation of a martingale is constant over time (has no tendency to rise or fall).

The most natural application of martingales is in the context of gambling, but there are also numerous other stochastic applications, including the binomial asset-pricing model. We will use gambling to further understand the martingale property. Let M_n be a gambler's total wealth after gambling in a fair game n times, and we will condition on the gambler's wealth after each gamble. In other words, $\mathcal{F}_n$ gives all of the information about his past gambles up to time n. A *fair game* means that the odds of winning and losing are equal. Thus, regardless of the outcomes of past gambles, the expected *change* in his wealth is zero because his wealth will increase or decrease by the same amount with the same probability. Therefore, the best prediction of his wealth after his next gamble ($n + 1$ gambles) is his current wealth (at time n), just as the martingale property states. The "multi-step-ahead" version states that at any point in time, the gambler's expected wealth is the same as his initial wealth because, on average, in a fair game, he will neither win nor lose money.

Now, we will look at an application to the binomial asset-pricing model. Recall equation (3.3) from chapter 3:

$$S_0 = \frac{1}{1+r}[\tilde{p}S_1(H) + \tilde{q}S_1(T)],$$

which can be generalized to take the form:

$$S_n = \frac{1}{1+r}[\tilde{p}S_{n+1}(H) + \tilde{q}S_{n+1}(T)].$$

Another way to write this is to recognize that S_n is equal to the discounted expected value of S_{n+1} with respect to the risk-neutral probability measure:

$$S_n = \frac{1}{1+r}E[S_{n+1}|\mathcal{F}_n].$$

Next, if we divide both sides of the equation by $(1+r)^n$, we get:

$$M_n = \frac{S_n}{(1+r)^n} = \frac{1}{(1+r)^{n+1}}\tilde{E}[S_{n+1}|\mathcal{F}_n].$$

This equation identifies the discounted stock price as a martingale. In other words, it says that the best prediction of the discounted stock price at time $n+1$ with respect to the risk-neutral probabilitiies is the discounted stock price at time n. This is another reason that the risk-neutral probabilties are chosen because using the actual probabilties will not preserve this equality.

In order to see that this is indeed true, we will go through a numerical example and follow it with the proof that the discounted stock price is a martingale. Recall from section 3.3 that $\tilde{p} = \tilde{q} = \frac{1}{2}$ and $\frac{1}{1+r} = \frac{4}{5}$. Thus, we want to check that the martingale equation $\left(\frac{4}{5}\right)^n S_n = \left(\frac{4}{5}\right)^{n+1} \tilde{E}[S_{n+1}|\mathcal{F}_n]$ holds regardless of the value of n. We will verify this for a couple values of n. First, suppose $n = 2$ and we know which events in $\mathcal{F}_2$ have occurred, namely $\{HHH, HHT\}$. In other words, we know the first two coin tosses both resulted in heads. Then,

$$\begin{aligned}
\left(\frac{4}{5}\right)^2 S_2 &= \left(\frac{4}{5}\right)^3 \tilde{E}[S_3|\mathcal{F}_2] \\
\left(\frac{16}{25}\right)16 &= \frac{64}{125}\left(32 \times \frac{1}{2} + 8 \times \frac{1}{2}\right) \\
10.24 &= 10.24\checkmark.
\end{aligned}$$

Next, suppose $n = 3$ and we know which events in $\mathcal{F}_3$ have occurred, namely $\{HHTH, HHTT\}$. In other words, we know that the first three coin tosses were HHT. Then,

$$\begin{aligned}
\left(\frac{4}{5}\right)^3 S_3 &= \left(\frac{4}{5}\right)^4 \tilde{E}[S_4|\mathcal{F}_3] \\
\left(\frac{64}{125}\right)8 &= \frac{256}{625}\left(16 \times \frac{1}{2} + 4 \times \frac{1}{2}\right) \\
4.10 &= 4.10\checkmark.
\end{aligned}$$

So far, we have shown that the martingale equation holds for $n = 2$ and $n = 3$. However, since we cannot analytically show that it holds for all values of n, we must provide a mathemtical proof.

Theorem 3. *Consider the general binomial model with* $0 < d < 1 + r < u$. *Let the* *risk-neutral probabilities be given by* $\tilde{p} = \frac{1+r-d}{u-d}$, $\tilde{q} = \frac{u-1-r}{u-d}$. *Then, under the risk-neutral probability measure, the discounted stock price is a martingale.*

Proof. The goal of this proof is to show that the martingale equation $\tilde{E}\left[\frac{S_{n+1}}{(1+r)^{n+1}}|\mathcal{F}_n\right] = \frac{S_n}{(1+r)^n}$ holds regardless of the value of n. First, let n be an arbitrary natural number. Then,

$$\tilde{E}\left[\frac{S_{n+1}}{(1+r)^{n+1}}|\mathcal{F}_n\right] = \tilde{E}\left[\frac{S_{n+1}}{(1+r)^{n+1}} \cdot \frac{S_n}{S_n}|\mathcal{F}_n\right],$$

which is found by multiplying the conditional expectation by the fraction $\frac{S_n}{S_n}$. Then, we rearrange the terms to arrive at

$$\tilde{E}\left[\frac{S_{n+1}}{(1+r)^{n+1}}|\mathcal{F}_n\right] = \tilde{E}\left[\frac{S_n}{(1+r)^n(1+r)} \times \frac{S_{n+1}}{S_n}|\mathcal{F}_n\right]. \tag{4.2}$$

Next, we "take out what is known." In (4.2), we are conditioning on the coin tosses up to time n, so at this time, we would already know the value of the stock at time n (S_n), and the value of $\frac{1}{(1+r)^n}$ as well as the value of the constant $\frac{1}{1+r}$. This allows us to remove them from the conditional expectation to produce

$$\tilde{E}\left[\frac{S_{n+1}}{(1+r)^{n+1}}|\mathcal{F}_n\right] = \frac{S_n}{(1+r)^n} \times \frac{1}{1+r}\tilde{E}\left[\frac{S_{n+1}}{S_n}\right].$$

By examining this equation, we should notice that the random variable $\frac{S_{n+1}}{S_n}$ only depends on toss number $n+1$. The reason for this is that S_{n+1} either equals uS_n with probability $\tilde{p}$ or dS_n with probability $\tilde{q}$. Since we already know the value of S_n, we only need to know the outcome of the next coin toss so that we know whether the stock went up or down. Another way to see this is that the quantity $\frac{S_{n+1}}{S_n}$ will either equal $\frac{uS_n}{S_n}$ or $\frac{dS_n}{S_n}$, which allows us to cancel the stock price at time n. Thus, taking the expectation will yield

$$\tilde{E}\left[\frac{S_{n+1}}{(1+r)^{n+1}}|\mathcal{F}_n\right] = \frac{S_n}{(1+r)^n} \times \frac{\tilde{p}u + \tilde{q}d}{1+r}.$$

Finally, since $\tilde{p}u + \tilde{q}d = 1 + r$, this term equals one, leaving us with

$$\tilde{E}\left[\frac{S_{n+1}}{(1+r)^{n+1}}|\mathcal{F}_n\right] = \frac{S_n}{(1+r)^n}.$$

Therefore, under the risk-neutral probability measure, the discounted stock price is a martingale. □

Earlier, it was mentioned that the discounted stock price is not necessarily a martingale under the actual probability measure. Conceptually, we can see why this is true. We know that martingales have no tendency to rise or fall. However, in reality, stock prices have a tendency to rise and often rise faster than investments in the money market in order to compensate investors for the risk associated with the stock market. For example, suppose the actual proabilities are $p = \frac{2}{3}$ and $q = \frac{1}{3}$, $n = 2$, and we know which events in $\mathcal{F}_2$ have occurred, namely, $\{HHH, HHT\}$. Now, we examine the martingale property of the discounted stock price, which says: $\left(\frac{1}{1+r}\right)^{n+1} E[S_{n+1}|\mathcal{F}_n] = \left(\frac{1}{1+r}\right)^n S_n$. First, we will look at the right-hand side of the equation:

$$\left(\frac{1}{1+r}\right)^n S_n = \left(\frac{4}{5}\right)^2 S_2 = \left(\frac{16}{25}\right) \times 16 = \$10.24.$$

Next, we will look at the left-hand side of the equation:

$$\begin{aligned}\left(\frac{1}{1+r}\right)^{n+1} E[S_{n+1}|\mathcal{F}_n] &= \left(\frac{4}{5}\right)^3 E\left[S_3|\mathcal{F}_2\right] \\ &= \left(\frac{64}{125}\right)\left[32 \times \frac{2}{3} + 8 \times \frac{1}{3}\right] \\ &= \left(\frac{64}{125}\right) \times 24 = \$12.29.\end{aligned}$$

Thus, $\left(\frac{1}{1+r}\right)^{n+1} E[S_{n+1}|\mathcal{F}_n] \geq \left(\frac{1}{1+r}\right)^{n} S_n$, so under the actual probability measure, the discounted stock price is not a martingale.

In fact, there is a special name for sequences of random variables where the equality does not hold, but rather one of the inequalities applies, instead. Under the actual probability measure, the discounted stock price has a tendency to increase over time. Mathematically, if for all values of n,

$$M_n \leq E\left[M_{n+1}|\mathcal{F}_n\right],$$

we say the process is a **submartingale**. On the other hand, if the process has a tendency to decrease over time, that is if for all values of n,

$$M_n \geq E\left[M_{n+1}|\mathcal{F}_n\right],$$

we say the process is a **supermartingale**.

4.2 Discrete-Time Stochastic Integral

The final concept we will discuss in this chapter is a **discrete-time stochastic integral**, which is sometimes called a martingale transform.

Suppose $M_0, M_1, ...M_N$ is a martingale, and let $\Delta_0, \Delta_1, ...\Delta_{N-1}$ be adapted to $\mathcal{F}_n$ with $E[(\Delta_n)^2] < \infty$ for all values of n. We define the discrete-time stochastic integral to be $I_0 = 0$ and:

$$I_n = \sum_{j=0}^{n-1} \Delta_j (M_{j+1} - M_j), n = 1, ..., N. \tag{4.3}$$

An interesting fact is that $I_0, I_1, ..., I_N$ is still a martingale, which we will prove in the next theorem. At first glance, it may be difficult to see the application of this to the finance theory that has been discussed. However, each I_n can be thought of as our monetary position in the stock at time n, where M_j is the discounted stock price and Δ_j is the amount of shares we own, each at time j. Notice that in the definition of the discrete stochastic integral we required that the M_n sequence be a martingale. Recall that this is indeed satisfied by the discounted stock price (as proven earlier in Theorem 3), so even through this intepretation, I_n is a discrete stochastic integral. We will continue to use $\mathcal{F}_n$ as the σ-algebra representing all of the information up to time n.

Theorem 4. *The discrete time stochastic integral defined by (4.3) is a martingale.*

Proof. The goal of this proof is to show that $E[I_{n+1}|\mathcal{F}_n] = I_n$ by using the definition of martingale. Assume $M_0, M_1, ..., M_N$ is a martingale, and let $\Delta_0, \Delta_1, ...\Delta_{N-1}$ be an adapted proces with $E[(\Delta_n)^2] < \infty$ for all values of n. First, we check that the first two properties hold, meaning that I_n must be adapted to $\mathcal{F}_n$ and $E[F_n^2] < \infty$. Since M_n is a martingale, it is adapted to $\mathcal{F}_n$. Further, Δ_n is also adapted to $\mathcal{F}_n$ by assumption. Thus, I_n is a sum of products of $\mathcal{F}_n$ measurable terms, so I_n is also adapted. We can easily check that $E[I_n^2] < \infty$ by using the **Cauchy-Schwarz Inequality**, which can be found in Durrett [6]. One result of this inequality is stated in the theorem below.

Theorem 5. *If X and Y are random variables with $E[X^2] < \infty$ and $E[Y^2] < \infty$, then*

$$|E[XY]|^2 \leq (E[X^2])(E[Y^2]).$$

Using this inequality, since M_n and Δ_n are both square integrable for all values of n, then I_n is also square integrable for all values of n. In other words, $E[I_n^2] < \infty$, so the first two parts of the definition of martingale are satisfied. Finally, to check part (3), let n be an arbitrary natural number. Then, for $0 \leq n \leq N-1$,

$$I_{n+1} = \sum_{j=0}^{n} \Delta_j (M_{j+1} - M_j) = I_n + \Delta_n (M_{n+1} - M_n). \tag{4.4}$$

In this step, we are simplyfing the equation for I_{n+1}. Since each random variable in the sequence is a summation of the previous random variables, the next random variable in the sequence, I_{n+1} is found by simply adding the very next term in the sum to the previous I_n. The next term in the sum is given by $\Delta_n(M_{n+1} - M_n)$. This gives us (4.4). Next, we take the conditional expectaion of both sides and apply the distributive property, producing

$$\begin{aligned} E[I_{n+1}|\mathcal{F}_n] &= E[I_n + \Delta_n(M_{n+1} - M_n)|\mathcal{F}_n] \\ &= E[I_n + \Delta_n M_{n+1} - \Delta_n M_n|\mathcal{F}_n]. \end{aligned}$$

Now, we are able to use linearity of expectations to separate the terms of the conditional expectation, yielding

$$E[I_{n+1}|\mathcal{F}_n] = E[I_n|\mathcal{F}_n] + E[\Delta_n M_{n+1}|\mathcal{F}_n] - E[\Delta_n M_n|\mathcal{F}_n].$$

In order to arrive at the next step, (4.5), we must realize that if we are given the first n coin tosses, we already know the result of I_n, Δ_n, and M_n. In other words, we know our position in the stock at time n because we know how many shares we purchased as well as the price we paid for each share. Thus,

$$E[I_{n+1}|\mathcal{F}_n] = I_n + \Delta_n E[M_{n+1}|\mathcal{F}_n] - \Delta_n M_n. \tag{4.5}$$

Lastly, we use the definition of martingale. Since $M_0, M_1, ..., M_N$ is a martingale, $E[M_{n+1}|\mathcal{F}_n] = M_n$, which yields

$$\begin{aligned} E[I_{n+1}|\mathcal{F}_n] &= I_n + \Delta_n M_n - \Delta_n M_n \\ &= I_n. \end{aligned}$$

The final step arises naturally, and the proof is complete. Therefore, by definition of martingale, $I_0, I_1, ..., I_N$ is a martingale. □

One final interesting fact is the relation of this discrete-time stochastic integral to the concept of Riemann sums, which are discussed in intro calculus classes. Riemann sums are used as a preview to integrals, and the same can be said for stochastic integrals. The idea of the Riemann integral is to multiply the length of an interval by the function value at a point inside the interval.

Traditionally, the point can be chosen to be any point inside the subinterval, such as the left or right-hand end points or the midpoints of the intervals, and in intro calculus, all three can be used to approximate the integral.

However, the idea changes in stochastic calculus because it is most common to use the left-hand end point of the interval. The reason for this is because stochastic calculus is time-oriented, so when making the switch to random variables, it is necessary for them be adapted in order for the application to make sense. For example, in the discrete stochastic integral defined by (4.3), each Δ_j is a random variable representing the number of shares we own at time j. In reference to the Riemann integral, each Δ_j represents the point chosen inside an interval at which we evaluate the function. Mathematically, it is adapted because we only have information up to time j. In other words, Δ_j is $\mathcal{F}_j$-measurable. Thus, this is the left-hand end point of the interval $M_{j+1} - M_j$. This type of stochastic integral, which is also a martingale, is called an *Ito integral*.

On the other hand, we can also consider the stochasic integral defined by

$$\tilde{I}_n = \sum_{j=0}^{n-1} \Delta_{j+1}(M_{j+1} - M_j), n = 1, ..., N.$$

Continuing with the Riemann integral comparison, $\tilde{I}_n$ corresponds to the right-hand end point being chosen, Δ_{j+1}. However, the drawback to this definition is that the stochastic integral would no longer be a martingale. The intuitive reason that the left-hand end point should be used is that $\tilde{I}_n$ implies a sort of "psychic knowledge," that is, using future unknown information in the trading strategy. In this case, we would have future knowledge of how many shares of the stock we should purchase in the next time period by using Δ_{j+1} in the calculation. This is prohibted by the binomial model because we do not want to consider situations where investors have supernatural abilities.

Chapter 5

Applications

5.1 Theoretical Price vs. Real-World Price

The most natural application of the binomial model is the pricing of options in a real-world setting. More specifically, this paper compares the theoretical price to the actual price of two companies: Netflix and Johnson and Johnson. These two companies were chosen because of their respective stabilities. Netflix is generally regarded as a potentially risky investment, whereas Johnson and Johnson is a relatively stable choice. This section explores how well the binomial model performs in the pricing of options for these two companies. The process of pricing an option can be described by the following steps:

1. Analyze the price history of the company in order to calculate the up and down factors.

2. Create the binomial price tree, beginning with the current price and

using the up and down factors to compute future prices.

3. Calculate the value of the option for each of the possible ending prices in the price tree.

4. Use Equation (3.4) to determine the value of the option at each prior period, eventually ending with the time zero value (price) of the option.

First, we will analyze pricing a Netflix call option that will expire in January of 2015 with a strike price of $80.00. Remember that this gives us the right but not the obligation to purchase one share of Netflix stock for $80 in January 2015, regardless of its actual price. We assume that our current period is January of 2014. We are only concerned with the month (not the day) because we are choosing months as the time periods. This option will expire after 12 periods, so the multi-period binomial model should be used. Step 1 of the process is to calculate the up and down factors. Intuitively, the magnitude of this movement should depend on the size of the time interval as well as the volatility of the stock price. Cox-Ross-Rubinstein [4] determined an effective way to capture this information is to use the following formulas:

$$u = e^{\sigma \sqrt{t}}, \tag{5.1}$$

$$d = e^{-\sigma \sqrt{t}} = \frac{1}{u}, \tag{5.2}$$

where t is the size of the time period and σ represents the volatility of the stock price. **Volatility** is a variable showing the extent to which the return of the underlying asset will fluctuate between now and the option's expiration. In other words, it is a measure of the amount of uncertainty (risk) associated with the size of changes in the stock price. This poses another question: how should we calculate the volatility of the stock price? The following steps were utilized in the calculation of volatility:

1. List the historical *monthly* stock prices.

2. Calculate the percentage change from each period (month) to the next.

3. Compute the standard deviation of the percentage price changes.

4. Multiply the value from step 3 by the square root of the number of periods per year (in this case, $\sqrt{12}$).

In the end, this process provides an annual percentage volatility for the given company's stock. The value found can then be substituted for σ in Equations (5.1) and (5.2). Since Netflix was chosen as a riskier stock, we expect it to have a higher volatility. The volatility for Netflix was 0.7859, so the up and down factors were 1.255 and 0.7970, respectively.

Now that we have calculated the up and down factors, the next step is to create the price tree. The current price of Netflix stock (January 2014) is $S_0 = \$366.81$. In order to simplify the number of computations in the creation of the price tree, we should notice that an upward movement followed by a downward movement is the same as the stock going down then up. Thus, in order to calculate the stock price at time n, we use the formula:

$$S_n = (S_0)(u)^L(d)^{n-L},$$

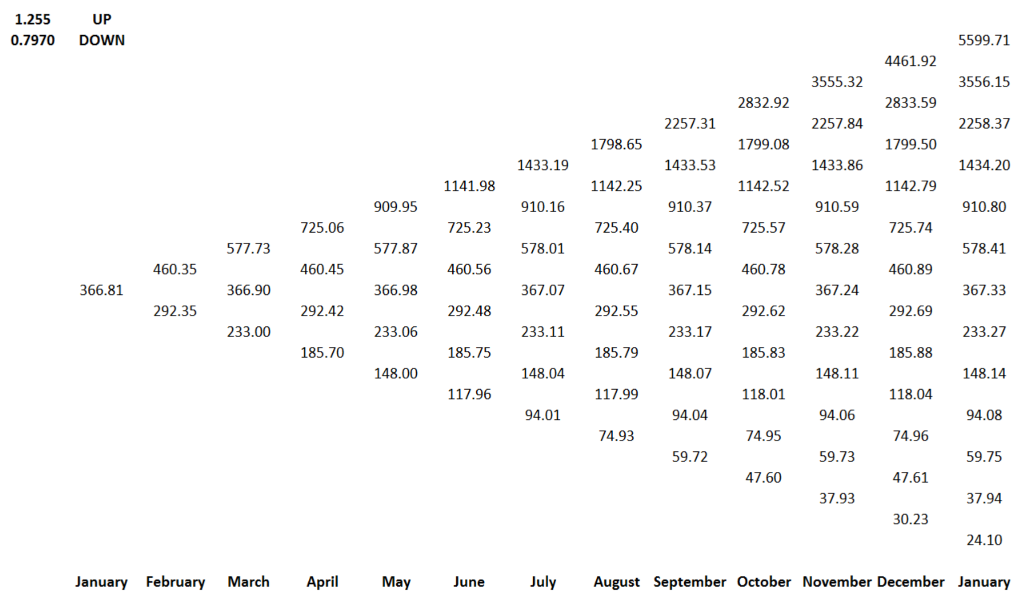

Figure 5.1: This is the price tree for Netflix beginning in January 2014 and ending in January 2015.

where L is the number of upward movements. The number of upward movements could also be seen as a binomial random variable with probability of success $\tilde{p}$ and 12 trials: $L \sim \text{Binomial}(12, \tilde{p})$. The price tree for Netflix is shown in Figure 5.1.

Recall that the next step in the process is to calculate the value of the option for each of the possible ending prices in the price tree. In other words, how much money would we make by exercising the option if the stock price was a given amount. For example, notice the top right node in Figure 5.1, which corresponds to the stock price increasing every time period for one year. According to the model, if this occurred, the stock price would be \$5599.71. With this in mind, what is the value of the option with an \$80 strike price? Without the option, we would have to pay the full \$5599.71 if we wanted one share of Netflix stock. Instead, we have the right to buy that same share for \$80.00. Thus, the value of the option is 5599.71 − 80.00 = \$5519.71. In step 3, we continue this calculation for every end node (every possible stock price in January 2015).

However, let us try this computation on the bottom right node in Figure 5.1, which corresponds to the stock price decreasing every time period for one year. According to the model, if this occurred, Netflix stock would trade for a mere \$24.10, a substantial decrease from the current price of \$366.81. In this case, the stock price is less than the strike price of \$80, so the option is worthless, giving it a value of \$0.00. Thus, we would not exercise the option.

In summary, the value of the option at the exercise date (January 2015 or period 12) can be computed using the following formula:

$$V_{12} = (S_{12} - K)^+,$$

where K is the strike price. The notation $(\cdots)^+$ simply means we take the maximum of the quantity inside the parentheses and zero, allowing us to include the situations when we do not exercise the option. In other words, this notation is another way to write:

$$V_{12} = \max\{S_{12} - K, 0\}.$$

However, remember that V_{12} and S_{12} are random variables that depend on the outcome of all 12 coin tosses. Typically, there are 2^{12} possible outcomes of 12 coin flips. However, as we mentioned earlier, since the order of the coin tosses does not matter ($HT = TH$), the number of outcomes drops dramatically to 13.

The 13 possible values of the Netflix call option at time 12 are given in Figure 5.1.

Table 5.1: Time 12 Option Values

# of Up's	Time 12 Option Values
12	5519.71
11	3476.15
10	2178.37
9	1354.20
8	830.80
7	498.41
6	287.33
5	153.27
4	68.14
3	14.08
2	0.00
1	0.00
0	0.00

The fourth and final step of the process involves going backward in time, starting with the value of the option on the exercise date and ending with the time zero price of the option. To do this, we use Equation (3.4), which requires a few more parameters that we have not yet computed: the risk-neutral probabilities $\tilde{p}$ and $\tilde{q}$, and the risk-free interest rate r. The standard choice for the risk-free interest rate is the interest rate on U.S. Treasury Bills, giving us a choice of the 3-month, 6-month, or one year rate. Since the call option we are analyzing expires in one year, the one-year rate is used, which, in January 2014, was only 0.13%. In Section 3.3, we used Equation (3.3) to solve for formulas for $\tilde{p}$ and $\tilde{q}$. These are repeated below, making the necessary substitutions for Netflix stock,

$$\begin{aligned}\tilde{p} &= \frac{1+r-d}{u-d} \qquad (5.3)\\ &= \frac{1+.0013-.7970}{1.255-.7970} = .4461,\\ \tilde{q} &= 1-\tilde{p} = .5539. \qquad (5.4)\end{aligned}$$

Now, we are ready to use Equation (3.3) to back-track. The equation is shown again below, followed by an example of how we go backward in time to compute the value of the option at time 11. Recall that Equation (3.3) is

$$V_n = \frac{1}{1+r}[\tilde{p}V_{n+1}(H) + \tilde{q}V_{n+1}(T)].$$

For this example, we will assume the stock price has increased all 11 time periods, meaning that, at time 12, we will end at one of the top 2 values in Figure 5.1. Using the equation allows us to make the following computations:

$$\begin{aligned}V_{11} &= \frac{1}{1+r}[\tilde{p}V_{12}(H) + \tilde{q}V_{12}(T)]\\ &= \frac{1}{1.0013}[(.4461)(5519.71) + (.5539)(3476.15)]\\ &= \$4393.43.\end{aligned}$$

The above example assumed we had 11 upward movements. Next, we would assume 10 upward movements and perform the same calculation. This process would continue to complete all 12 possible time 11 values. Then, we would move back to time 10, 9, 8,..., until we reach time 0. After completing this process, the model determined that the theoretical price in January 2014 of a Netflix call option with a strike price of $80.00 expiring in January 2015 is $298.30. The real-world price of this option is $286.84. Thus, the percent error of the model was approximately 4%.

The second option we will analyze is a call option for Johnson and Johnson expiring in January of 2015 with a strike price of $80.00. This stock was chosen because it is considered to be a relatively stable stock. However, when using the same process as before, its volatility is 0.5677, giving up and down factors of 1.089 and 0.9186, respectively. Even though this is a lower volatility than that of Netflix, indicating a more stable stock in comparison, it is still fairly high. One reason for this may be because volatility is typically calculated in terms of *daily* historical prices. In our model, recall that we chose *months* as the time period to be used, so volatility was also calculated in terms of monthly historical prices. This may have skewed the volatility of Johnson and Johnson because we are only using the opening price at the beginning of each month.

In order to price the Johnson and Johnson call option, we assume that our current period is February 2014, so we use a current price of $86.78. Using the same process as was used to price the Netflix call option, our model's theoretical price for the Johnson and Johnson call option is $13.98. However, the real-world price of this option is $9.10. Thus, the percent error of the model was approximately 53%, a very high percent error in comparison to our model's price of the Netflix call option. Why might this have happened? Our first guess may be that the volatility used in our model is higher than the volatility used in the real-world price calculation. However, since the option is more expensive according to our model, that says that the buyer of the option is more likely to make money by purchasing the option (causing higher demand for that option and a rise in the price). In order for the buyer of the option to make money, the stock price of Johnson and Johnson must be higher than $80.00 in January 2015, allowing him or her to exercise the option (buying shares of Johnson and Johnson for the lower price). Since the volatility of our model seems to be higher than the real-world volatility, this may account for some of the error. However, there may be other factors at work in the real-world model.

The key fact might be that Johnson and Johnson is a dividend paying stock, which is not factored into our model. When a dividend is declared, it gets discounted in the stock price, decreasing the price of the stock. As we mentioned before, the buyer of the option will be less likely to exercise the option at lower prices (because eventually it will drop below the strike price of $80.00). According to this logic, incorporating dividends into the model would cause a cheaper option price because the stock price will be lower. This may explain the high percent error of our model in comparison to the real-world model — the real-world model includes dividends in the calculation of the option.

One final possibility for the higher percent error arises from analyzing exactly how percent error is calculated, specifically, the denominators. The denominator for the percent error in this context is the real-world price of the option. The Netflix option price is much higher than the Johnson and Johnson option price ($286.84 > $9.10). In other words, by using a strike price of $80 for both companies, the Netflix option is much more valuable because the current stock price of Netflix is so much higher than the strike price. This is another factor that may inflate the percent error of Johnson and Johnson when compared to Netflix.

5.2 Which NBA Players were Worth Their Rookie Salaries

The 2003 NBA draft was considered to be one of the most hyped-up drafts in the history of the NBA. Many analysts believed that there were numerous players in the draft class that could change the face of the NBA, which was slowly running out of stars since the retirement of Michael Jordan. On draft day, four of the top five picks were used to draft LeBron James, Carmelo Anthony, Chris Bosh, and Dwayne Wade. These four players have combined for more than 30 All-Star appearances and nearly 70,000 points. James, Bosh, and Wade have now all joined the Miami Heat and have won two consecutive NBA titles (and this is only the top 5 picks). There is little question that the hype surrounding the 2003 NBA draft has been validated. However, how soon could we have known that these players were worth the hype? This section attempts to use the binomial asset pricing model to answer the question: which NBA players from the 2003 NBA draft were worth their rookie salaries? This idea was inspired by a similar study performed on cricketers in Saikia, Bhattacharjee, and Bhattacharjee [9].

In order to use the binomial model to answer this question, we need to establish meaning to each of its parameters. In a way, we will be thinking of a player's performance as the stock price. After each season (each time period), we will assume he either improves or worsens, just as the stock price can either increase or decrease. The NBA structures all first round salaries the same way. The player receives a set salary (decreases steadily throughout the draft) depending on the pick he is drafted. Also, in 2003, all first round picks were signed to three-year contracts. Thus, our model will contain three time periods.

In the binomial model, we use a risk-free interest rate and a risk-neutral probability measure, which we will again incorporate into this application. One way to use the risk-free interest rate in any situation is to compute the value of an investment in the next period. We will make a similar calculation here, assuming that a person's wealth can either go up or down by that interest rate (depending on lending or borrowing). Since we are analyzing salaries from the 2003 NBA draft, it is appropriate to use the one year interest rate on U.S. Treasury Bills in 2003, which was 1.07% at the time the draft took place. Thus, a $1 at time zero (the time of the draft) could be compounded to

$$M_u = \$1 \times (1 + .0107) = 1.0107$$

at time one, or reduced to

$$M_d = \frac{\$1}{1 + .0107} = 0.9894.$$

This enables us to calculate a risk-neutral probability measure using a formula found in [9] . Using the factors M_u and M_d, we arrive at the equation

$$
\begin{aligned}
r &= M_u \times \tilde{p} + (1 - \tilde{p}) \times (-M_d) \qquad (5.5)\\
&= 1.0107\tilde{p} + (-0.9894)(1 - \tilde{p}),\\
\tilde{p} &= \frac{r + 0.9894}{1.0107 + 0.9894}\\
&= \frac{1.0001}{2.0001} \approx 0.500,\\
\tilde{q} &\approx 0.499.
\end{aligned}
$$

The next step is to decide how we will measure a player's performance. **Win Shares** (WS) is an advanced statistic that is commonly used to measure performance. Intuitively, it is a metric that estimates the number of wins a player produces for his team. This can be calculated on its own, or by finding a player's Offensive Win Shares and Defensive Win Shares, and adding those together. The idea was first developed for baseball by Bill James, but was eventually converted to basketball. However, the formula for calculating Win Shares is quite complicated and will not be listed here. One thing to note about Win Shares is that this statistic can be negative. Negative Win Shares simply means that the player actually *hurt* the teams chances of winning. This happened to one player in the top ten picks, Darko Milicic. Darko totaled −0.20 Win Shares in each of his first two seasons. Table 5.2 gives the total win shares for the top ten 2003 draft picks in their first three years. The data was taken from Forman [7].

The official performance measure we use for a given player is found by

Table 5.2: Total Win Shares

Pick	Player	Total WS	WS/Season
1	LeBron James	35.7	11.90
2	Darko Milicic	0.8	0.27
3	Carmelo Anthony	20.4	6.80
4	Chris Bosh	22.8	7.60
5	Dwayne Wade	30.0	10.0
6	Chris Kaman	8.2	2.73
7	Kirk Hinrich	18.9	6.30
8	T.J. Ford	7.5	2.50
9	Michael Sweetney	7.3	2.43
10	Jarvis Hayes	2.9	0.97

$$p_t = \frac{WS_t - \min\{WS_i\}}{\max\{WS_i\} - \min\{WS_i\}},$$

where t represents the season (time period) and WS_i represents a sequence of numbers correponding to the Win Shares of each player in the NBA during season t. Thus, p_t will be normalized between 0 and 1, with the player having the highest number of Win Shares receiving a performance score of 1, and the player having the least Win Shares receiving a performance score of 0. Each season, we compute new performance scores for each player in the study. The players in this study will consist of the top ten picks of the 2003 NBA draft.

Next, we must decide on the up and down factors. When using the model to compute the price of options, it made sense to include certain variables like volatility in the computation of the up and down factors. However, the same variables no longer apply in the context of NBA salaries. The up and down factors determine the magnitude that the stock price will increase or decrease. In this study of player salaries, it makes sense to utilize the player's performance from the current season to calculate how much the player can improve or worsen. Thus, we use the following factors:

$$u_t = 1 + p_t,$$
$$d_t = p_t.$$

The choice of these factors is supported by the following argument. A "better" player will work to continue to improve more than an average player because he has already seen some of the benefits of being an elite player. Also, if a player with a higher performance score does worsen throughout the offseason, his value should not decrease quickly because he was playing at a better than average level. These factors are able to capture the essence of this argument because better players will have higher performance measures. One thing to note is that unlike the up and down factors used in Section 5.1, these factors change after each season in order to account for a rapid improvement or declination of a player's performance. Another difference between the two applications is the stock application is a predictive model, whereas this application uses the binomial model as an evaluative tool.

The only model input remaining is the initial value of the player (initial stock price). Since we are attempting to answer the question of whether or not a particular player was worth his rookie salary, we use the rookie salary as the initial player value. Then, the steps of the option pricing process follow accordingly, eventually using the risk-neutral probabilities calculated earlier to back-track to a theoretical original valuation. The results of the first ten 2003 draft picks are summarized in Table 5.3.

Table 5.3: Player Valuation Summary

Pick	Player	Rookie Salary	Model Valuation
1	LeBron James	12.96	19.96
2	Darko Milicic	11.59	1.97
3	Carmelo Anthony	10.51	7.58
4	Chris Bosh	9.34	7.88
5	Dwayne Wade	8.50	10.08
6	Chris Kaman	7.72	2.30
7	Kirk Hinrich	7.09	4.76
8	T.J. Ford	6.46	1.90
9	Michael Sweetney	5.94	1.76
10	Jarvis Hayes	5.64	1.14

The results shown in Table 5.3 should be read in a relative sense. In other words, even though LeBron's model valuation is 19.96, this does not necessarily mean that he should have received $19.96 million. Instead, it shows that he out-performed the other nine draft picks in their first three seasons. On the other hand, Darko's plunge from 11.59 to 1.97 signifies that he was not worth a number two pick. One final point is that it is difficult to gain much insight into the value of Jarvis Hayes because he was drafted tenth and remained tenth. We would be able to learn more if we extended the study to include more picks such as the entire first round. The new order of players as well as their change in draft position is given by Table 5.4.

Table 5.4: New Draft Order

Model Pick	Real Pick	Player	Difference
1	1	LeBron James	-
2	5	Dwayne Wade	+3
3	4	Chris Bosh	+1
4	3	Carmelo Anthony	-1
5	7	Kirk Hinrich	+2
6	6	Chris Kaman	-
7	2	Darko Milicic	-5
8	8	T.J. Ford	-
9	9	Michael Sweetney	-
10	10	Jarvis Hayes	-

5.3 Discussion

Throughout this I.S., we have explored the binomial asset pricing model, and we end with a short discussion section. We have seen that the binomial model is able to estimate the price of an option in discrete time, and the Black-Scholes model is able to give the price of an option in continuous time. Are these models equal or are there certain pros and cons of using one in comparison with the other?

The first and obvious advantage to using the binomial model is that there is no calculus (or any advanced math) associated with the pricing process. On the other hand, at the very minimum, the Black-Scholes model requires Brownian motion, which can be a difficult concept on its own. The second pro of the binomial model is that it can account for dividends with a few modifications. For example, informally, consider a situation where a stock pays a dividend as a percentage of its price on a given date. The binomial model can account for this in the price tree by multiplying the stock price at that date by the percent of the price that the stockholder will not receive. It is possible that this type of calculation could reduce the percent error of Johnson and Johnson in Section 5.2. This would be an area that this I.S. could be extended in the future. In addition to accounting for dividends, the binomial model can also be used to price **American options**. An American option differs from a European option because rather than only having the ability to exercise the option on a specific date (European option), an American option allows its holder to exercise at *any* date prior to the expiriation date. Another possible extension of this I.S. would be to analyze the differences between European and American options.

There are also a few cons to using the binomial model. Most importantly, it does not give *exact* answers. As mentioned time and time again, the binomial model is a discrete time model, and time in real life runs continuously. Essentially, the binomial model estimates the option price of the Black-Scholes model, and the estimation depends on the size of the time intervals. Smaller intervals give better estimations because the time intervals in continuous time can be thought of as of size 0. In other words, as the size of the time intervals approaches 0, the option price given by the binomial model approaches the Black-Scholes price. A second con of the binomial model is the fact that it is not realistic to calculate by hand. As the size of the time intervals decreases, more and more computations are needed in order to back-track from the final period to the starting period.

Now, we discuss a limitation of the application of Section 5.2. Equation (5.5) was used to calculate the risk-neutral probability measure used to back-track to a player's valuation. However, in the binomial model, the risk-neutral probabilities were given by Equations (5.3) and (5.4), which depend on the up and down factors. The equation used to calculate the risk-neutral probabilities for the player valuation application does not use the up and down factors. Instead, it uses a combination of the risk-free interest rate and the potential values of \$1. It is not clear why this equation is appropriate, so an attempt was made to use Equation (5.3) to calculate $\tilde{p}$ in Section 5.2. However, the equation reduced to

$$\tilde{p} = 1 + r - p \tag{5.6}$$

because the down factor in the valuation model is the performance measure p. Thus, players with a very high performance measure have a very low probability of improving the next season. For example, a player with a performance measure of 0.90 has a $1 + r - 0.90$ probability of improving the next season, and since r will always be small relative to a high p (because it is an interest rate), $\tilde{p}$ will also end up being small.

The results of using Equation (5.6) in the valuation model did not make sense, intuitively, and were discarded. A reason that this may not be an optimal way to calculate the risk-neutral probability measure in a player valuation context may have to do with the derivation of Equation (5.6). It comes from Equation (5.3), which originated from Equation (3.3):

$$V_n = \frac{1}{1+r}[\tilde{p}V_{n+1}(H) + \tilde{q}V_{n+1}(T)].$$

Recall that in the context of stocks, this equation states that the average rate of growth in the money market should equal the average rate of growth in the stock market. However, if we convert this to the player valuation context, it says that the average rate of growth of the money market should equal the average rate of growth of the player. Intuitively, this does not make sense because teams do not have the option to invest in the money market instead of paying a player's salary. For future research, there may be a better way to calculate a risk-neutral probability measure that better evaluates player salaries.

This I.S. introduced the binomial asset pricing model and explored a few of its mathematical properties. Also, it demonstrated that it can be used for more creative applications than simply pricing options. However, there are plenty of concepts that we did not explore. The world of financial mathematics is ever-changing and improving, so let this serve as proof that the creativity of mathematics is unlimited.

Bibliography

[1] Richard F. Bass. The basics of financial mathematics. Department of Mathematics, University of Connecticut, 2003.

[2] Phelim Boyle and Feidhlim Boyle. *Derivatives: The tools that changed finance*. Risk Books, 2001.

[3] Goddard Consulting. Binomial-one-step. Digital image. http://www.goddardconsulting.ca/option-pricing-binomial-index.html.

[4] John C Cox, Stephen A Ross, and Mark Rubinstein. Option pricing: A simplified approach. *Journal of financial Economics*, 7(3):229–263, 1979.

[5] Tuan Do-Hong. Random variable. Digital image. http://cnx.org/content/m26841/latest/.

[6] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, third edition, 2010.

[7] Sean Forman. Basketball reference, January 2014. http://www.basketball-reference.com/.

[8] Ralf Korn and Stefanie Muller. Binomial trees in option pricing—history, practical applications and recent developments. 2010.

[9] Hemanta Saikia, Dibyojyoti Bhattacharjee, and Atanu Bhattacharjee. Performance based market valuation of cricketers in ipl. *Sport, Business and Management: An International Journal*, 3(2):127–146, 2013.

[10] Steven E. Shreve. *Stochastic calculus for finance. I*. Springer Finance. Springer-Verlag, New York, 2004. The binomial asset pricing model.

[11] R. J. Williams. *Introduction to the mathematics of finance*, volume 72 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2006.