Spring 5-1-2017

Mathematical Modeling of Financial Derivative Pricing

Kelly L. Cosgrove
University of Connecticut, cosgrove.kelly@gmail.com

Mathematical Modeling of Financial Derivative Pricing

Kelly Cosgrove

May 6, 2017

Abstract

The binomial asset-pricing model is used to price financial derivative securities. This text will begin by going over the probability concepts necessary to understand this discrete-time model. It then develops the theory behind the binomial model and different properties that arise. It shows how to use the binomial model to predict future stock prices, and then uses this information to price derivative securities. It initially focuses on the European call option, but goes on to provide a pricing method for the American put option. However, many of the theorems developed are applicable to all derivative securities. The text wraps up by considering a different method used in pricing derivative securities, the Black-Scholes model, which is based on continuous-time concepts.

Contents

Introduction

The binomial asset-pricing model is used to price financial derivative securities. In this text, we will mostly use the example of the European call option to illustrate the function the binomial model serves. This type of derivative is one that allows its owner the right (but not the obligation) to buy a stock at a specified strike price on a specified expiration date. A similar derivative is the European put option which gives the owner the right to sell stock at a specific price on a specific date.

We begin by reviewing general probability concepts needed to develop the binomial asset-pricing model in Chapter 1. Topics include finite probability spaces, random variables, and properties of conditional expectations. Chapter 1 then goes on to develop the properties of certain adapted stochastic processes. Namely, we shall look at martingales and Markov processes.

We develop the idea of the binomial asset-pricing model and how to use it to price European derivative securities in Chapter 2. In order to accomplish this, we will need to figure out how to replicate the derivative security in the stock and money markets. Once we are able to do this, we can determine a fair price for the derivative security by setting its value at time n equal to that of our replicated portfolio.

In Chapter 3 we will explore American derivative securities which are defined similarly to their European counterparts except for the fact that its owner has the right to exercise the option at any point up to or on the expiration date. This complicates the process of building a replicating portfolio because there will exist an optimal exercise date on which the owner of the derivative security should exercise it (which we must find). The concept of stopping times will be developed in this chapter and will be crucial to figuring out the optimal exercise date of an American derivative security.

Chapter 4 wraps up the paper by introducing an alternate method for pricing derivative securities known as the Black-Scholes model. It begins by describing the random walk, whose continuous-time counterpart, Brownian motion, is an underlying assumption of the Black-Scholes model. It then provides the general idea behind Black-Scholes and how it is related to the binomial model.

Several sources were referenced in order to complete this paper, but Steven E. Shreve's *Stochastic Calculus for Finance I The Binomial Asset Pricing Model* [Shreve, 2004] was the most prominently referenced, as the source of the material in Chapters 1, 2, and 3, and Section 4.1. Further, Rossitsa Yalamova's *Simple heuristic approach to introduction of the Black-Scholes model* [Yalamova, 2010] was referenced to complete Section 4.2, and Noga Alon and Joel H. Spencer's *The Probabilistic Method* [Alon and Spencer, 2015] was referenced for Theorem 4.2.1 (Azuma's Inequality).

Chapter 1

Probability Theory

1.1 Finite Probability Spaces

Because the binomial asset-pricing model is an application of probability theory, it is crucial to have a good understanding of general probability concepts in order to understand the model itself. In this section we will build up our understanding of a finite probability space. We will use the example of a coin toss to illustrate the components of a finite probability space.

Say we toss a coin 3 times. Let H denote a heads flip and T denote a tails flip. The *sample space*, Ω, is the finite set of all possible outcomes that could result from tossing that coin, or

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\} \tag{1.1}$$

Sequences of outcomes are denoted by $\omega = \omega_1\omega_2\omega_3$.

We call subsets of Ω *events*. For example, the event A that we get at least two tails:

$$A = \{HTT, THT, TTH, TTT\}$$

If we let $\#\mathrm{T}(\omega_1\omega_2\omega_3)$ be defined as the number of tails that appear in our sequence of 3 coin tosses, then we can rewrite A as:

$$A = \{\omega \in \Omega; \#T(\omega_1\omega_2\omega_3) \geq 2\} \tag{1.2}$$

Finally, we must determine the likelihood of each sequence occurring. We can do this by letting p be the probability of the coin landing on heads and $q = 1 - p$ be the probability of tails. Then, the probabilities of each ω occurring are

$$\begin{gathered}\mathbb{P}(HHH) = p^3, \mathbb{P}(HHT) = p^2q, \mathbb{P}(HTH) = p^2q, \mathbb{P}(THH) = p^2q, \\ \mathbb{P}(HTT) = pq^2, \mathbb{P}(THT) = pq^2, \mathbb{P}(TTH) = pq^2, \mathbb{P}(TTT) = q^3\end{gathered} \tag{1.3}$$

Further, we can figure out the probability of an event by adding up the individual probabilities of each possible sequence in the event:

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(HTT) + \mathbb{P}(THT) + \mathbb{P}(TTH) + \mathbb{P}(TTT) \\ &= pq^2 + pq^2 + pq^2 + q^3\end{aligned} \tag{1.4}$$

Note that $\sum_{\omega\in\Omega} \mathbb{P}(\omega) = 1$, or $\mathbb{P}(\Omega) = 1$.

Definition 1.1.1. *A* **finite probability space** *consists of a sample space* Ω *and a probability measure* $\mathbb{P}$ *that takes each element* $\omega \in \Omega$ *and assigns it a value in the interval [0,1]. We denote it by* $(\Omega, \mathbb{P})$.

1.2 Random Variables

The finite probability space models a situation in which a random experiment is conducted. These experiments typically produce numerical data which can be written as *random variables.*

Definition 1.2.1. *Let* $(\Omega, \mathbb{P})$ *be a finite probability space. A* **random variable** *is a function that maps* Ω *onto* $\mathbb{R}$.

It is important to note that random variables themselves do not depend on the probability measure $\mathbb{P}$. The *distribution* of a random variable uses $\mathbb{P}$ to determine the probabilities of the random variable taking different values. We can keep our sample space and random variable function consistent between two finite probability spaces, but if their probability measures differ, the same random variable can have two different distributions. This is illustrated in the example below.

Example 1.2.2. Suppose we toss a coin 3 times such that $\Omega = \{$HHH, HHT, HTH, THH, HTT, THT, TTH, TTT$\}$. Now define the random variable X to be the total number of heads and Y to be the total number of tails. Therefore,

$$\begin{aligned} X(HHH) &= 3 \\ X(HHT) = X(HTH) = X(THH) &= 2 \\ X(HTT) = X(THT) = X(TTH) &= 1 \\ X(TTT) &= 0 \end{aligned} \tag{1.5}$$

$$\begin{aligned} Y(HHH) &= 0 \\ Y(HHT) = Y(HTH) = Y(THH) &= 1 \\ Y(HTT) = Y(THT) = Y(TTH) &= 2 \\ Y(TTT) &= 3 \end{aligned} \tag{1.6}$$

We have not yet set the probability measure $\mathbb{P}$ for our finite probability space and we were still able to determine the various values for our random variables X and Y.

Now, let us specify $\mathbb{P}_1$ such that the probability of getting heads is $\frac{1}{2}$ (and therefore the probability of getting tails is $1 - \frac{1}{2} = \frac{1}{2}$). Then, $\mathbb{P}_1(\omega) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$ for all $\omega \in \Omega$. Let $\mathbb{P}\{X = i\} := \mathbb{P}\{\omega \in \Omega \,;\, X(\omega) = i\}$. The distribution would be as follows:

$$\begin{aligned} \mathbb{P}_1\{X = 3\} = \mathbb{P}_1\{HHH\} &= \frac{1}{8} \\ \mathbb{P}_1\{X = 2\} = \mathbb{P}_1\{HHT, HTH, THH\} &= \frac{3}{8} \\ \mathbb{P}_1\{X = 1\} = \mathbb{P}_1\{HTT, THT, TTH\} &= \frac{3}{8} \\ \mathbb{P}_1\{X = 0\} = \mathbb{P}_1\{TTT\} &= \frac{1}{8} \end{aligned} \tag{1.7}$$

Now let us use a different $\mathbb{P}$, say $\mathbb{P}_2$, such that the probability of getting heads is $\frac{2}{3}$ and the probability of getting tails is $\frac{1}{3}$. By similar calculations, we get the following distribution:

$$\begin{aligned} \mathbb{P}_2\{X = 3\} = \mathbb{P}_2\{HHH\} &= \frac{8}{27} \\ \mathbb{P}_2\{X = 2\} = \mathbb{P}_2\{HHT, HTH, THH\} &= \frac{12}{27} \\ \mathbb{P}_2\{X = 1\} = \mathbb{P}_2\{HTT, THT, TTH\} &= \frac{6}{27} \\ \mathbb{P}_2\{X = 0\} = \mathbb{P}_2\{TTT\} &= \frac{1}{27} \end{aligned} \tag{1.8}$$

Clearly these distributions differ, even though we are measuring the same random variable.

Knowing the probability distribution of a random variable, we can calculate a single value of what we expect the result of our random experiment to be if we were to actually conduct it.

Definition 1.2.3. *Let X be a random variable defined on a finite probability space $(\Omega, \mathbb{P})$. The* **expected value***, $\mathbb{E}$, of X is defined as*

$$\mathbb{E}X = \sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega) \tag{1.9}$$

The **variance** *of X is*

$$Var(X) = \mathbb{E}(X - \mathbb{E}X^2) \tag{1.10}$$

Using our sample space $(\Omega, \mathbb{P}_2)$, we can expect the number of heads that show up after three coin tosses to be

$$\begin{aligned}\mathbb{E}X =& \sum_{\omega \in \Omega} X(\omega)\mathbb{P}_2(\omega) \\ =& X(HHH)\mathbb{P}_2\{HHH\} + X(HHT)\mathbb{P}_2\{HHT\} + X(HTH)\mathbb{P}_2\{HTH\} + \\ & X(THH)\mathbb{P}_2\{THH\} + X(HTT)\mathbb{P}_2\{HTT\} + X(THT)\mathbb{P}_2\{THT\} + \\ & X(TTH)\mathbb{P}_2\{TTH\} + X(TTT)\mathbb{P}_2\{TTT\} \\ =& 3 \times \frac{8}{27} + 2 \times \frac{4}{27} + 2 \times \frac{4}{27} + 2 \times \frac{4}{27} + 1 \times \frac{2}{27} + 1 \times \frac{2}{27} + 1 \times \frac{2}{27} + 0 \times \frac{1}{27} \\ =& 2\end{aligned} \tag{1.11}$$

This agrees with our intuition that if we have a $\frac{2}{3}$ probability of getting heads, then every 3 tosses, we should expect to see 2 heads.

This is the best estimate if we are provided with no additional information. However, say we change the scenario so that we have already flipped the coin two times, and both times have turned up tails! Surely, getting two heads by our third toss would be impossible. We can adjust our expected value using the information about the first two tosses that we now have. This is called the *conditional expectation of X based on the information we have.*

Definition 1.2.4. *Given n coin tosses such that $1 \leq n \leq N$, there are 2^{N-n} possible continuations $\omega_{n+1} \ldots \omega_N$ of the sequence $\omega_1 \ldots \omega_n$. Let p be the probability of getting heads and $q = 1 - p$ be the probability of getting tails. Let $\#H(\omega_{n+1} \ldots \omega_N)$ be the number of heads in $\omega_{n+1} \ldots \omega_N$ and $\#T(\omega_{n+1} \ldots \omega_N)$ be the number of tails in $\omega_{n+1} \ldots \omega_N$. Then, our* **conditional expectation of X based on the information at time *n*** *is*

$$\begin{aligned}&\mathbb{E}_n[X](\omega_1 \ldots \omega_n) \\ &= \sum_{\omega_{n+1} \ldots \omega_N} p^{\#H(\omega_{n+1} \ldots \omega_N)} q^{\#T(\omega_{n+1} \ldots \omega_N)} X(\omega_1 \ldots \omega_n \omega_{n+1} \ldots \omega_N)\end{aligned} \tag{1.12}$$

Theorem 1.2.5 (Fundamental properties of conditional expectations)**.** *Let N be a positive integer and let X, Y be random variables depending on the first N coin tosses. Let $0 \leq n \leq N$ be given. Then, the following properties hold.*

(i) ***Linearity of conditional expectations****. For all constants c_1 and c_2,*

$$\mathbb{E}_n[c_1 X + c_2 Y] = c_1 \mathbb{E}_n[X] + c_2 \mathbb{E}_n[Y] \tag{1.13}$$

(ii) ***Taking out what is known****. If X actually depends on the first n coin tosses, then*

$$\mathbb{E}_n[XY] = X \cdot \mathbb{E}_n[Y] \tag{1.14}$$

(iii) ***Iterated conditioning****. If* $0 \leq n \leq m \leq N$*, then*

$$\mathbb{E}_n[\mathbb{E}_m[X]] = \mathbb{E}_n[X] \tag{1.15}$$

(iv) ***Independence****. If* X *depends only on tosses* $n + 1$ *through* N*, then*

$$\mathbb{E}_n[X] = \mathbb{E}[X] \tag{1.16}$$

The proof of this theorem results largely from the definition of conditional expectation and will be left to the reader to work through if desired.

1.3 Martingales and Markov Processes

Definition 1.3.1. *Consider the coin toss scenario. Let* $M_0, M_1, \ldots, M_N$ *be a sequence of random variables such that each* M_n*,* $0 \leq n \leq N$*, depends only on the first* n *coin tosses. Then, we call this sequence an* **adapted stochastic process***.*

We can further classify adapted stochastic processes by how we expect them to change from one coin toss to the next. Given M_n, and calculating M_{n+1} using our definition of conditional expectation, we will find that certain adapted stochastic processes can be expected to rise, others can be expected to fall, and others can be expected to remain constant. Such classifications are defined formally below.

Definition 1.3.2. *Let the sequence of random variables* $M_0, M_1, \ldots, M_N$ *be an adapted stochastic process.*

If $M_n = \mathbb{E}_n[M_{n+1}], n = 0, 1, \ldots, N-1$*, this process is a* **martingale***.*

If $M_n \leq \mathbb{E}_n[M_{n+1}], n = 0, 1, \ldots, N-1$*, this process is a* **submartingale***.*

If $M_n \geq \mathbb{E}_n[M_{n+1}], n = 0, 1, \ldots, N-1$*, this process is a* **submartingale***.*

Though the martingale property is "one-step-ahead" by only specifying the expectation of the immediate next variable in the sequence, we can extend the property for any variable appearing after M_n using the *iterated conditioning* property of conditional expectation. For example, if the sequence $M_0, M_1, \ldots, M_N$ is a martingale and $n \leq N - 2$, we know that

$$M_{n+1} = \mathbb{E}_{n+1}[M_{n+2}]. \tag{1.17}$$

Applying the iterated conditioning property, we see that

$$\mathbb{E}_n[M_{n+1}] = \mathbb{E}_n[\mathbb{E}_{n+1}[M_{n+2}]] = \mathbb{E}_n[M_{n+2}]. \tag{1.18}$$

And since $M_n = \mathbb{E}_n[M_{n+1}]$,

$$M_n = \mathbb{E}_n[M_{n+2}]. \tag{1.19}$$

The notion of conditional expectations give rise to many algorithms that serve as powerful predictive tools in different scenarios. But, if an adapted stochastic process is *path independent*; namely, if a variable in the sequence only depends on the immediate former variable, often times these algorithms can be greatly simplified. This type of process is known as a *Markov process*.

Definition 1.3.3. *Let* $X_0, X_1, \ldots, X_N$ *be an adapted stochastic process. If for* $0 \leq n \leq N-1$ *there exists a function g(x) for every f(x) depending on n and f such that*

$$\mathbb{E}_n[f(X_{n+1})] = g(X_n) \tag{1.20}$$

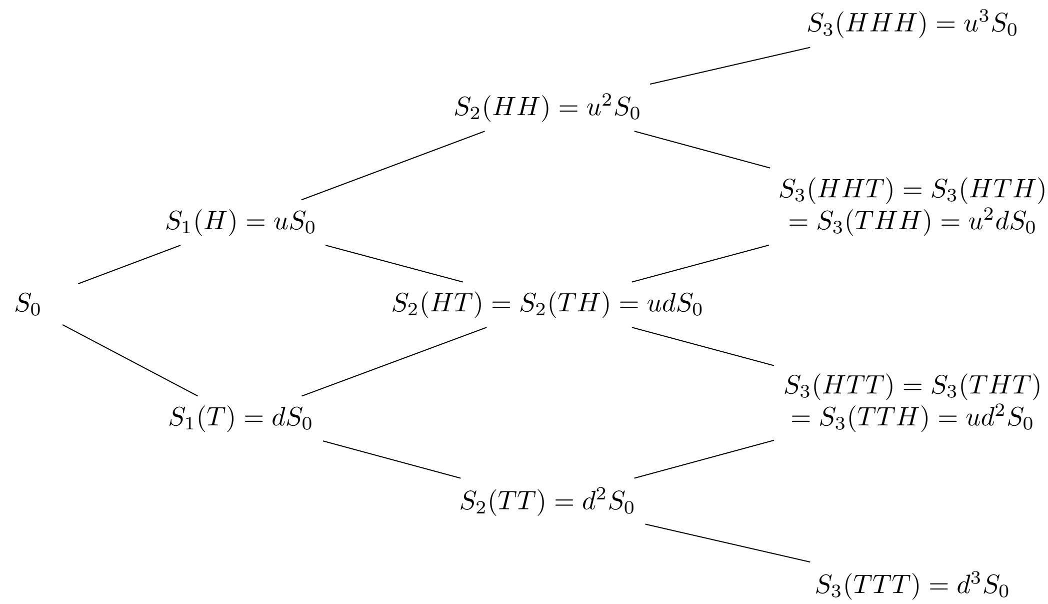
We call $X_0, X_1, \ldots, X_N$ *a* **Markov process***.*

The martingale is a special case of (1.20) with $f(x) = x$ and $g(x) = x$. But since there must exist a $g(x)$ for *every* $f(x)$ such that equation (1.20) holds in order for a process to be Markov, not every martingale is a Markov process. Of course, since $g(x)$ does not have to equal x, not every Markov process is martingale.

Chapter 2

The Binomial Model

2.1 Structure



A three-period binomial model.

The binomial model above displays the possible paths of a stock price, with $S_0 > 0$ being the initial price per share at time $t = 0$. The model is broken up into periods, with time $t = 1$ being the end of the first period, $t = 2$ the end of the second, and $t = 3$ the end of the third. We extend our coin toss scenario onto this model and let heads represent the stock price increasing and tails represent the stock price decreasing. At the end of each period the stock can take on two possible values, one of which is greater than its previous value and one of which is less. Here we can introduce two new variables, u and d, that serve as the two possible ratios of the new stock price to the former stock price. If we have a total of N periods, then for each n, $1 < n \leq N$ and for every possible $\omega_1 \dots \omega_{n-1}$,

$$u = \frac{S_n(\omega_1 \dots \omega_{n-1} H)}{S_{n-1}(\omega_1 \dots \omega_{n-1})}, \quad d = \frac{S_n(\omega_1 \dots \omega_{n-1} T)}{S_{n-1}(\omega_1 \dots \omega_{n-1})} \tag{2.1}$$

And when $n = 1$,

$$u = \frac{S_1(H)}{S_0}, \quad d = \frac{S_1(T)}{S_0} \tag{2.2}$$

Clearly, we will have $u > 1$ and $0 < d < 1$. It is important to note that stock prices can never be negative. Multiplying both sides by the denominator will result in the equations given in the binomial model. Since each S_n is a random variable that depends on the first n coin tosses, $S_0, S_1, \ldots, S_N$ is an adapted stochastic process. We still use the probabilities p and $q = 1 - p$ for tossing heads and tails, respectively.

We will also introduce a variable r to be the money market *interest rate*. If we invest one dollar in the money market at time $t = 0$, we will get $1 + r$ dollars back at time $t = 1$. If we borrow one dollar from the money market at time $t = 0$, we will owe $1 + r$ dollars at time $t = 1$. We use the binomial asset-pricing model to price derivatives by replicating them in the stock and money markets, so we must assume that $0 < d < 1 + r < u$ in order to avoid *arbitrage*.

Definition 2.1.1. Arbitrage *is a trading strategy by which participants start with no money, have a positive probability of making money, and have zero probability of losing money.*

We must be careful to avoid arbitrage in our model because otherwise, analyzing it would result in nonsensical conclusions. Though money markets in real life sometimes have arbitrage, it is always quickly discovered and solved through trading. Assuming a no-arbitrage model results in the inequalities $0 < d < 1 + r < u$ for the following reasons:

> $\mathbf{d < 1 + r}$. Otherwise, one could start with zero dollars, borrow from the money market, and guarantee they make *at least* what they will owe in interest by investing in the stock. There would be no chance of losing money, hence there would be arbitrage.
>
> $\mathbf{1 + r < u}$. Otherwise, one could sell the stock short and invest the money in the money market. Then, when the security expires, it will cost *at most* what was made from the money market investment to replace the stock. Again, there would be no chance of losing money, and there would be arbitrage.

It is common, but not necessary, to have d and u such that $d = \frac{1}{u}$.

2.2 Pricing Derivatives

In order to ensure no arbitrage when pricing our derivatives, there exists an *arbitrage pricing theory* by replicating it through trading in the stock and money markets. Firstly, one should assign a price to the security to prevent the possibility of arbitrage. Secondly, for an expiration time N, at any time $n < N$ one should be able to imagine selling the derivative for a price and invest that money in the stock and money markets such that the value of this portfolio at time N matches the payoff of the security.

Let's go through the process using the example of a European call option that expires at time $t = 2$. We define K to be the *strike price* at which the owner may buy one stock of the share at the expiration date. Therefore, the payoff of the deal will be $V_2 = (S_2 - K)^+$, where $(A)^+$ means the maximum value of the set $(0, A)$. We are trying to determine V_0 which is the no-arbitrage price for the option at time $t = 0$.

Now we suppose an agent sells the option for V_0 at time $t = 0$ and buys Δ_0 shares of stock (priced at S_0 per share). She invests $V_0 - \Delta_0 S_0$ dollars in the money market. This will be a negative quantity, taken to mean that she has borrowed $\Delta_0 S_0 - V_0$ dollars from the money market to finance her stock purchase. At time $t = 1$, the agent has a portfolio valued at

$$X_1 = \Delta_0 S_1 + (1 + r)(V_0 - \Delta_0 S_0) \tag{2.3}$$

But, because the value of S_1 depends on the outcome of the first coin toss, we really have two equations:

$$X_1(H) = \Delta_0 S_1(H) + (1+r)(V_0 - \Delta_0 S_0) \tag{2.4}$$
$$X_1(T) = \Delta_0 S_1(T) + (1+r)(V_0 - \Delta_0 S_0) \tag{2.5}$$

The agent can adjust her portfolio based on the outcome of the first coin toss. She now decides to hold Δ_1 shares of stock and invests $X_1 - \Delta_1 S_1$ in the money market. Because we are replicating the option, we want the portfolio valued at V_2. In other words, we want

$$V_2 = \Delta_1 S_2 + (1+r)(X_1 - \Delta_1 S_1) \tag{2.6}$$

But, S_2 and V_2 depend on the first two coin tosses so we really have four equations:

$$V_2(HH) = \Delta_1(H)S_2(HH) + (1+r)(X_1(H) - \Delta_1(H)S_1(HH)) \tag{2.7}$$
$$V_2(HT) = \Delta_1(H)S_2(HT) + (1+r)(X_1(H) - \Delta_1(H)S_1(HT)) \tag{2.8}$$
$$V_2(TH) = \Delta_1(T)S_2(TH) + (1+r)(X_1(T) - \Delta_1(T)S_1(TH)) \tag{2.9}$$
$$V_2(TT) = \Delta_1(T)S_2(TT) + (1+r)(X_1(T) - \Delta_1(T)S_1(TT)) \tag{2.10}$$

Now, we have six equations and six unknowns: V_0, Δ_0, $\Delta_1(H)$, $\Delta_1(T)$, $X_1(H)$, $X_1(T)$.

Let's first find $X_1(T)$ and $\Delta_1(T)$ by looking at equations (2.9) and (2.10). We can easily solve for $\Delta_1(T)$ by subtracting (2.10) from (2.9) and solving for it. The solution we obtain is called the *delta-hedging formula*.

$$\Delta_1(T) = \frac{V_2(TH) - V_2(TT)}{S_2(TH) - S_2(TT)} \tag{2.11}$$

Next we solve for $X_1(T)$. Since $S_2(TH)$, $S_2(TT)$, $S_1(TH)$, and $S_1(TT)$ are random variables, we know that if our first coin toss is tails, we have some probability $\tilde{p}$ and $\tilde{q} = 1 - \tilde{p}$ probability of getting tails. In other words, we have $\tilde{p}$ probability of equation (2.9) and $\tilde{q}$ probability of equation (2.10). In order to find the expected value of $X_1(T)$, we multiply (2.9) by $\tilde{p}$ and (2.10) $\tilde{q}$ and add them together. We are also going to divide all terms by $(1+r)$.

$$\begin{aligned}\frac{1}{1+r}\Big(\tilde{p}V_2(TH) + \tilde{q}V_2(TT)\Big) = \Delta_1(T)\Big(&\frac{1}{1+r}[\tilde{p}S_2(TH) + \tilde{q}S_2(TT)] - \\ &(\tilde{p} - \tilde{q})S_1(T)\Big) + (\tilde{p} - \tilde{q})X_1(T)\end{aligned} \tag{2.12}$$

Since $(\tilde{p} + \tilde{q}) = 1$, we can simplify (2.12) to get

$$\frac{1}{1+r}\Big(\tilde{p}V_2(TH) + \tilde{q}V_2(TT)\Big) = \Delta_1(T)\Big(\frac{1}{1+r}[\tilde{p}S_2(TH) + \tilde{q}S_2(TT)] - S_1(T)\Big) + X_1(T) \tag{2.13}$$

If we choose $\tilde{p}$ such that

$$S_1(T) = \frac{1}{1+r}[\tilde{p}S_2(TH) + \tilde{q}S_2(TT)] \tag{2.14}$$

We can greatly simplify (2.13). Use $S_1(T) = dS_0$, $S_2(TH) = duS_0$, and $S_2(TT) = ddS_0$, and $\tilde{q} = 1 - \tilde{p}$. Then, (2.14) becomes

$$dS_0 = \frac{1}{1+r}[\tilde{p}duS_0 + (1-\tilde{p})ddS_0] \tag{2.15}$$

Divide both sides by dS_0 to get

$$1 = \frac{1}{1+r}[\tilde{p}u + (1-\tilde{p})d] \tag{2.16}$$

And from there we can easily solve for $\tilde{p}$. Similarly, we can solve (2.14) for $\tilde{q}$.

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

These probabilities $\tilde{p}$ and $\tilde{q}$ are called *risk neutral probabilities.* Normally, the average growth rate of a stock exceeds that of the money market; otherwise, it would not make sense to risk investing in stock. So, the *actual probabilities* p and q of a given stock should satisfy

$$(1+r)S_1(T) < pS_2(TH) + qS_2(TT)$$

We instead chose $\tilde{p}$ and $\tilde{q}$ to satisfy (2.14) to assist us in our calculations.

Upon simplifying equation (2.13) to reflect our strategic choice of $\tilde{p}$ and $\tilde{q}$, we got the value the replicating portfolio should have at time $t = 1$ if the stock price goes down (our coin toss results in tails):

$$X_1(T) = \frac{1}{1+r}\Big(\tilde{p}V_2(TH) + \tilde{q}V_2(TT)\Big) \tag{2.17}$$

In terms of our stock option, we define this to be the *price of the option at time t=1 if the first coin toss results in tail*, denoted by $V_1(T)$. It is an instance of the *risk-neutral pricing formula*, which will be formally defined later in this chapter.

$$V_1(T) = \frac{1}{1+r}\Big(\tilde{p}V_2(TH) + \tilde{q}V_2(TT)\Big) \tag{2.18}$$

By going through a similar procedure with equations (2.7) and (2.8), we come up with the following two equations for $\Delta_1(H)$ and $V_1(H) = X_1(H)$:

$$\Delta_1(H) = \frac{V_2(HH) - V_2(HT)}{S_2(HH) - S_2(HT)} \tag{2.19}$$

$$V_1(H) = \frac{1}{1+r}\Big(\tilde{p}V_2(HH) + \tilde{q}V_2(HT)\Big) \tag{2.20}$$

The latter equation is the *price of the option at time t=1 if the first toss results in head.*

Finally, we can solve for V_0 and Δ_0 by plugging our values for $X_1(H) = V_1(H)$ and $X_1(T) = V_1(T)$ into equations (2.4) and (2.5).

Clearly, we have three stochastic processes: (Δ_0, Δ_1), (X_0, X_1, X_2), and (V_0, V_1, V_2). We recursively defined our portfolio, and if we specify values for X_0, Δ_0, $\Delta_1(H)$, and $\Delta_1(T)$, we can recursively define a replicating portfolio with any number of periods by the *wealth equation*

$$X_{n+1} = \Delta_n S_{n+1} + (1+r)(X_n - \Delta_n S_n) \tag{2.21}$$

and we can price our stock option in a similar manner as we did for the two-period model.

Theorem 2.2.1 (Replication in the multiperiod binomial model)**.** *Consider an N-period binomial asset-pricing model with $0 < d < 1 + r < u$ and with*

$$\tilde{p} = \frac{1+r-d}{u-d}, \quad \tilde{q} = \frac{u-1-r}{u-d} \tag{2.22}$$

Let V_N be a random variable (a derivative security paying off at time N) depending on the first N coin tosses $\omega_1\omega_2\dots\omega_N$. Define recursively backward in time the sequence of random variables $V_{N-1}, V_{N-2}, \dots V_0$, by

$$V_n(\omega_1\omega_2\dots\omega_n) = \frac{1}{1+r}[\tilde{p}V_{n+1}(\omega_1\omega_2\dots\omega_n H) + \tilde{q}V_{n+1}(\omega_1\omega_2\dots\omega_n T)] \tag{2.23}$$

such that each V_n depends on the first n coin tosses $\omega_1\omega_2\ldots\omega_n$, $0 \leq n \leq N-1$. Next, define

$$\Delta_n(\omega_1\omega_2\ldots\omega_n) = \frac{V_{n+1}(\omega_1\omega_2\ldots\omega_n H) - V_{n+1}(\omega_1\omega_2\ldots\omega_n T)}{S_{n+1}(\omega_1\omega_2\ldots\omega_n H) - S_{n+1}(\omega_1\omega_2\ldots\omega_n T)} \tag{2.24}$$

where again $0 \leq n \leq N-1$. If we set $X_0 = V_0$ and define recursively forward in time the portfolio values $X_1, X_2, \ldots, X_N$ by the wealth equation, then we will have

$$X_N(\omega_1\omega_2\ldots\omega_N) = V_N(\omega_1\omega_2\ldots\omega_N) \tag{2.25}$$

for all $\omega_1\omega_2\ldots\omega_N$.

Proof. We will use a proof of induction on n that

$$X_n(\omega_1\omega_2\ldots\omega_n) = V_n(\omega_1\omega_2\ldots\omega_n) \tag{2.26}$$

for all $\omega_1\omega_2\ldots\omega_n$, $0 \leq n \leq N$

Base Case: n = 0 The base case is given by the definition $X_0 = V_0$.

Inductive Step Assume the theorem holds for some $n < N$. We wish to show that it also holds for $n+1$. We know by the hypothesis that $X_n(\omega_1\omega_2\ldots\omega_n) = V_n(\omega_1\omega_2\ldots\omega_n)$. Now, consider $\omega_1\omega_2\ldots\omega_{n+1}$ to be fixed and arbitrary. We do not know whether ω_{n+1} is H or T, so we must consider both cases. However, the steps in both cases are very similar, so we will only prove the case when $\omega_{n+1} \coloneqq H$. From the wealth equation, we know

$$\begin{aligned} X_{n+1}(\omega_1\omega_2\ldots\omega_n H) =&\Delta_n(\omega_1\omega_2\ldots\omega_n)uS_n(\omega_1\omega_2\ldots\omega_n)+ \\ &(1+r)\Big(X_n(\omega_1\omega_2\ldots\omega_n) - \Delta_n(\omega_1\omega_2\ldots\omega_n)S_n(\omega_1\omega_2\ldots\omega_n)\Big) \end{aligned}$$

For simplification purposes, suppress the sequence $\omega_1\omega_2\ldots\omega_n$ so this equation becomes

$$X_{n+1}(H) = \Delta_n uS_n + (1+r)(X_n - \Delta_n S_n) \tag{2.27}$$

From equation (2.24), and using our simplified notation, we know that

$$\Delta_n = \frac{V_{n+1}(H) - V_{n+1}(T)}{S_{n+1}(H) - S_{n+1}(T)} \tag{2.28}$$

We may rewrite this as

$$\Delta_n = \frac{V_{n+1}(H) - V_{n+1}(T)}{(u-d)S_n} \tag{2.29}$$

Now, we can rewrite equation (2.27) and substitute this in for Δ_n, along with using our hypothesis that $V_n = X_n$:

$$\begin{aligned} X_{n+1}(H) &= (1+r)X_n + \Delta_n S_n(u-(1+r)) \\ &= (1+r)V_n + \frac{(V_{n+1}(H) - V_{n+1}(T))(u-(1+r)}{u-d} \end{aligned}$$

We can also use equation (2.22) to substitute $\tilde{q}$ into the equation:

$$\begin{aligned} X_{n+1}(H) &= (1+r)V_n + \tilde{q}V_{n+1}(H) - \tilde{q}V_{n+1}(T) \\ &= \tilde{p}V_{n+1}(H) + \tilde{q}V_{n+1}(T) + \tilde{q}V_{n+1}(H) - \tilde{q}V_{n+1}(T) \\ &= V_{n+1}(H) \end{aligned}$$

A similar argument shows that $X_{n+1}(T) = V_{n+1}(T)$. So, no matter what ω_{n+1} is, we have

$$X_{n+1}(\omega_1\omega_2\ldots\omega_{n+1}) = V_{n+1}(\omega_1\omega_2\ldots\omega_{n+1})$$

The induction step is complete. □

This theorem was prefaced with the example of a European call option for which the payoff only depends on the final stock price. The theorem also applies to *path dependent* options whose payoff depends on the different values the stock takes on between its initial value and its final value. We will explore the path dependent example of American derivatives in Chapter 3.

2.3 Properties

In the previous section we came up with risk-neutral probabilities $\tilde{p}$ and $\tilde{q}$. We can define these probabilities as the probability measure $\tilde{\mathbb{P}}$. Similarly, our risk-neutral expected value for random variable X under these risk-neutral probabilities is $\tilde{\mathbb{E}}X$. Recall that

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

It is easy to check that

$$\frac{\tilde{p}u + \tilde{q}d}{1+r} = 1$$

So, for all n and for every coin toss sequence $\omega_1 \dots \omega_n$, we can write

$$S_n(\omega_1 \dots \omega_n) = \frac{1}{1+r}\Big(\tilde{p}S_{n+1}(\omega_1 \dots \omega_n H) + \tilde{q}S_{n+1}(\omega_1 \dots \omega_n T)\Big) \tag{2.30}$$

But, using our knowledge of conditional expectation from Definition 1.2.4, we can rewrite the the above equation as follows:

$$S_n = \frac{1}{1+r}\tilde{\mathbb{E}}_n[S_{n+1}] \tag{2.31}$$

If we divide both sides of (2.31) by $(1+r)^n$, we get the equation

$$\frac{S_n}{(1+r)^n} = \tilde{\mathbb{E}}_n\Big[\frac{S_{n+1}}{(1+r)^n n+1}\Big] \tag{2.32}$$

where $\frac{S_n}{(1+r)^n}$ is the discounted stock price at time n. This equation expresses a key fact that under the risk-neutral probability measure $\tilde{\mathbb{P}}$, *the best estimate of the discounted stock price at time $n+1$ is the discounted stock price at time n*. In other words, the discounted stock price process is a martingale.

Theorem 2.3.1. *Consider the general binomial model with $0 < d < 1 + r < u$. Under the risk neutral measure, the discounted stock price is a martingale, i.e. equation (2.32) holds at every time n and for every sequence of coin tosses.*

By definition of $\tilde{\mathbb{P}}$, our expected rate of growth of a stock is equal to the interest rate, r, of the money market (as was defined through equation (2.14)). Because of this, no matter how an agent divides up the wealth of their replicating portfolio between the stock and the money market, it will have an average rate of growth r. This is restated in the theorem below.

Theorem 2.3.2. *Consider the binomial model with N periods. Let $\Delta_0, \Delta_1, \dots, \Delta_{N-1}$ be an adapted portfolio process, let X_0 be a real number, and let $X_1, \dots, X_N$ be generated recursively by the wealth equation,*

$$X_{n+1} = \Delta_n S_{n+1} + (1+r)(X_n - \Delta_n S_n) \quad n = 0, 1, \dots, N-1$$

Then the discounted wealth process $\frac{X_n}{(1+r)^n}, n = 0, 1, \dots, N$ is a martingale under the risk-neutral measure. In other words,

$$\frac{X_n}{(1+r)^n} = \tilde{\mathbb{E}}_n\Big[\frac{X_{n+1}}{(1+r)^{n+1}}\Big], \quad n = 0, 1, \dots, N-1 \tag{2.33}$$

Proof.

$$
\begin{aligned}
\tilde{\mathbb{E}}_n\Big[\frac{X_{n+1}}{(1+r)^{n+1}}\Big] &= \tilde{\mathbb{E}}_n\Big[\frac{\Delta_n S_{n+1} + (1+r)(X_n - \Delta_n S_n)}{(1+r)^{n+1}}\Big] && \text{(Wealth equation)}\\
&= \tilde{\mathbb{E}}_n\Big[\frac{\Delta_n S_{n+1}}{(1+r)^{n+1}} + \frac{X_n - \Delta_n S_n}{(1+r)^n}\Big]\\
&= \tilde{\mathbb{E}}_n\Big[\frac{\Delta_n S_{n+1}}{(1+r)^{n+1}}\Big] + \tilde{\mathbb{E}}_n\Big[\frac{X_n - \Delta_n S_n}{(1+r)^n}\Big] && \text{(Linearity)}\\
&= \Delta_n \tilde{\mathbb{E}}_n\Big[\frac{S_{n+1}}{(1+r)^{n+1}}\Big] + \frac{X_n - \Delta_n S_n}{(1+r)^n} && \text{(Taking out what is known)}\\
&= \Delta_n \frac{S_n}{(1+r)^n} + \frac{X_n - \Delta_n S_n}{(1+r)^n} && \text{(Theorem 2.3.1)}\\
&= \Delta_n \frac{S_n}{(1+r)^n} - \Delta_n \frac{S_n}{(1+r)^n} + \frac{X_n}{(1+r)^n}\\
&= \frac{X_n}{(1+r)^n}
\end{aligned}
$$

□

The expression $\frac{X_n}{(1+r)^n}$ is called the $\tilde{\mathbb{P}}$-*martingale*. Because of the iterated conditioning property, we can infer from Theorem 2.3.2 that

$$
\tilde{\mathbb{E}}_n \frac{X_n}{(1+r)^n} = X_0, \quad n = 0, 1, \ldots, N \tag{2.34}
$$

A consequence of this is known as the *first fundamental theorem of asset pricing*. If we have $X_0 = 0$, then $\tilde{\mathbb{E}}_n \frac{X_n}{(1+r)^n} = 0$ and therefore $\tilde{\mathbb{E}}_n X_n = 0$. Therefore, there is no arbitrage in the asset-pricing model if we can find a risk neutral measure in it. Another consequence of Theorem 2.3.2 is the *Risk-Neutral Pricing Formula*, stated below as a theorem.

Theorem 2.3.3 (Risk-Neutral Pricing Formula)**.** *Consider an N-period binomial asset-pricing model with $0 < d < 1 + r < u$ and with risk-neutral probability measure $\tilde{\mathbb{P}}$. Let V_N be a random variable (a derivative security paying off at time N depending on coin tosses. For $0 \leq n \leq N$, the price of the derivative security at time n is given by the risk-neutral pricing formula*

$$
V_n = \tilde{\mathbb{E}}_n\Big[\frac{V_N}{(1+r)^{N-n}}\Big] \tag{2.35}
$$

Furthermore, the discounted price of the derivative security is a martingale under $\tilde{\mathbb{P}}$, i.e.

$$
\frac{V_n}{(1+r)^n} = \tilde{\mathbb{E}}_n\Big[\frac{V_{n+1}}{(1+r)^{n+1}}\Big], \quad n = 0, 1, \ldots, N-1 \tag{2.36}
$$

In addition to the discounted price of a derivative security being a martingale, we can show that the price process of a derivative security is Markov. Recall Definition 1.3.3. If we can find a function $g(x)$ for every $f(x)$ and for every n between 0 and $N-1$ such that $\tilde{\mathbb{E}}[f(X_{n+1})] = g(X_n)$, then $X_0, X_1, \ldots, X_N$ is a Markov process. We will first show that the stock price process is Markov.

Example 2.3.4 (Stock Price)**.** We know we can compute the stock price at time $n+1$ using the following formula

$$
S_{n+1}(\omega_1 \ldots \omega_n \omega_{n+1}) = \left\{ \begin{array}{ll} uS_n(\omega_1 \ldots \omega_n) & \text{if } \omega_{n+1} = H \\ dS_n(\omega_1 \ldots \omega_n) & \text{if } \omega_{n+1} = T \end{array} \right\} \tag{2.37}
$$

Therefore,

$$\mathbb{E}_n[f(S_{n+1})](\omega_1 \dots \omega_n) = pf[uS_n(\omega_1 \dots \omega_n)] + qf[dS_n(\omega_1 \dots \omega_n)] = g[S_n(\omega_1 \dots \omega_n)] \tag{2.38}$$

Where $g(x)$ is defined by $g(x) = pf(ux) + qf(dx)$. Hence, we have an algorithm for finding $g(x)$ for every $f(x)$ and the stock price process is Markov. And, it is Markov under the risk-neutral probability measure or the actual probability measure.

We know that the payoff of a derivative security at time N is a function v_N of the stock price at time N. In other words, $V_N = v_N(S_N)$. If we take equation (2.36) and divide both sides by $(1+r)^n$, we get

$$V_n = \frac{1}{1+r}\tilde{\mathbb{E}}_n[V_{n+1}], \quad n = 0, 1, \dots, N-1 \tag{2.39}$$

Since the stock price process is Markov, we can define V_{N-1} as

$$V_{N-1} = \frac{1}{1+r}\tilde{\mathbb{E}}_{N-1}[v_N(S_N)] = v_{N-1}(S_{N-1}) \tag{2.40}$$

for some function v_{N-1}. Similarly, we can define V_{N-2}:

$$V_{N-2} = \frac{1}{1+r}\tilde{\mathbb{E}}_{N-2}[v_{N-1}(S_{N-1})] = v_{N-2}(S_{N-2}) \tag{2.41}$$

In general, we can define recursively backwards the function v_n such that $V_n = v_n(S_n)$ by the algorithm

$$v_n(s) = \frac{1}{1+r}\Big[\tilde{p}v_{n+1}(us) + \tilde{q}v_{n+1}(ds)\Big], \quad n = N-1, N-2, \dots, 0 \tag{2.42}$$

This shows that the price process of any derivative security is Markov under the risk-neutral probability measure. We can restate our findings as a theorem.

Theorem 2.3.5. *Let $X_0, X_1, \dots, X_N$ be a Markov process under the risk-neutral probability measure $\tilde{\mathbb{P}}$ in the binomial model. Let $v_N(x)$ be a function of the dummy variable x, and consider a derivative security whose payoff at time N is $v_N(X_N)$. Then, for each n between 0 and N, the price of V_n of this derivative security is some function v_n of X_n, i.e.*

$$V_n = v_n(X_n), \quad n = 0, 1, \dots, N \tag{2.43}$$

There is a recursive algorithm for computing v_n whose exact formula depends on the underlying Markov process $X_0, X_1, \dots, X_N$.

Chapter 3

Application to American Derivative Securities

3.1 Introduction

So far, we have only looked at European derivative securities of which the owner can only choose to exercise on a given expiration date. This chapter discusses *American derivative securities* that can be exercised at *any point on or before the expiration date*. This means that there will exist an *optimal exercise date*, after which the derivative security will tend to lose value. As a result, the discounted price process of American derivative securities is a supermartingale (unlike the European case, for which the discounted price process is a martingale).

We will start by looking at those American derivative securities that are not path dependent and then move on to those that are path dependent. We will find that we can still come up with a pricing algorithm by replicating the derivative using the stock and money markets much like we did with European derivative securities.

3.2 Path-Independent

Recall the pricing algorithm for a European derivative security. We use an N-period binomial model with up factor u, down factor d, and interest rate r such that $0 < d < 1 + r < u$. Let $g(S_N)$ be the function that tells us the payoff of the derivative security at time N. As was stated in Theorem 2.3.5, we can write the value V_n of this derivative security as a function v_n of the price of the stock. Using equation (2.42) and using the risk-neutral measure $\widetilde{\mathbb{P}}$, we can define v_n recursively backwards:

$$\begin{aligned} v_N(s) &= [g(s)]^+, \\ v_n(s) &= \frac{1}{1+r}\Big[\tilde{p}v_{n+1}(us) + \tilde{q}v_{n+1}(ds)\Big], \quad n = N-1, N-2, \ldots, 0 \end{aligned} \tag{3.1}$$

We also define the replicating portfolio using equation (2.24) to be

$$\Delta_n = \frac{v_{n+1}(uS_n) - v_{n+1}(dS_n)}{(u-d)S_n}, \quad n = 0, 1, \ldots, N \tag{3.2}$$

For an American derivative security, we will still have a payoff function $g(s)$. But, since the owner of the security can choose to exercise it at any time up to the expiration date and receive payment $g(S_n)$, we need to make sure our replicating portfolio *always* has a value X_n of at least $g(S_n)$ for $n \leq N$, i.e.

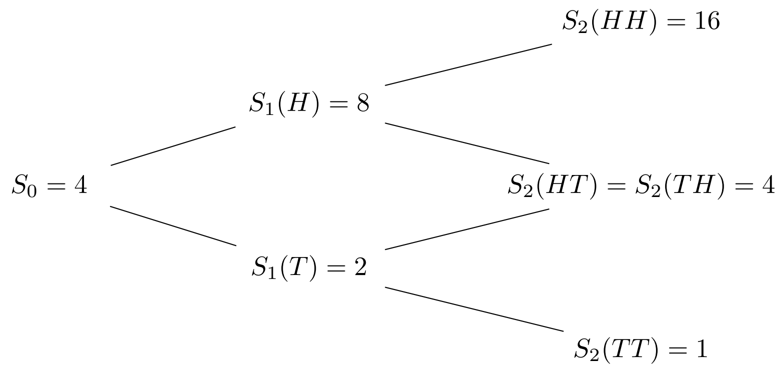
$$X_n \geq g(S_n), \quad n = 0, 1, \ldots, N \tag{3.3}$$

The function $g(S_n)$ tells us the *intrinsic value* of the derivative security at time n. Therefore, the derivative security will always have a value of at least $g(S_n)$, and since the value of the replicating portfolio must *match* that of the derivative security, it must also have a value of at least $g(S_n)$. Our definition of v_n is modified in the case of the American derivative security to account for this.

$$
\begin{aligned}
&v_N(s) = [g(s)]^+, \\
&v_n(s) = max\left[g(s), \frac{1}{1+r}\left[\tilde{p}v_{n+1}(us) + \tilde{q}v_{n+1}(ds)\right]\right], \quad n = N-1, N-2, \ldots, 0
\end{aligned}
\tag{3.4}
$$

with $V_n = v_n(S_n)$ being the derivative security price at time n.

Example 3.2.1 (American put option)**.** Let the following binomial model represent the possible stock prices of a given stock expiring at time $t = 2$:



A two-period binomial model.

Consider an American put option. Let $r = 4$, meaning our risk-neutral probabilities are $\tilde{p} = \tilde{q} = \frac{1}{2}$. Also let the strike price of the option be 5, meaning the owner will receive $(5 - S_n)$ if exercised at time n. So, $g(s) = 5 - s$, and we use equation (3.4) to define the function v_n as

$$
\begin{aligned}
&v_2(s) = [5 - s]^+, \\
&v_n(s) = max\left[5 - s, \frac{4}{5}\left[\frac{1}{2}v_{n+1}(2s) + \frac{1}{2}v_{n+1}(\frac{1}{2}s)\right]\right], \quad n = 1, 0
\end{aligned}
\tag{3.5}
$$

Plugging in our values for the stock price at time $n = 2, 1, 0$, this gives us

$$
\begin{aligned}
&v_2(16) = [5 - 16]^+ = 0, \\
&v_2(4) = [5 - 4]^+ = 1, \\
&v_2(1) = [5 - 1]^+ = 4, \\
&v_1(8) = max\left[5 - 8, \frac{2}{5}(0 + 1)\right] = max(-3, 0.40) = 0.40 \\
&v_1(2) = max\left[5 - 2, \frac{2}{5}(1 + 4)\right] = max(3, 2) = 3 \\
&v_0(4) = max\left[5 - 4, \frac{2}{5}(0.40 + 3)\right] = max(1, 1.36) = 1.36
\end{aligned}
$$

Had we calculated the values of v_n for the European equivalent of this put option, we would get $v_1(2) = 2$ instead of 3 and $v_0(4) = \frac{2}{5}(0.40 + 2) = 0.96$ instead of 1.36.

Next, we construct the replicating portfolio using equation (3.2), starting with initial capital $v_0 = 1.36$.

$$\Delta_0 = \frac{v_1(uS_0) - v_1(dS_0)}{(u-d)S_0} = \frac{v_1(8) - v_1(2)}{8-2} = -0.43 \tag{3.6}$$

We can verify this is correct by checking that $X_1(H) = v_1(S_1(H))$ and $X_1(T) = v_1(S_1(T))$ using the wealth equation (2.21):

$$\begin{aligned} X_1(H) &= \Delta_0 S_1(H) + (1+\frac{1}{4})(X_0 - \Delta_0 S_0) \qquad & X_1(T) &= \Delta_0 S_1(T) + (1+\frac{1}{4})(X_0 - \Delta_0 S_0) \\ &= (-0.43)(8) + \frac{5}{4}(1.36 + 0.43 \times 4) & &= (-0.43)(2) + \frac{5}{4}(1.36 + 0.43 \times 4) \\ &= (-3.44) + \frac{5}{4}(3.08) & &= (-0.86) + \frac{5}{4}(3.08) \\ &= 0.40 & &= 3 \\ &= v_1(S_1(H)) & &= v_1(S_1(T)) \end{aligned} \tag{3.7}$$

So, no matter what the result of our coin toss, the value of our replicating portfolio at time 1 will match that of the put option if we were to sell it at time 1.

The owner could choose to exercise the option at time $t = 1$, but let's assume otherwise. First, we consider the scenario where the first coin toss results in tails. The risk-neutral pricing formula tells us that at time $t = 1$, we want our portfolio to be valued at

$$\begin{aligned} X_1 = V_1 &= \tilde{\mathbb{E}}\Big[\frac{V_2}{1+r}\Big] \\ &= \frac{4}{5}\Big(\frac{1}{2}v_2(4) + \frac{1}{2}v_2(1)\Big) \\ &= 2 \end{aligned}$$

Since our portfolio is currently valued at 3 but we only need the value of our portfolio to be 2, we can consume a dollar and invest the rest. We again use equation (3.2) to calculate the value of $\Delta_1(T)$:

$$\Delta_1(T) = \frac{v_2(uS_1(T)) - v_2(dS_1(T))}{(u-d)S_1(T)} = \frac{v_2(4) - v_2(1)}{4-1} = -1$$

We once again verify this value of $\Delta_1(T)$ is correct by using the wealth equation to check if $X_2(TH) = v_2(S_2(TH))$ and $X_2(TT) = v_2(S_2(TT))$

$$\begin{aligned} X_2(TH) &= \Delta_1(T)S_2(TH) + (1+\frac{1}{4})(X_1(T) - \Delta_1(T)S_1(T)) \\ &= (-1)(4) + \frac{5}{4}(2 + 1 \times 2) \\ &= (-4) + \frac{5}{4}(4) \\ &= 1 \\ &= v_2(S_2(TH)) \\ X_2(TT) &= \Delta_1(T)S_2(TT) + (1+\frac{1}{4})(X_1(T) - \Delta_1(T)S_1(T)) \\ &= (-1)(1) + \frac{5}{4}(2 + 1 \times 2) \\ &= (-1) + \frac{5}{4}(4) \\ &= 4 \\ &= v_2(S_2(TT)) \end{aligned} \tag{3.8}$$

Finally, we consider the scenario where the first coin toss results in heads. With $X_1(H) = 0.40$, we first find $\Delta_1(H)$:

$$\Delta_1(H) = \frac{v_2(uS_1(H)) - v_2(dS_1(H))}{(u-d)S_1(H)} = \frac{v_2(16) - v_2(4)}{16 - 4} = -\frac{1}{12}$$

Then we may easily verify that this value for our replicating portfolio is correct by similar calculations as those in equation (3.8).

Lastly, we take a look at the discounted American put prices.

$$\begin{tikzcd}[column sep=small]
 & & (\frac{4}{5})^2 v_2(16) = 0 \\
 & \frac{4}{5}v_1(8) = 0.32 \arrow[ur, no head] \arrow[dr, no head] & \\
v_0(4) = 1.36 \arrow[ur, no head] \arrow[dr, no head] & & (\frac{4}{5})^2 v_2(4) = 0.64 \\
 & \frac{4}{5}v_1(2) = 2.40 \arrow[ur, no head] \arrow[dr, no head] & \\
 & & (\frac{4}{5})^2 v_2(1) = 2.56
\end{tikzcd}$$

Discounted American put prices.

Clearly the discounted American put price process is a supermartingale under the risk-neutral probability measure $\tilde{\mathbb{P}}$ because the average of the two rightward branches of each node is less than or equal to the node itself.

We can now formalize our findings from the previous example with a theorem.

Theorem 3.2.2 (Replication of path-independent American derivatives)**.** *Consider an N-period binomial asset-pricing model with $0 < d < 1 + r < u$ and with*

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

Let a payoff function $g(s)$ be given, and define recursively backward in time the sequence of functions $v_N(s), v_{N-1}(s), \ldots, v_0(s)$ by equation (3.4). Next define

$$\Delta_n = \frac{v_{n+1}(uS_n) - v_{n+1}(dS_n)}{(u-d)S_n}, \tag{3.9}$$

$$C_n = v_n(S_n) - \frac{1}{1+r}\Big[\tilde{p}v_{n+1}(uS_n) + \tilde{q}v_{n+1}(dS_n)\Big] \tag{3.10}$$

where $n = 0, 1, \ldots, N-1$. If we set $X_0 = v_0(S_0)$ and define recursively forward the portfolio values $X_1, X_2, \ldots, X_N$ by

$$X_{n+1} = \Delta_n S_{n+1} + (1+r)(X_n - C_n - \Delta_n S_n) \tag{3.11}$$

then we will have $X_n(\omega_1 \ldots \omega_n) = v_n(S_n(\omega_1 \ldots \omega_n))$ for all n and for all $\omega_1 \ldots \omega_n$. In particular, $X_n \geq g(S_n)$ for all n.

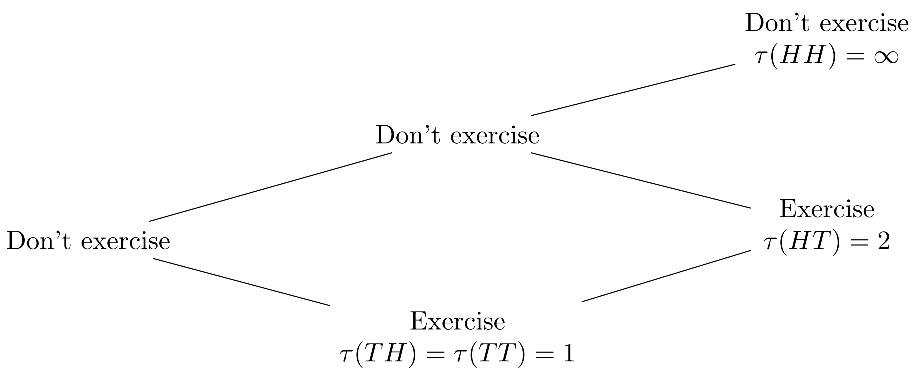
Note that equation (3.10) is the wealth equation with the possibility of consumption.

3.3 Stopping Times

When considering the general case of American derivative securities, which may be path-dependent, we must develop the notion of stopping times. This will help us determine when the optimal time for an owner to exercise the derivative security is based on the path of the stock price. If we look back at example 3.2.1, we can conclude that if the first coin toss results in tails, the owner should exercise the put option at time $t = 1$ because the value of our replicating portfolio was a dollar more than what the risk-neutral pricing formula determined it should be. If the first coin toss was heads, the owner would be out of the money since $S_1(H) = 8$ and the strike price was 5, so it would be wise for them to wait. Then, if the second coin toss resulted in tails they would be back in the money and they could exercise the option at time $t = 2$. But, if the second coin toss resulted in heads the owner would still be out of the money and should not exercise the option at all (i.e. let it expire). We can use the random variable τ to rewrite our findings as follows:

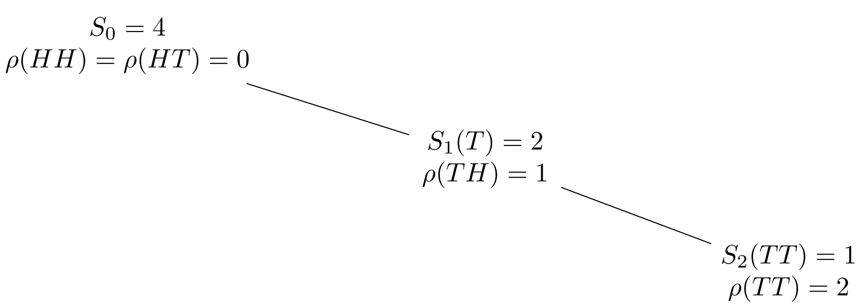
$$\tau(HH) = \infty, \quad \tau(HT) = 2, \quad \tau(TH) = 1, \quad \tau(TT) = 1 \tag{3.12}$$

And this can be displayed using the binomial model:



Exercise rule τ.

Where τ taking on the value of ∞ means we allow the option to expire. The random variable τ is defined on the sample space $\Omega = HH, HT, TH, TT$ and takes values in the set $0, 1, 2, \infty$. It is known as an exercise time. The values were determined based on the fact that the owner did not know what the next coin toss was to result in. If they had foreknowledge of the coin tossing, then they surely would have exercised the option at time $t = 0$ if the first coin toss were heads. And, if the coin toss sequence was TT, the owner would have waited until $t = 2$ to exercise instead of exercising the option at time $t = 1$. We can define another random variable, ρ, which is an exercise time that reflects this.



Exercise rule ρ.

But, ρ cannot be implemented without *insider information*. So, while it is an exercise time, it is not a stopping time. The random variable τ is a stopping time. We give the formal definition of *stopping* *time* below.

Definition 3.3.1. *In an N-period binomial model, a* **stopping time** *is a random variable* τ *that takes values* $0, 1, \ldots, N$ *or* ∞ *and satisfies the condition that if* $\tau(\omega_1\omega_2 \ldots \omega_n\omega_{n+1} \ldots \omega_N) = n$*, then* $\tau(\omega_1\omega_2 \ldots \omega_n\omega\prime_{n+1} \ldots \omega\prime_N) = n$ *for all* $\omega\prime_{n+1} \ldots \omega\prime_N$*.*

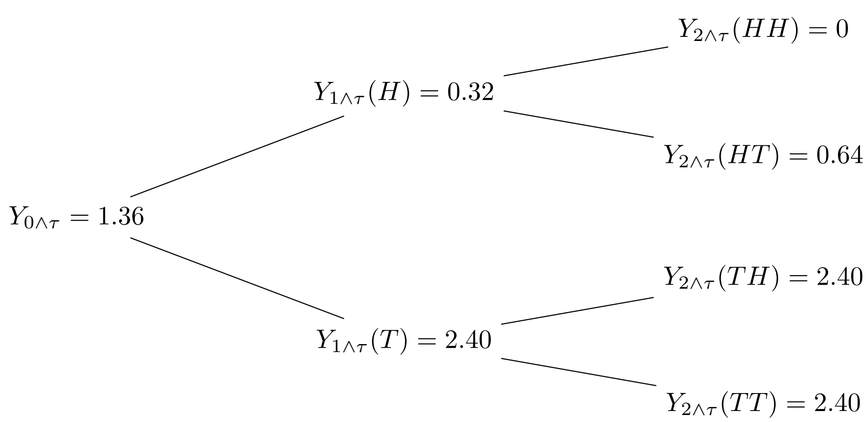
In other words, the stopping is based only on available information. Whenever we have a stochastic process and a stopping time, we can define a *stopped process*. For example, we can let Y_n be the process of the discounted American put prices from example 3.2.1.

$$Y_0 = 1.36, \quad Y_1(H) = 0.32, \quad Y_1(T) = 2.40,$$
$$Y_2(HH) = 0, \quad Y_2(HT) = Y_2(TH) = 0.64, \quad Y_2(TT) = 2.56$$

We can take our stopping time, τ, and let $n \wedge \tau$ mean the minimum of n and τ. We can let $Y_{0\wedge\tau} = Y_0 = 1.36$ and $Y_{1\wedge\tau} = Y_1$ because regardless of the coin toss outcome, we will always hold onto our option until at least time $t = 1$. $2 \wedge \tau$ will depend on the coin tossing:

$$Y_{2\wedge\tau}(HH) = Y_2(HH) = 0, \quad Y_{2\wedge\tau}(HT) = Y_2(HT) = 0.64,$$
$$Y_{2\wedge\tau}(TH) = Y_1(T) = 2.40, \quad Y_{2\wedge\tau}(TT) = Y_1(T) = 2.40$$

Note that even if the stopping time is less than 2, the process continues until $t = 2$. The value of the process merely freezes upon reaching the stopping time. Therefore, we may illustrate a stopped process using a binomial model, albeit slightly modified from what we are used to since $Y_{2\wedge\tau}(TH)$ does not equal $Y_{2\wedge\tau}(HT)$.



A stopped process.

This process is a martingale (which is also a supermartingale by definition). This fact is true for all stopped price processes of an American put option under the risk-neutral probability measure $\tilde{\mathbb{P}}$. Additionally, $\mathbb{E}Y_{n\wedge\tau} \geq \mathbb{E}Y_n$.

Theorem 3.3.2 (Optional Sampling Part I)**.** *A martingale stopped at a stopping time is a martingale. A supermartingale (or submartingale) stopped at a stopping time is a supermartingale (or submartingale, respectively).*

Theorem 3.3.3 (Optional Sampling Part II)**.** *Let* X_n*,* $n = 0, 1, \ldots, N$ *be a supermartingale and let* τ *be a stopping time. Then* $\mathbb{E}X_{n\wedge\tau} \geq \mathbb{E}X_n$*. If* X_n *is a submartingale, then* $\mathbb{E}X_{n\wedge\tau} \leq \mathbb{E}X_n$*. If* X_n *is a martingale, then* $\mathbb{E}X_{n\wedge\tau} = \mathbb{E}X_n$*.*

This does not hold for all exercise times, such as ρ.

3.4 Path-Dependent

Using our newfound knowledge of stopping times, we may now work out how to price American derivative securities that are permitted to be path-dependent. First, we must redefine the price process V_n to include stopping times. We still have an N-period binomial model with u and d as the up-factor and down-factor respectively, along with interest rate r such that $0 < d < 1 + r < u$. Let us define P_n as the set of all stopping times τ taking on values in the set $n, n+1, \ldots, N, \infty$. Then, P_0 would be the set of all stopping times and P_N can take on only the values N and ∞. We also let G_n be the intrinsic value process of the derivative security. Obviously, we need to have the intrinsic value be a random variable unlike the path-independent equivalent $g(s)$ which only depended on the stock price S_n at time n.

Definition 3.4.1. *For each n, $0 \leq n \leq N$, let G_n be a random variable depending on the first n coin tosses. An* **American derivative security with intrinsic value process** *G_n is a contract that can be exercised at any time prior to or at time N. If exercised at time n, its payoff is G_n. We define the* **price process** *V_n by the* **American risk-neutral pricing formula***:*

$$V_n = \max_{\tau \epsilon P_n} \tilde{\mathbb{E}}_n \Big[\mathbb{I}_{\{\tau \leq N\}} \frac{1}{(1+r)^{\tau - n}} G_\tau \Big], \qquad n = 0, 1, \ldots, N \tag{3.13}$$

In the above definition, $\mathbb{I}_{\{\tau \leq N\}}$ tells us that $\mathbb{I}_{\{\tau \leq N\}} \frac{1}{(1+r)^{\tau-n}} G_\tau = 0$ when $\tau = \infty$, and otherwise it simply equals $\frac{1}{(1+r)^{\tau-n}} G_\tau$. To get a clearer understanding of this definition, suppose we are at time n. The owner can choose whether or not to exercise the derivative based on only the previous path (i.e. not based on a later date). This means the exercise date *is* indeed a stopping time τ. Since the owner has not yet exercised the security, the stopping time must be in the set P_n. If it is never exercised, the owner would receive no money which is why we use the notation $\mathbb{I}_{\{\tau \leq N\}}$. Otherwise, the owner should choose $\tau \epsilon P_n$ that makes the expected value of $\frac{1}{(1+r)^{\tau-n}} G_\tau$ as large as possible. We will now develop some properties of the American derivative security price process as defined in Definition 3.4.1.

Theorem 3.4.2. *The American derivative security price process given by Definition 3.4.1 has the following properties:*

(i) $V_n \geq max\{G_n, 0\}$ for all n

(ii) The discounted process $\frac{1}{(1+r)^n} V_n$ is a supermartingale

(iii) If Y_n is another process satisfying $Y_n \geq max\{G_n, 0\}$ for all n and for which $\frac{1}{(1+r)^n} Y_n$ is a supermartingale, then $Y_n \geq V_n$ for all n.

Proof. (i) Let n be given. Consider the stopping time $\hat{\tau}$ in P_n that only takes on the value n. This means $\tilde{\mathbb{E}}_n \Big[\mathbb{I}_{\{\hat{\tau} \leq N\}} \frac{1}{(1+r)^{\hat{\tau}-n}} G_{\hat{\tau}} \Big] = \tilde{\mathbb{E}}_n \Big[\frac{1}{(1+r)^{n-n}} G_n \Big] = G_n$. Since V_n is the maximum value of $\tilde{\mathbb{E}}_n \Big[\mathbb{I}_{\{\tau \leq N\}} \frac{1}{(1+r)^{\tau-n}} G_\tau \Big]$, it must be at least equal to G_n. But, we also know that if the security is left to expire, the owner will receive no money for it and we will have $V_n = 0$. Since V_n is at least equal to G_n and at least equal to 0, we have $V_n \geq max\{G_n, 0\}$.

(ii) Let n be given. Let τ^* attain the maximum in the definition of V_{n+1}, i.e. $V_{n+1} = \tilde{\mathbb{E}}_{n+1} \Big[\mathbb{I}_{\{\tau^* \leq N\}} \frac{1}{(1+r)^{\tau^* - (n+1)}} G_{\tau^*} \Big]$. But $\tau \epsilon P_n$. Using this fact as well as the iterated conditioning property of conditional expectation, we see that

$$\begin{aligned}
V_n &\geq \tilde{\mathbb{E}}_n\left[\mathbb{I}_{\{\tau^*\leq N\}}\frac{1}{(1+r)^{\tau^*-n}}G_{\tau^*}\right] \\
&= \tilde{\mathbb{E}}_n\left[\tilde{\mathbb{E}}_{n+1}\left[\mathbb{I}_{\{\tau^*\leq N\}}\frac{1}{(1+r)^{\tau^*-n}}G_{\tau^*}\right]\right] \\
&= \tilde{\mathbb{E}}_n\left[\frac{1}{1+r}\tilde{\mathbb{E}}_{n+1}\left[\mathbb{I}_{\{\tau^*\leq N\}}\frac{1}{(1+r)^{\tau^*-n-1}}G_{\tau^*}\right]\right] \\
&= \tilde{\mathbb{E}}_n\left[\frac{1}{1+r}V_{n+1}\right]
\end{aligned}$$

Finally we divide both sides by $(1+r)^n$ to see that $\frac{V_n}{(1+r)^n} \geq \frac{1}{(1+r)^{n+1}}V_{n+1}$.

(iii) To prove the last part of the theorem, let Y_n be a process satisfying conditions (i) and (ii). Let $n \geq N$ be given and τ be a stopping time in P_n. Because of condition (i), we have

$$\begin{aligned}
\mathbb{I}_{\{\tau\leq N\}}G_\tau &\leq \mathbb{I}_{\{\tau\leq N\}}max\{G_\tau, 0\} \\
&\leq \mathbb{I}_{\{\tau\leq N\}}max\{G_{N\wedge\tau}, 0\} + \mathbb{I}_{\{\tau=\infty\}}max\{G_{N\wedge\tau}, 0\} \\
&= max\{G_{N\wedge\tau}, 0\} \\
&\leq Y_{N\wedge\tau}
\end{aligned}$$

Next, we can use Theorem 3.3.2 along with condition (ii):

$$\begin{aligned}
\tilde{\mathbb{E}}_n\left[\mathbb{I}_{\{\tau\leq N\}}\frac{1}{(1+r)^{\tau}}G_\tau\right] &= \tilde{\mathbb{E}}_n\left[\mathbb{I}_{\{\tau\leq N\}}\frac{1}{(1+r)^{\tau\wedge N}}G_\tau\right] \\
&\leq \tilde{\mathbb{E}}_n\left[\frac{1}{(1+r)^{\tau\wedge N}}Y_{\tau\wedge N}\right] \\
&\leq \frac{1}{(1+r)^{\tau\wedge n}}Y_{\tau\wedge n} \\
&= \frac{1}{(1+r)^n}Y_n
\end{aligned}$$

with the last step due to the fact that $\tau\epsilon P_n$ is greater than or equal to n on every possible path. If we multiply both sides by $(1+r)^n$, we get the equation

$$\tilde{\mathbb{E}}_n\left[\mathbb{I}_{\{\tau\leq N\}}\frac{1}{(1+r)^{\tau-n}}G_\tau\right] \leq Y_n \tag{3.14}$$

And, since V_n is the maximum value of the left side of equation (3.14) for all $\tau\epsilon P_n$, it must be that $V_n \leq Y_n$.

□

Each part of Theorem 3.4.2 tells us something about building a replicating portfolio of a path-dependent American derivative security. Property (ii) tells us that it is possible to construct a portfolio whose agent starts with initial capital V_0 and the value of the portfolio at each time $t = n$ is V_n. Property (i) tells us that if the agent does this, they will have hedged a short position in the security, no matter the exercise time. Therefore, properties (i) and (ii) together guarantee that the price of the derivative security is acceptable to the seller. In addition, property (iii) shows that the price is the minimum required to be acceptable for the seller, and thus is a fair price for the buyer.

We will now provide a theorem for and prove the American pricing algorithm for path-dependent derivative securities.

Theorem 3.4.3. *We have the following American pricing algorithm for the path-dependent derivative security price process given by Definition 3.4.1:*

$$V_N(\omega_1 \dots \omega_N) = max\{G_N(\omega_1 \dots \omega_N), 0\}, \tag{3.15}$$

$$V_n(\omega_1 \dots \omega_n) = max\Big\{G_n(\omega_1 \dots \omega_n), \frac{1}{1+r}\Big[\tilde{p}V_{n+1}(\omega_1 \dots \omega_n H) + \tilde{q}V_{n+1}(\omega_1 \dots \omega_n T)\Big]\Big\}, \tag{3.16}$$

for $n = N-1, \dots, 0$.

Proof. In this proof we will show that this process satisfies all three properties of Theorem 3.4.2 and is therefore the same American risk-neutral pricing formula as defined in Definition 3.4.1.

(i) It is clear that V_N satisfies property (i) of Theorem 3.4.2. Now, we use backwards induction. Suppose for some n between 0 and $N-1$, we have $V_{n+1} \geq max\{G_{n+1}, 0\}$. From equation (3.16), it follows that $V_n(\omega_1 \dots \omega_n) \geq max\{G_n(\omega_1 \dots \omega_n), 0\}$.

(ii) From equation (3.16), it follows that

$$\begin{aligned} V_n(\omega_1 \dots \omega_n) &\geq \frac{1}{1+r}\Big[\tilde{p}V_{n+1}(\omega_1 \dots \omega_n H) + \tilde{q}V_{n+1}(\omega_1 \dots \omega_n T)\Big] \\ &= \tilde{\mathbb{E}}_n\Big[\frac{1}{1+r}V_{n+1}\Big](\omega_1 \dots \omega_n) \end{aligned}$$

Multiplying both sides of this equation by $\frac{1}{(1+r)^n}$ will lead to property (ii) being satisfied.

(iii) It is clear that V_N is the smallest random variable satisfying $V_N \geq max\{G_N, 0\}$ since $V_N = max\{G_N, 0\}$. We must therefore show by backwards induction that the same is true for all n. Suppose for some n between 0 and $N-1$, V_n is the smallest random variable satisfying $V_{n+1} \geq max\{G_{n+1}, 0\}$. We know from (ii) that $V_n(\omega_1 \dots \omega_n) \geq \frac{1}{1+r}\Big[\tilde{p}V_{n+1}(\omega_1 \dots \omega_n H) + \tilde{q}V_{n+1}(\omega_1 \dots \omega_n T)\Big]$ because this process is a supermartingale. We also know from (i) that $V_n \geq G_n$. Combining these two facts, we get

$$V_n(\omega_1 \dots \omega_n) \geq max\Big\{G_n(\omega_1 \dots \omega_n), \frac{1}{1+r}\Big[\tilde{p}V_{n+1}(\omega_1 \dots \omega_n H) + \tilde{q}V_{n+1}(\omega_1 \dots \omega_n T)\Big]\Big\} \tag{3.17}$$

But, equation (3.16) tells us that both sides of this inequality are equal. So, $V_n(\omega_1 \dots \omega_n)$ is as small as possible.

□

Finally, we will prove the process for replicating path-dependent American derivatives in the stock and money markets. The proof is another induction on n, so we will skip it.

Theorem 3.4.4. *Consider an N-period binomial asset-pricing model with* $0 < d < 1 + r < u$ *and with*

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

For each n*,* $0 \leq n \leq N$*, let* G_n *be a random variable depending on the first* n *coin tosses. With* V_n*,* $0 \leq n \leq N$ *given by Definition 3.4.1, we define*

$$\Delta_n(\omega_1 \dots \omega_n) = \frac{V_{n+1}(\omega_1 \dots \omega_n H) - V_{n+1}(\omega_1 \dots \omega_n T)}{S_{n+1}(\omega_1 \dots \omega_n H) - S_{n+1}(\omega_1 \dots \omega_n T)}, \tag{3.18}$$

$$C_n(\omega_1 \dots \omega_n) = V_n(\omega_1 \dots \omega_n) - \frac{1}{1+r}\Big[\tilde{p}V_{n+1}(\omega_1 \dots \omega_n H) + \tilde{q}V_{n+1}(\omega_1 \dots \omega_n T)\Big], \tag{3.19}$$

where $0 \leq n \leq N-1$. *We have* $C_n \geq 0$ *for all* n. *If we set* $X_0 = V_0$ *and define recursively forward the portfolio values* $X_1, X_2, \ldots X_N$ *by*

$$X_{n+1} = \Delta_n S_{n+1} + (1+r)(X_n - C_n - \Delta_n S_n) \tag{3.20}$$

then we will have $X_n(\omega_1 \ldots \omega_n) = V_n(\omega_1 \ldots \omega_n)$ *for all* n *and all* $\omega_1 \ldots \omega_n$. *In particular, we have* $X_n \geq G_n$ *for all* n.

This theorem shows gives us an algorithm that is acceptable for the seller. Now, we will consider the buyer. We must first establish another theorem regarding a series of payments.

Theorem 3.4.5 (Cash flow valuation)**.** *Consider an N-period binomial asset-pricing model with 0 < d < 1 + r < u and risk-neutral probability measure* $\tilde{\mathbb{P}}$. *Let* $C_0, C_1, \ldots, C_N$ *be a sequence of random variables such that each* C_n *depends only on* $\omega_1 \ldots \omega_n$. *The price at time* n *of the derivative security that makes payments* $C_n, \ldots, C_N$ *at times* $n, \ldots, N$, *respectively, is*

$$V_n = \tilde{\mathbb{E}}_n\left[\sum_{k=n}^{N} \frac{C_k}{(1+r)^{k-n}}\right], n = 0, 1, \ldots, N \tag{3.21}$$

The price process V_n, $n = 0, 1, \ldots, N$ *satisfies Equation (3.19) and we define* Δ_n *and* X_{n+1} *by Equations (3.18) and (3.20), respectively, from Theorem 3.4.4.*

We can imagine we are at a fixed time n and that the security has not yet been exercised. Denote $\tau^* \epsilon P_n$ as the stopping time that attains the maximum V_n, i.e.

$$V_n = \tilde{\mathbb{E}}_n\Big[\mathbb{I}_{\{\tau^* \leq N\}} \frac{1}{(1+r)^{\tau^*-n}} G_{\tau^*}\Big] \tag{3.22}$$

Also define

$$C_k = \mathbb{I}_{\{\tau^* \leq k\}} G_k, \qquad k = n, n+1, \ldots, N \tag{3.23}$$

as the cash flows that the owner receives if the derivative security is exercised according to τ^*. At most only one C_k will be nonzero, and the k will correspond with the exercise time if the security is exercised at all. But, by our definition of C_k, we have

$$V_n = \tilde{\mathbb{E}}_n\Big[\sum_{k=n}^{N} \mathbb{I}_{\{\tau^*=k\}} \frac{1}{(1+r)^{k-n}} G_k\Big] = \tilde{\mathbb{E}}_n\Big[\sum_{k=n}^{N} \frac{C_k}{(1+r)^{k-n}}\Big] \tag{3.24}$$

And, this is just the cash flows $C_n, \ldots, C_N$ received at time $n, \ldots, N$, respectively. Therefore, V_n is acceptable to the seller as well.

Now, we only need to define a method by which to determine an optimal exercise time. We let $n = 0$ so we can consider all possible stopping times from the time the security is obtained.

Theorem 3.4.6 (Optimal exercise)**.** *The stopping time*

$$\tau^* = min\{n \mid V_n = G_n\} \tag{3.25}$$

maximizes the right hand side of Equation (3.13) when $n = 0$*; i.e.*

$$V_0 = \tilde{\mathbb{E}}\Big[\mathbb{I}_{\{\tau \leq N\}} \frac{1}{(1+r)^{\tau-n}} G_\tau\Big] \tag{3.26}$$

Proof. First, we need to show that the stopped process

$$\frac{1}{(1+r)^{n\wedge\tau^*}} V_{n\wedge\tau^*}$$

is a martingale. Let the first n coin tosses result in the sequence $\omega_1 \ldots \omega_n$

(i) If, along this path, $\tau^* \geq n+1$, $V_n(\omega_1 \dots \omega_n) \geq G_n(\omega_1 \dots \omega_n)$ from our assumption of τ^*. Further, as a consequence of Equation (3.16),

$$\begin{aligned} V_{n\wedge\tau^*}(\omega_1 \dots \omega_n) &= V_n(\omega_1 \dots \omega_n) \\ &= \frac{1}{1+r}\Big[\tilde{p}V_{n+1}(\omega_1 \dots \omega_n H) + \tilde{q}V_{n+1}(\omega_1 \dots \omega_n T)\Big] \\ &= \frac{1}{1+r}\Big[\tilde{p}V_{(n+1)\wedge\tau^*}(\omega_1 \dots \omega_n H) + \tilde{q}V_{(n+1)\wedge\tau^*}(\omega_1 \dots \omega_n T)\Big] \end{aligned}$$

And thus the martingale property is satisfied.

(ii) If, along this path, $\tau^* \leq n$,

$$\begin{aligned} V_{n\wedge\tau^*}(\omega_1 \dots \omega_n) &= V_{\tau^*}(\omega_1 \dots \omega_n) \\ &= \tilde{p}V_{\tau^*}(\omega_1 \dots \omega_{\tau^*}) + \tilde{q}V_{\tau^*}(\omega_1 \dots \omega_{\tau^*}) \\ &= \tilde{p}V_{(n+1)\wedge\tau^*}(\omega_1 \dots \omega_n H) + \tilde{q}V_{(n+1)\wedge\tau^*}(\omega_1 \dots \omega_n T) \end{aligned}$$

And again, the martingale property is satisfied.

Therefore, we can say

$$\begin{aligned} V_0 &= \tilde{\mathbb{E}}\Big[\frac{1}{(1+r)^{N\wedge\tau^*}}V_{N\wedge\tau^*}\Big] \\ &= \tilde{\mathbb{E}}\Big[\mathbb{I}_{\{\tau^*\leq N\}}\frac{1}{(1+r)^{\tau^*}}G_{\tau^*}\Big] + \tilde{\mathbb{E}}\Big[\mathbb{I}_{\{\tau^*=\infty\}}\frac{1}{(1+r)^{N}}G_N\Big] \end{aligned}$$

But if $\tau^* = \infty$ then it must be that $V_n > G_n$ for all n, including for $n = N$. But, because of equation (3.15), it must be true that $G_N < 0$ and $V_N = 0$. So, $\mathbb{I}_{\{\tau^*=\infty\}}V_N = 0$ and we can simplify the equation for V_0:

$$V_0 = \tilde{\mathbb{E}}\Big[\mathbb{I}_{\{\tau^*\leq N\}}\frac{1}{(1+r)^{\tau^*}}G_{\tau^*}\Big]$$

Which is exactly Equation (3.26). □

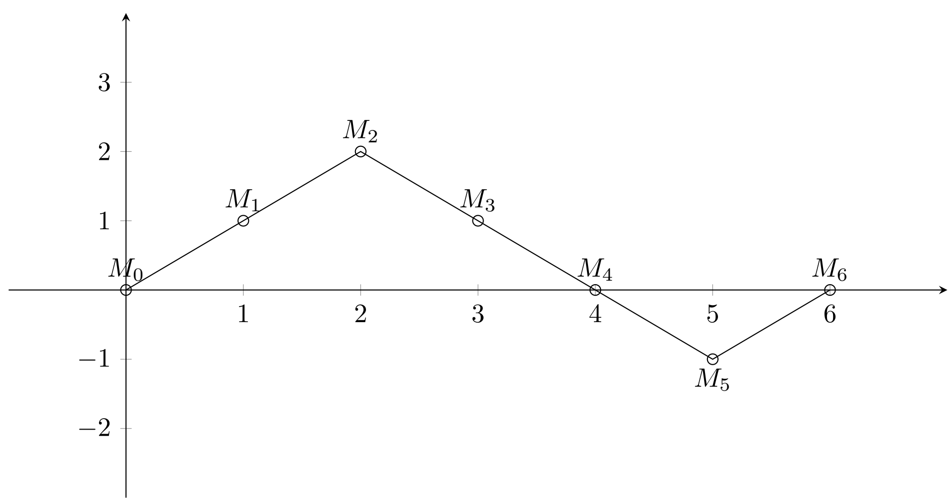
We have now seen how the binomial model can be used to price multiple different derivative securities, including European and American options. There are, of course, different means by which one may price such securities. The next chapter will give a brief overview of one such way: the Black-Scholes model.

Chapter 4

The Black-Scholes Model

4.1 Random Walk

The Black-Scholes model relies on Brownian motion. Though we will not explore this concept, we will develop its discrete-time counterpart known as the random walk. We will focus on the *symmetric* random walk wherein we repeatedly toss a fair coin ($p = q = \frac{1}{2}$). If we chose probabilities other than this, the resulting walk would be asymmetric. Let the process M_n be our random walk and $M_0 = 0$. If $\omega_1 \ldots \omega_n$ is our sequence of coin tosses, then for each ω_k that is heads, $M_k = M_{k-1} + 1$ and for each ω_k that is tails, $M_k = M_{k-1} - 1$. M_n is both a martingale and a Markov process.



A six-step random walk.

We can let τ_m be the first time the random walk reaches some integer m, i.e.

$$\tau_m = min\{n \,|M_n = m\} \tag{4.1}$$

We call τ_m the *first passage time* of the random walk to level m. If the walk never reaches level m, $\tau_m = \infty$. We would like to figure out the distribution of τ_m and we will do so by using the reflection principle.

Let's say we toss a coin an odd number of times. We can represent this number as $(2j-1)$ for some positive integer j. Out of all the possible paths that could result from different sequences of this coin toss, some will reach level 1 and others will not. For example, if we have a series of three coin tosses, there are 8 possible paths total. Five of these will reach level 1 (resulting from coin toss sequences HHH, HHT, HTH, THH, HTT).

Consider one of the paths that reach level 1. This means the first passage time would be no more than the total number of coin tosses; i.e. $\tau_1 \leq 2j-1$. We can create a *reflected path* that diverges from the original path at step τ_1 and mirrors all future steps. If the original path ends above 1, the reflected path ends below 1. If the original path ends at 1, then so does the reflected path. But, this covers all possible paths. And since there are the same amount of paths ending above 1 as there are reflected paths ending below 1, we can say:

$$\mathbb{P}(\tau_1 \leq 2j-1) = \mathbb{P}(M_{2j-1} = 1) + 2\mathbb{P}(M_{2j-1} \geq 3)$$

Note that since we have an odd number of coin tosses, it is impossible for the path to end on level 2, therefore in order to have ended above level 1 it must be at least at level 3. But, since there is an equal probability of getting heads as there is tails,

$$\mathbb{P}(M_{2j-1} \geq 3) = \mathbb{P}(M_{2j-1} \leq -3)$$

Which means

$$\mathbb{P}(\tau_1 \leq 2j-1) = \mathbb{P}(M_{2j-1} = 1) + \mathbb{P}(M_{2j-1} \geq 3) + \mathbb{P}(M_{2j-1} \leq -3)$$

This covers all cases except for when $M_{2j-1} = -1$. So,

$$\mathbb{P}(\tau_1 \leq 2j-1) = 1 - \mathbb{P}(M_{2j-1} = -1)$$

If $M_{2j-1} = -1$, the corresponding coin toss sequence contains $(j-1)$ heads and j tails. The number of possible paths whose coin toss fits this criteria is:

$$\binom{2j-1}{j} = \frac{(2j-1)!}{j!(j-1)!}$$

and each of these coin toss sequences have a probability of $\left(\frac{1}{2}\right)^{2j-1}$. Therefore,

$$\mathbb{P}(M_{2j-1} = -1) = \left(\frac{1}{2}\right)^{2j-1} \frac{(2j-1)!}{j!(j-1)!}$$

We come to similar conclusions for $(2j-3)$ coin tosses. In addition, it is clear that for $j \geq 2$,

$$\begin{aligned}
\mathbb{P}(\tau_1 = 2j-1) &= \mathbb{P}(\tau_1 \leq 2j-1) - \mathbb{P}(\tau_1 \leq 2j-3) \\
&= [1 - \mathbb{P}(M_{2j-1} = -1)] - [1 - \mathbb{P}(M_{2j-3} = -1)] \\
&= \mathbb{P}(M_{2j-3} = -1) - \mathbb{P}(M_{2j-1} = -1) \\
&= \left(\frac{1}{2}\right)^{2j-3} \frac{(2j-3)!}{(j-1)!(j-2)!} - \left(\frac{1}{2}\right)^{2j-1} \frac{(2j-1)!}{j!(j-1)!} \\
&= \left(\frac{1}{2}\right)^{2j-1} \frac{(2j-3)!}{j!(j-1)!} \Big[4j(j-1) - 2(j-1)(2j-2)\Big] \\
&= \left(\frac{1}{2}\right)^{2j-1} \frac{(2j-3)!}{j!(j-1)!} \Big[2j(2j-2) - 2(j-1)(2j-2)\Big] \\
&= \left(\frac{1}{2}\right)^{2j-1} \frac{(2j-3)!}{j!(j-1)!} (2j-2) \\
&= \left(\frac{1}{2}\right)^{2j-1} \frac{(2j-2)!}{j!(j-1)!}
\end{aligned}$$

We can state this as a theorem.

Theorem 4.1.1. *Let τ_1 be the first passage time to level 1 of a symmetric random walk. Then,*

$$\mathbb{P}(\tau_1 = 2j-1) = \left(\frac{1}{2}\right)^{2j-1} \frac{(2j-2)!}{j!(j-1)!} \qquad j = 1, 2, \ldots \tag{4.2}$$

4.2 The Black-Scholes Model

Another popular method for pricing derivative securities is Black-Scholes. The stock price model of Black-Scholes can be thought of as a natural extension of the binomial model, since it is based on a continuous-time model and the binomial model is discrete. It is as if we divide the periods of the binomial model into smaller and smaller parts, approaching a period length of zero. This will give us a model with continuous prices. The following theorem helps to estimate the distribution of stock prices as an assumption of the Black-Scholes model.

Theorem 4.2.1 (Azuma's Inequality)**.** *Let* $0 = X_0, \ldots, X_m$ *be a martingale with* $|X_{i+1} - X_i| \leq 1$ *for all* $0 \leq i \leq m-1$*. Let* $\lambda > 0$ *be arbitrary. Then, the probability that* $X_m > \lambda\sqrt{m}$ *is less than* $e^{-\frac{\lambda^2}{2}}$*.*

By Theorem 2.3.1, we know that the discounted stock price is a martingale. Black-Scholes instead takes into account the logarithm of stock prices, and this process is also a martingale. Azuma's Inequality states that, after many periods, the stock prices in the middle of our binomial model have a much higher probability than those at the very top and very bottom. And, in fact, the logarithm of the stock price process follows a normal distribution. This means our stock prices have a *log normal distribution* with mean μ and standard deviation σ.

This also allows for the Black-Scholes model to follow Brownian motion in estimating stock prices, which assumes that (1) the returns of a stock are normally distributed and (2) the standard deviation of these returns can be determined from historical data. In this scenario, we equate a stock's standard deviation to its *volatility*. The less volatile a stock is, the greater our ability will be to predict future prices.

The Black-Scholes formula uses z-scores of a stock price's log normal distribution. A z-score can be defined as follows:

Definition 4.2.2. *Consider a standard normal distribution, define* μ *to be the mean,* σ *to be the standard deviation, and* x *to be an observed value. Then, the probability of* x *can be defined in terms of its z-score:*

$$z = \frac{x - \mu}{\sigma} \tag{4.3}$$

which measures how far away x *is from the mean in terms of standard deviations.*

When pricing a call option, Black-Scholes looks for probabilities that are smaller than the stock's z-score.

Acknowledgements

Researching for and writing this thesis has taken up a good part of my senior year as an undergraduate at the University of Connecticut, and I definitely could not have completed it without the help of several people.

First and foremost, I would like to express my gratitude to my thesis advisor, Professor Alexander Russell, for guiding me to the research topic of financial derivative pricing and helping me every step of the way after that. Without his continual involvement and assistance I surely would not have been able to complete this thesis in the time frame that I did.

I would also like to thank my honors advisor, Professor Luke Rogers. He helped me navigate the requirements of the honors program for the last several years and was a great help in making sure I stayed on track to graduate with honors.

Lastly, I would like to thank all of my friends in the honors program. We mutually supported each other in completing our senior theses, especially at times where the task seemed particularly daunting and out of reach. We did it!

Bibliography

[Alon and Spencer, 2015] Alon, N. and Spencer, J. H. (2015). *The Probabilistic Method*. John Wiley & Sons, Inc., Hoboken, New Jersey, 4 edition.

[Shreve, 2004] Shreve, S. E. (2004). *Stochastic Calculus for Finance I The Binomial Asset Pricing Model*. Springer-Verlag, New York.

[Yalamova, 2010] Yalamova, R. (2010). Simple heuristic approach to introduction of the black-scholes model. *American Journal of Business Education*, 3(2):31–42.