Risk-Neutral Probabilities in a Binomial Tree Model

Masaru Okada

October 14, 2025

Contents

1 Introduction

The ultimate goal of this book is to explore the limits of arbitrage. To that end, we'll develop a mathematical framework that serves as a realistic model for actual financial markets and allows us to build pricing techniques.

2 The Binomial Branch Model

There are two minimum requirements for our purposes:

- A representation of the time value of money
- A random element to represent stocks

Without these two things, no model can be considered relevant to actual financial markets. So, we'll start by considering the simplest possible model, one consisting of bonds and stocks.

2.1 Stocks

Let's consider a very short period starting at time $t = 0$ and ending at $t = \delta t$. A variable representing a stock price should ideally have an unpredictable, random component. Over this short interval δt, we'll assume the stock price can only change in two ways: it can either go up or go down.

We'll introduce some structure to this randomness by assigning probabilities to the price going up or down. Let the probability of an upward movement over the short interval δt be p, and the probability of a downward movement be $(1 - p)$. The initial price at $t = 0$ is denoted as s_1. This is the price at which we can buy or sell any quantity of the stock without limit. We can then hold the stock for the duration of δt. Nothing happens while we hold the stock; that is, there is no cost to holding any quantity, whether positive or negative. However, after the elapsed time δt, the stock is assumed to have a new price. The price becomes s_2 if it goes down, and s_3 if it goes up.

2.2 Bonds

We also need to consider something that represents the time value of money, namely, a cash bond. Just as with the stock, we consider a very short period starting at $t = 0$ and ending at $t = \delta t$. We'll assume there is a continuous compound interest rate r such that one dollar at time $t = 0$ becomes $\exp(r\delta t)$ after the elapsed time δt. Any amount can be lent or borrowed at this interest rate. To represent this, we introduce a cash bond B, which can be traded at a price B_0 at time $t = 0$. The value of this bond becomes $B_0 \exp(r\delta t)$ after the elapsed time δt.

These two tools are all that our financial world contains. Although simplified, this world still presents uncertainty to investors. For a market participant, only one of the two possible stock prices after δt is favorable. (For example, if one is shorting the stock, a price drop is favorable.) Whether the outcome is favorable or not depends on a random result. An investor's demand for a payoff based on the future stock price can be expressed as a function f, which maps the two possible stock prices s_2 and s_3 at $t = \delta t$ to payoffs $f(2)$ and $f(3)$, respectively. For instance, a forward contract with a delivery price of k is represented by $f(2) = s_2 - k$ and $f(3) = s_3 - k$.

3 Risk-Free Construction

Let's consider what specific function f might be provided by a suitable strategy. As mentioned in Chapter 1, in a forward contract, you buy the stock at a price of s_1 at $t = 0$, sell cash bonds for the same amount s_1 to finance the purchase, and hold the position until maturity at δt. Therefore, the amount to be claimed should be $s_1 \exp(r\delta t)$. This leads to the forward contract price being

$$k = s_1 \exp(r\delta t).$$

This is price determination through arbitrage.

It may seem at first glance that we can't find a price-building strategy for a more complex f. After the very short time δt, the stock price takes on one of two random values, and similarly, derivatives generally take on a price different from their $t = 0$ price. If we knew the probabilities for each derivative payout, we could also find the expected price of f at maturity, which would be $(1 - p)f(2) + pf(3)$. But in reality, at time $t = 0$, we don't know the probability p of the price transitioning to its future state at time $t = \delta t$.

3.1 A Bond-Only Strategy

Let's consider a portfolio consisting only of cash bonds. Over the elapsed time δt, this cash bond increases by a factor of $\exp(r\delta t)$. This means that if you buy a cash bond at a price of $\exp(-r\delta t)[(1 - p)f(2) + pf(3)]$ at $t = 0$, its price will become $(1 - p)f(2) + pf(3)$ after δt. This price is chosen as the target because it is the expected price of the derivative.

Let S be a binomial process with an initial price s_1 at time $t = 0$, a downward value of s_2, and an upward value of s_3. Under the upward probability p, the expected value (or expected price) of S_1 at time 1, $E_p(S_1)$, is:

$$E_p(S_1) = (1 - p)s_2 + ps_3.$$

The contract f on S is a random variable, just like S_1. This means the expected value of the contract f is similarly defined:

$$E_p[f(1)] = (1 - p)f(2) + pf(3).$$

Also, the expected value discounted by the interest rate, under probability p, is

$$\exp(-r\delta t)[(1 - p)f(2) + pf(3)].$$

This is called the discounted expected value (or discounted expected price).

However, this price-building strategy is highly unlikely to be a good one. It's the law of large numbers from Chapter 1 appearing again in a different form, and just as before, it overlooks the price enforcement that comes from arbitrage. We already saw in Chapter 1 that this expectation cannot be applied to a forward contract, at least. The forward price is not expressed by $f(2), f(3)$ (i.e., s_2, s_3); rather, it's enforced by the interest rate r associated with the bond, resulting in $s_1 \exp(r\delta t)$.

3.2 Combining Bonds and Stocks

Let's think about a better combination. We'll use both bonds and stocks to construct a short-term portfolio. The tool that is more strongly linked to the performance of both stocks and derivatives than

a simple cash bond is, in fact, the stock itself.

As a general portfolio (ϕ, ψ), we assume we hold ϕ units of stock S (at a price of ϕs_1) and ψ units of cash bond B (at a price of ψB_0). If we purchase this portfolio at $t = 0$, the acquisition cost would be $\phi s_1 + \psi B_0$. However, at $t = \delta t$, this portfolio would take one of two possible values: Price of contract f if the stock price goes up:

$$f(3) = \phi s_3 + \psi B_0 \exp(r \delta t)$$

Price of contract f if the stock price goes down:

$$f(2) = \phi s_2 + \psi B_0 \exp(r \delta t)$$

Now, we have two possible prices and two free variables, ϕ and ψ. Under the condition that the stock moves appropriately, we have the two prices we want to match, $f(3)$ and $f(2)$, in hand. (Note: This is a problem of determining how much stock and cash (bonds) to hold if the values $s_2, s_3, f(2), f(3)$ are already known, and we want to find the values of ϕ and ψ.) If $s_3 \neq s_2$, we can transform the equations as follows:

$$\phi = \frac{f(3) - f(2)}{s_3 - s_2}$$

$$\psi = B_0^{-1} \exp(-r \delta t) \left\{ f(3) - s_3 \frac{f(3) - f(2)}{s_3 - s_2} \right\}$$

4 Fair Pricing

Any derivative f can be constructed from a suitable portfolio of bonds and stocks. This fact should have some influence on the contract, and indeed, the market recognizes this as a rational price, not an expected price. If we purchase this portfolio at $t = 0$, its acquisition cost is $V = \phi s_1 + \psi B_0$. Substituting the previous results, we get:

$$V = \frac{f(3) - f(2)}{s_3 - s_2} s_1 + \exp(-r \delta t) \left\{ f(3) - s_3 \frac{f(3) - f(2)}{s_3 - s_2} \right\}$$

Therefore, if we know the continuous compound interest rate r, the contract values $f(2), f(3)$, and the future stock prices s_2, s_3, we can replicate any derivative.

4.1 If a derivative is offered for sale at a price lower than V

Let's assume a derivative is offered for sale at a price P which is lower than this V. Another market participant can buy any quantity of the cheap derivative from them at price P. They can also sell a portfolio of (ϕ, ψ) for the same amount as the purchased derivative. After the short period δt has passed, no matter what the stock price is, the price P of the derivative can offset the value of the portfolio. This transaction generates a profit of $V - P$ for each unit of the purchased derivative and portfolio traded. This means anyone can earn a risk-free profit of any size. Therefore, P is not a rational price for market participants to accept, and the market will quickly correct to the correct price V in an effort to capture this free profit.

4.2 If a derivative is offered for sale at a price higher than V

Let's assume a derivative is offered for sale at a price P which is higher than this V. Another market participant can sell any quantity of the expensive derivative to them at price P. They can also buy a

portfolio of (ϕ, ψ) for the same amount as the sold derivative. After the short period δt has passed, no matter what the stock price is, the price P of the derivative can offset the value of the portfolio. This transaction generates a profit of $P - V$ for each unit of the sold derivative and portfolio traded. This means anyone can earn a risk-free profit of any size. Therefore, P is not a rational price for market participants to accept, and the market will quickly correct to the correct price V in an effort to capture this free profit.

The only way to prevent the counterparty from making a risk-free profit is to offer the price V. Thus, only V is the rational price for the derivative at time $t = 0$.

4.3 Example - The Big Picture in One Step

4.3.1 Problem

There's a non-interest-bearing bond and a stock, both at 1 dollar at $t = 0$. In the future, at $t = \delta t$, the stock price will be either 2 dollars or 0.5 dollars. What is the fair price for a bet that pays 1 dollar if the stock price goes up?

4.3.2 Solution

Let B be the bond price, S be the stock price, and X be the desired payoff for the bet. Referring to the expression for V from earlier, for this problem, we have $r = 0$, $f(3) = 1$, $f(2) = 0$, $s_3 = 2$, $s_2 = 0.5$, and $B_0 = 1$.

$$X = \frac{f(3) - f(2)}{s_3 - s_2} s_1 + \exp(-r\delta t) \left\{ f(3) - s_3 \frac{f(3) - f(2)}{s_3 - s_2} \right\}$$

Substituting the values, we find $X = 1/3$, so the answer to this problem is 0.33 dollars. Also, since

$$\phi = \frac{f(3) - f(2)}{s_3 - s_2}$$

$$\psi = B_0^{-1} \exp(-r\delta t) \left\{ f(3) - s_3 \frac{f(3) - f(2)}{s_3 - s_2} \right\}$$

substituting gives us $\phi = 2/3$ and $\psi = -1/3$. This means that to replicate this derivative (or "bet"), you should buy 2/3 units of the bond and sell 1/3 units of the stock. Indeed, at $t = 0$, the value is $2/3 \times 1 - 1/3 \times 1 = 1/3$ (dollars). If the stock price goes up at $t = \delta t$, the value becomes $2/3 \times 2 - 1/3 \times 1 = 1$ (dollar), and if the stock price goes down, the value becomes $2/3 \times 0.5 - 1/3 \times 1 = 0$, which are exactly the same payouts as the "bet" in this problem.

5 The Return of Expected Value

I've already stated that the approach based on the law of large numbers is futile, and that the expected price calculated using probabilities p and $(1 - p)$ would provide an arbitrage opportunity, except by sheer coincidence. Here, let's unexpectedly consider the following number q:

$$q = \frac{s_1 \exp(r\delta t) - s_2}{s_3 - s_2}$$

We'll assume $s_3 > s_2$ from now on. (This doesn't result in any loss of generality.)

5.1 The case where $q < 0$

If $q < 0$, then the inequality $s_1 \exp(r\delta t) < s_2 < s_3$ would hold. However, $s_1 \exp(r\delta t)$ is the price you get when you buy a bond worth s_1 at $t = 0$. Therefore, if $s_1 \exp(r\delta t) < s_2$, you could sell any quantity of bonds at $t = 0$ and use the proceeds to buy any quantity of stock at price s_1, thereby earning an infinite profit. In other words, under the assumption of no arbitrage, $q < 0$ cannot hold.

5.2 The case where $q > 1$

If $q > 1$, then the inequality $s_2 < s_3 < s_1 \exp(r\delta t)$ would hold. $s_1 \exp(r\delta t)$ is the price you get when you buy a bond worth s_1 at $t = 0$. Therefore, if $s_3 < s_1 \exp(r\delta t)$, you could buy any quantity of bonds at $t = 0$ and use the proceeds to sell any quantity of stock at price s_1, thereby earning an infinite profit. In other words, under the assumption of no arbitrage, $q > 1$ cannot hold.

From the above, in a rational market structure (i.e., under the no-arbitrage condition), the value q must satisfy $0 \le q \le 1$, so we can view q as some kind of probability.

Surprisingly, the price V of the (ϕ, ψ) portfolio,

$$V = \frac{f(3) - f(2)}{s_3 - s_2} s_1 + \exp(-r\delta t) \left\{ f(3) - s_3 \frac{f(3) - f(2)}{s_3 - s_2} \right\},$$

can be transformed using $q = \frac{s_1 \exp(r\delta t) - s_2}{s_3 - s_2}$ into the following expression:

$$V = \exp(-r\delta t) \left\{ (1 - q)f(2) + qf(3) \right\}$$

It's the very same thing, but when we express it using the expected value operator $E_p[f(1)] = \exp(-r\delta t) \left\{ (1 - p)f(2) + pf(3) \right\}$ that appeared earlier, we get:

$$V = E_q[f(1)]$$

In other words, the rational price V determined by the arbitrage condition of the (ϕ, ψ) portfolio was, in fact, the discounted expected price under probability q.

5.3 The payoff for a forward contract with a delivery price of k

The payoff for a forward contract with a delivery price of k is:

$$f(2) = s_2 - k$$
$$f(3) = s_3 - k$$

Let's find the correct delivery price k. In this transaction, we don't need to hold the portfolio at $t = 0$ (we borrow a bond and buy a stock at $t = 0$, then get money by delivering the stock to the counterparty at maturity, and finally repay the debt made at $t = 0$ with that money). Since $V = 0$, we have:

$$V = \frac{f(3) - f(2)}{s_3 - s_2} s_1 + \exp(-r\delta t) \left\{ f(3) - s_3 \frac{f(3) - f(2)}{s_3 - s_2} \right\}$$

$$= \frac{(s_3 - k) - (s_2 - k)}{s_3 - s_2} s_1 + \exp(-r\delta t) \left\{ (s_3 - k) - s_3 \frac{(s_3 - k) - (s_2 - k)}{s_3 - s_2} \right\}$$

$$= \frac{s_3 - s_2}{s_3 - s_2} s_1 + \exp(-r\delta t) \left\{ s_3 - k - s_3 \frac{s_3 - s_2}{s_3 - s_2} \right\} = s_1 + \exp(-r\delta t)(s_3 - k - s_3) = s_1 - k \exp(-r\delta t)$$

Therefore, when $V = 0$, we get $k = s_1 \exp(r\delta t)$.

6 Binomial Trees

We'll continue our discussion by moving from binomial branches to binomial trees. A single time step is easy to analyze, but it only presents the bare minimum of a model. In the binomial branch case, even with a cash bond and a random stock, the stock price after the time step was limited to two patterns. Let's see what we can say when we expand the discussion from binomial branches to binomial trees.

In the following, we'll look at the time evolution over multiple periods, not just one. We'll still assume we only have cash bonds and a random stock, with no transaction costs, no risk of default, and the ability to buy and sell infinitely without a bid-ask spread.

6.1 Stocks

Just as in the previous section, we'll assume the stock price S at $t = 0$ has a value of $S_0 = s_1$, and after a minimal time unit δt has elapsed, S_1 can only take on two values, s_2 or s_3. However, after $2\delta t$ has passed, s_2 can become $S_2 = s_4$ or s_5, and s_3 can become $S_2 = s_6$ or s_7. For a natural number i, the stock can take on 2^i values after time $i\delta t$ has elapsed. However, given the value at time $(i - 1)$, there are only two possibilities. The next value of s_j is either s_{2j} or $s_{(2j+1)}$. Also, if the probability of a price increase (i.e., s_j becoming $s_{(2j+1)}$) is p_j, then the probability of a price decrease (i.e., s_j becoming s_{2j}) is $1 - p_j$.

6.2 Cash Bonds

6.2.1 General Theory

We can acknowledge that there may be a random element to the performance of a cash bond. (However, we're not particularly interested in its exact shape right now.) In the case of a cash bond, the randomness is completely different from that of a stock. A cash bond has the same structure as the time value of money. The interest rate to be paid can only change with time. However, the value of the cash bond at the next instant is always known. This is because this value depends only on the interest rate, which is known at $t = 0$.

6.2.2 Assumption

Here, we'll introduce a constant interest rate r that is the same everywhere in the tree. This means the value of the cash bond at time $n\delta t$ is $B_0 \exp(rn\delta t)$.

6.3 The Tree is Complex

Our ultimate goal is to understand the limits of risk-free construction when the underlying asset takes on continuous values in continuous time. And in the limit as $\delta t \to 0$, our model aligns very well with this ultimate goal. Rather than giving up because the tree is simple, we should say that it's not so complex for our purpose of advancing the analysis.

6.4 Backward Induction

In fact, most of the groundwork has already been laid in our discussion of the one-period binomial branch. The key point of this multi-period binomial tree is backward induction, where we expand the replicating portfolio by going backward in time, from the final moment when the payoff occurs to the starting time when we want to find the price, in very short time steps. Just as in the one-period binomial branch discussion, we had a payoff function f at the final time (maturity). In a multi-period binomial tree, we can extend our thinking to include contracts that exist at intermediate points between the start time $t = 0$ and maturity.

The structure of the binomial tree is one where there is a one-to-one relationship between a 'node' and the 'path the stock takes to reach that node.' If you follow a different path to a certain node, you won't reach that node. This condition allows us to correspond 'a contract's payoff f' with 'a specific endpoint.' A general contract can be thought of as a function f on the nodes at the maturity of that contract.

6.5 Two-Period Example

The expected value operator $E_p[]$ was defined for a single branch (a one-period binomial branch). Here, we'll calculate for a two-period tree, which has three branches. We'll also assume the interest rate r is constant for all branches. From the one-period binomial branch discussion, there exists a suitable combination of q_j that gives the derivative price $f(j)$ at node j at time i:

$$f(j) = \exp(-r\delta t)(q_j f(2j + 1) + (1 - q_j)f(2j))$$

This is the discounted expected value of the contract prices $f(2j + 1)$ and $f(2j)$ at time $(i + 1)$ under the probability (measure) q_i. For $j = 3, 2$ respectively, we have:

$$f(3) = \exp(-r\delta t)(q_3 f(7) + (1 - q_3)f(6))$$

$$f(2) = \exp(-r\delta t)(q_2 f(5) + (1 - q_2)f(4))$$

Here, the probability q_j is:

$$q_j = \frac{s_j \exp(r\delta t) - s_{2j}}{s_{(2j+1)} - s_{2j}}$$

That is,

$$q_3 = \frac{s_3 \exp(r\delta t) - s_6}{s_7 - s_6}$$

$$q_2 = \frac{s_2 \exp(r\delta t) - s_4}{s_5 - s_4}$$

And since

$$f(1) = \exp(-r\delta t)(q_1 f(3) + (1 - q_1)f(2))$$

substituting everything gives us:

$$f(1) = \exp(-2r\delta t)\{q_1 q_3 f(7) + q_1(1 - q_3)f(6) + (1 - q_1)q_2 f(5) + (1 - q_1)(1 - q_2)f(4)\}$$

From this, we can conclude the following.

6.6 Expected Value in a Tree

The expected value of the claim f at the final nodes of the tree is the sum of the payoffs multiplied by the probabilities of each path.

In a two-period tree, there were four paths to the final nodes. Therefore, $f(1)$ is composed of four terms. The probabilities of the paths were $q_1 q_3$, $q_1(1-q_3)$, $(1-q_1)q_2$, and $(1-q_1)(1-q_2)$, respectively. The corresponding claims were the endpoints $f(7), f(6), f(5), f(4)$. The period is $2\delta t$, and the discount factor $\exp(-2r\delta t)$ is applied.

6.7 The Inductive Step

Consider one node at time $(n-1)$ and the two nodes at time n (a total of three nodes). From the results of the previous section, a risk-free portfolio (ϕ, ψ) consisting of stocks and bonds can replicate the claims at time n. And when we consider only these three nodes, it's indistinguishable from the one-period binomial branch model. (It's exactly the same as the one-period binomial branch model.)

6.8 The Result of Induction

By induction, by going backward through the tree (i.e., from contract maturity to $t = 0$), we can build (replicate) a portfolio (ϕ, ψ) for each branch. A single value will arrive at the root of the entire tree at $t = 0$. This is the price of the derivative at $t = 0$. Just as in the one-period binomial branch model, no matter what path the stock actually takes, we can build (replicate) a unique portfolio (ϕ, ψ) under the strong constraint of arbitrage. This portfolio (ϕ, ψ) changes with each time unit. This replicating portfolio gives not just one stock holding ϕ at one node, but a stock holding ϕ for every node. So, when the coin of fate is flipped (a coin that can only land heads or tails), and the stock moves on the tree, this holding (the amount that should be held) ϕ also changes. This replicating portfolio (ϕ, ψ) is random just like the stock, but the most important difference is that "unlike the stock, the amount of the replicating portfolio (ϕ, ψ) that can be traded at that exact moment is known one time unit in advance." (For example, the problem on page 28 of the previous section). All contracts can be replicated by a portfolio of stocks and bonds. And all contracts have a no-arbitrage price.

6.9 Expected Values, Again

Until now, we haven't needed the probabilities p_j (the probabilities of the stock price increasing, i.e., s_j becoming $s_{(2j+1)}$). It's not a coincidence that we didn't need p_j in the formulas; there's a solid logic behind it. We will now see that the conclusion of the previous section is equivalent to saying that the expected value operator, with respect to a suitable 'probability,' provides the correct local (time-dependent) hedge.

6.10 A Concrete Example

Let's consider a 3-period binomial tree where the probability of a stock price increase at each node is 3/4, and the probability of a decrease is 1/4, as shown in Figure 2.5. The problem is to find the price of an option that allows you to buy the stock for a price of 100 at time 3.

The prices at the final time (maturity) in Figure 2.5 are 160, 120, 80, and 40, respectively. Since we'll exercise the option if the price is greater than 100 and abandon it if it's less than 100, the maturity prices of the option in the binomial tree are $60(= 160 - 100)$, $20(= 120 - 100)$, 0, and 0.

Here, we need equations for a new probability q and a contract price f. They can be found respectively by:

$$q = \frac{s_{now} - s_{down}}{s_{up} - s_{down}}$$

$$f_{now} = q f_{up} + (1 - q) f_{down}$$

This new probability q is 1/2 at every node. For example, for the branch at time 2 where the stock price is 140,

$$q = \frac{140 - 120}{160 - 120} = 1/2$$

And for the branch at time 1 where the stock price is 80,

$$q = \frac{80 - 60}{100 - 60} = 1/2$$

The contract price, for example, for the node at time 2 where the stock price was 140, is:

$$f_{now} = (1/2) \times 60 + (1/2) \times 20 = 40$$

And for the branch at time 1 where the stock price was 80, is:

$$f_{now} = (1/2) \times 10 + (1/2) \times 0 = 5$$

By proceeding with these steps, we can find the risk-neutral probabilities and contract prices at all nodes.

6.10.1 On the practical use of q and f

Let's consider a specific hedging strategy. We'll calculate ϕ, which is defined as:

$$\phi = \frac{f_{up} - f_{down}}{s_{up} - s_{down}}$$

This is the amount of stock in the replicating portfolio at the beginning (at $t = 0$) in the binomial branch model.

- **Strategy at $t = 0$ (Node with stock price=100, option price=15)** From the previous procedure, the option price at time 0 was found to be 15. (This 'option price' has the same meaning as the 'game entry fee paid to the bookmaker' in the example, where we're the bookmaker.) First, the option has a value of 15. (So, the entry fee for this 'game' is 15.) Now, at this $t = 0$, $\phi = (25 - 5)/(120 - 80) = 1/2$. The stock price at $t = 0$ is 100, and buying 1/2 unit costs 50. This means that (initial option price) - (stock to be held) $= 15 - 50 = -35$, so we need to borrow 35 to hold $\phi(= 0.5)$ units of stock. Therefore, the borrowing at time 0 is 35.
- **Strategy at $t = 1$ (if we reach the node with stock price=120, option price=25)** The ϕ at this node at time 1 is $\phi = (40 - 10)/(140 - 100) = 3/4$. Since ϕ was 1/2 at $t = 0$, to match the new ϕ, we must increase our stock holding by $3/4 - 1/2 = 1/4$. The stock price at this node is 120, so we want to buy an additional $120 \times (1/4) = 30$ worth of stock to match the new ϕ. Therefore, to match the new ϕ, we need to borrow an additional 30. Since we had borrowed 35 at $t = 0$, our total borrowing at this node becomes 65.

- **Strategy at $t = 2$ (if we reach the node with stock price=140, option price=40)** The ϕ at this node at time 2 is $\phi = (60 - 20)/(160 - 120) = 1$. At the previous node, ϕ was $3/4$, so to match the new ϕ, we must increase our stock holding by $1 - 3/4 = 1/4$. The stock price at this node is 140, so we want to buy an additional $140 \times (1/4) = 35$ worth of stock to match the new ϕ. Therefore, to match the new ϕ, we need to borrow an additional 35. Since we had borrowed 65 at the previous node, our total borrowing at this node becomes 100.
- **Strategy at $t = 3$ (at maturity, if we reach the node with stock price=120, option price=20)** Since $120 > 100$ (maturity stock price ¿ option strike price), the option is exercised. We hand over $\phi = 1$ unit of stock. To repay the total of 100 that we borrowed at the previous step, we receive 100 in cash from the person to whom we handed over the stock. From the above, we see that the option price must be 15 at $t = 0$, otherwise the person selling the option would lose money, so 15 is a reasonable price.

That's how ϕ is used. Let's consider another path on the same tree, different from the one in the example above.

- **Strategy at $t = 0$ (Node with stock price=100, option price=15)** At $t = 0$, $\phi = (25-5)/(120-80) = 1/2$, and Option Price = Replicating Portfolio Bond + Replicating Portfolio Stock. Bond = Option - Stock = Option - Stock Holding ϕ * Stock Price = $15 - (1/2) \times 100 = -35$. This means we first need to borrow 35 at $t = 0$ to make $\phi = 1/2$. (This is exactly the same argument as in the previous example.)
- **Strategy at $t = 1$ (if we reach the node with stock price=80, option price=5)** $\phi = (10 - 0)/(100 - 60) = 1/4$. Bond = Option - Stock = $5 - 80 \times 1/4 = -15$. Our borrowing was 35 at the previous step, but at this node, we reduce our borrowing to 15. (In other words, we partially sell the stock we held for a price of 20 to reduce the borrowed amount to 15.)
- **Strategy at $t = 2$ (if we reach the node with stock price=100, option price=10)** $\phi = (20 - 0)/(120 - 80)$. Bond = Option - Stock = $10 - 100 \times (1/2) = -40$. Our borrowing was 15 at the previous step, but we borrow an additional 25 to increase our borrowing to 40. (This additional 25 in borrowing is spent on buying more stock. Since there are only bonds and stocks in the world, the bonds are exchanged for stocks. So if the amount of our bond assets decreases, the amount of our stock holdings increases.)
  (Why do we consider just 'stocks' and 'bonds'? Why not 'rocks', 'water', and 'rice', for example? This question can be split into two: first, why are there only two types? And second, why are the types limited to just 'stocks' and 'bonds'? → First, we're considering the simplest possible model. To exchange things, you only need a minimum number of types greater than one (i.e., two types). Furthermore, the initial derivative price is the contract payoff price reflecting the randomness of the underlying asset. Randomness, from the perspective of a bond, means the bond's price moves randomly. Conversely, from the perspective of a stock, the bond moves randomly. For example, the price of one stock changes from 100 yen in bonds to 110 yen on one day, and 105 yen on another. If you have one asset that moves randomly from the perspective of another asset, you can replicate a cash flow. You could have three types of things, but by not using one of them, you can reduce it to a two-asset model. So, while you could consider n types of goods where n is a natural number greater than or equal to 3, you wouldn't need to use $n - 2$

of them in your analysis.)

- **Strategy at** $t = 3$ **(at maturity, if we reach the node with stock price=80, option price=0)** The option is not exercised. (Stock price=80 ¡ Strike price=100.) Our stock holding from the previous step was $\phi = 1/2$, so the value of our stock assets is $80 \times (1/2) = 40$. Our borrowing was 40 at the previous step, so we can close the transaction with no profit or loss.

Here, too, we can use expected values. Under the risk-neutral probability q of an increase, the probabilities of reaching each final node are $1/8, 3/8, 3/8, 1/8$ for the four possible maturity prices. The expected value of f under this probability q becomes the option's no-arbitrage price of 15. However, the probabilities of reaching each final node under the original stock probabilities are $27/64, 27/64, 9/64, 1/64$ for the four possible maturity prices. The expected value of f under this probability p is 33.75.

The expected value of S under this probability p is $(160 \times 27 + 120 \times 27 + 80 \times 9 + 40 \times 1)/64 = 130$, which is a value that has never appeared in any previous calculations and is meaningless. It seems that only the combination of the expected value of f under probability q is meaningful.

7 The Binomial Martingale Representation Theorem

We can summarize our finding—that any derivative can be priced using a simple discounted expected value operator under a risk-neutral up-probability q_j that exists at a given node j—as a theorem. This theorem may seem formal, but it's useful when we look at continuous models. In continuous models, things can be counterintuitive. However, the expected value results are still applicable and make things simpler even in continuous models.

7.1 Definitions Based on Diagrams

We'll provide seven definitions for concepts we've been using implicitly.

7.2 Definition of the Stock Price Process, S

The possible stock price combinations and their interconnections shown at each node of a tree are called the stock price process, S. The random variable S_i represents the value of the stock price process at time i. The S_1 in Figure 2.9 is a random variable that can be either 60 or 120, depending on whether it's at node 2 or node 3.

7.3 Definition of Probability Measures P and Q

Separate from the stock price process, the set of probabilities (p_i) or (q_i) is called a probability measure P or a probability measure Q. This probability measure indicates the rate at which an upward or downward movement occurs at each node. In Figure 2.9, it's defined on the paths between the nodes.

"Probability of an increase" and "magnitude of an increase" are distinct concepts within the binomial tree, and they are independent. We didn't need the real-world probability measure P to find the probability measure that allows for a risk-free construction. The probability measure Q that enables

risk-free construction is a function of the stock price process S, and the real-world probability measure P was unrelated to it. The real-world probability measure P does not affect the pricing of derivatives. It's the magnitude of the stock price's increase (or decrease) and their interrelationships that affect the derivative price, not the probability (the pre-predicted probability (p_i)) of an increase.

The reason the law of large numbers didn't work was because we were focused on the stock price process S and the real-world probability measure P. For the purpose of arbitrage-free discussion, we should have discarded the real-world probability measure P and focused only on the stock price process S.

7.4   Definition of Filtration (F_i)

Filtration (F_i) represents the history of stock prices up to time i on the tree. The filtration F_0 at time 0 consists of a single node, node 1 (at time 0 it's a node, not a path). For example, F_1 is (1,2) if the stock price first moves downwards, and F_1 is (1,3) if it first moves upwards.

In a binomial tree, there is only one path to reach a given node. Therefore, filtration F_i can be associated with the node reached at time i. Filtration determines the history of a stock price's upward/downward movements. In other words, filtration represents the node. In a non-recombining binomial tree model like the one we're considering, having a given node is equivalent to having a given filtration.

7.5   Definition of a Contract X

A contract X is a function on the nodes at maturity T. Since the nodes in a non-recombining binomial tree model like the one we're considering correspond one-to-one with the filtrations, it can also be said to be a function on the filtration F_T. For example, the stock price S_2 at time 2 (maturity) is the contract $X = S_2$.

A contract X is a different concept from the stock price process S. The contract X is determined only by the nodes at maturity. The stock price process S is determined by all time points leading up to maturity.

7.6   Conditional Expectation Operator $E_Q(\cdot \,|F_i)$

The conditional expectation operator $E_Q(\cdot \,|F_i)$ is different from the traditional expectation $E(\cdot)$ in that it depends on the probability measure Q and the filtration F_i. When we used the traditional expectation $E(\cdot)$, for example, we were interested in taking the expectation along the entire path from time 0. The conditional expectation $E_Q(\cdot \,|F_i)$ is suited for the purpose of taking the expectation with a later time as the starting value.

The conditional expectation $E_Q(X|F_i)$ is the expected value of the contract X along future paths (filtrations) from F_i. The conditional expectation $E_Q(X|F_i)$ is itself a random variable. The value of $E_Q(\cdot \,|F_i)$ at each node at time i is the expected value of X if that node is reached.

7.7 Examples of Conditional Expectation

$$E_P(S_0|F_0 = (1)) = 80$$

At the time of filtration $F_0 = (1)$, the value of S_0 is fixed at 80, and the expected value is 80.

$$E_P(S_1|F_0 = (1)) = (120 + 60)/2$$

At the time of filtration $F_0 = (1)$, S_1 can only take two values (120,60), so we take the probability average of those two values.

$$E_P(S_1|F_1 = (1,3)) = 120$$

At filtration $F_1 = (1,3)$, the value of S_1 is fixed at 120.

$$E_P(S_1|F_1 = (1,2)) = 60$$

At filtration $F_1 = (1,2)$, the value of S_1 is fixed at 60.

$$E_P(S_2|F_0 = (1)) = (180 + 80 + 72 + 36) \times \frac{1}{2} \times \frac{1}{2} = 92$$

At filtration $F_0 = (1)$, S_2 can only take on four values (180,80,72,36) at maturity, so we take the probability average of those four values.

$$E_P(S_2|F_1 = (1,3)) = (180 + 80)/2 = 130$$

At filtration $F_1 = (1,3)$, S_2 can only take on two values (180,80) at maturity, so we take the probability average of those two values.

$$E_P(S_2|F_1 = (1,2)) = (72 + 36)/2 = 54$$

At filtration $F_1 = (1,2)$, S_2 can take on two values (72,36), so we take the probability average of those two values.

$$E_P(S_2|F_2 = (1,3,7)) = 180$$

At filtration $F_2 = (1,3,7)$, S_2 is already fixed at 180.

$$E_P(S_2|F_2 = (1,3,6)) = 80$$

At filtration $F_2 = (1,3,6)$, S_2 is already fixed at 80.

$$E_P(S_2|F_2 = (1,2,5)) = 72$$

At filtration $F_2 = (1,2,5)$, S_2 is already fixed at 72.

$$E_P(S_2|F_2 = (1,2,4)) = 36$$

At filtration $F_2 = (1,2,4)$, S_2 is already fixed at 36.

So, the conditional expectation at $t = 0$, $E_P(S_2|F_0 = (0))$, is the same value as the expectation we've used before, $E_P(S_2)$. However, the expectation under the filtration F_2 that has reached maturity, $E_P(S_2|F_2)$, is the value of S_2 itself.

Now, the expectation of a contract X under probability measure P at time i, $E_P(X|F_i)$, can itself be regarded as a process at time i (just like the stock price process S_2). In fact, as shown in Figure 2.11, the expectation process $E_P(S_2|F_i)$ is a different process from the stock price S_2. In this way, by determining some probability measure (for example, Q), a contract X can be transformed into a process with index i, $E_Q(X|F_i)$.

(Note: A contract only has a value at maturity, whereas a process has values at all nodes of the binomial tree.)

7.8   Predictable Process $\phi = \phi_i$

This is a process where the value at a node at time i depends only on the filtration F_{i-1} up to time $(i-1)$. The value at each node is defined for nodes after time 0. For example, the random bond price process B_i ($i \geq 1$) is $B_i = \phi_i = S_{i-1}$, and it's predictable. (Figure 2.12 is the lagged process of Figure 2.9. The value is fixed at each point in time.)

7.9   Definition of a Martingale

A process S is said to be a martingale with respect to a probability measure P and filtration F_i if $E_P(S_j|F_i) = S_i$ for all $i \leq j$. It is also sometimes simply called "S is a P-martingale under F_i". For a process S to be a martingale with respect to a probability measure P, the conditional expectation of the process S_j at time j, given the filtration F_i up to time i, is simply the process S_i at time i. In other words, if S is a P-martingale under F_i, the process S has no upward or downward bias under the probability measure P. If the process S_i is 100 at some point i, the P-conditional expectation is also 100.

7.10   Martingale Example 1

Trivially, a process that only takes on a constant value is a martingale with respect to any probability measure.

7.11   Martingale Example 2

The process S shown in Figure 2.9 is a martingale under the probability measure Q given in Figure 2.10b.

For example,
$$E_Q(S_1|F_0 = (0)) = (1/3) \times 120 + (2/3) \times 60 = 80$$

and since $S_0 = 80$, $E_Q(S_1|F_0 = (0)) = S_0$ holds. Furthermore,
$$E_Q(S_2|F_1 = (1,3)) = (2/5) \times 180 + (3/5) \times 80 = 120$$

and for this filtration, $S_1 = 120$. Therefore, $E_Q(S_2|F_1 = (1,3)) = S_1$ holds. To show it's a martingale, we must check all cases, starting with the case for $F_1 = (1,2)$.

7.12 Martingale Example 3

The conditional expectation process $N_i = E_P(S_2|F_i)$ is a P-martingale. (Check) We find N_1 according to the definition.

$$N_1 = E_P(S_2|F_1 = (1,3)) = 180 \times (1/2) + 80 \times (1/2) = 130$$

or,

$$N_1 = E_P(S_2|F_1 = (1,2)) = 72 \times (1/2) + 36 \times (1/2) = 54$$

Thus, we find that

$$E_P(N_1|F_0) = 130 \times (1/2) + 54 \times (1/2) = 92$$

On the other hand,

$$
\begin{aligned}
N_0 &= E_P(S_2|F_0) \\
&= 180 \times \frac{1}{4} + 80 \times \frac{1}{4} + 72 \times \frac{1}{4} + 36 \times \frac{1}{4} \\
&= 92
\end{aligned}
$$

Therefore, $E_P(N_1|F_0) = N_1$ holds, and N_i is a P-martingale.

(Why can we say that all N_i are a P-martingale just by showing $E_P(N_1|F_0) = N_1$? Ask someone.)

$\longrightarrow$ Kubo-san explained that this is simply the first step of an inductive proof (testing the case for $n = 1$, assuming it holds for $n = k$, and then proving it for $n = k + 1$), specifically the case for $n = 1$.

7.13 Conditional Expectation Process of a Contract (Tower Property)

For any contract X, the process $E_{\mathbf{P}}(X|F_i)$ is a $\mathbf{P}$-martingale. That is, if we set the process $N_i = E_{\mathbf{P}}(X|F_i)$, then for any $i \leq j$, $E_{\mathbf{P}}(N_j|F_i) = N_i$ holds. To summarize, $E_{\mathbf{P}}(E_{\mathbf{P}}(X|F_j)|F_i) = E_{\mathbf{P}}(X|F_i)$ holds. Interpreting the Tower property, it means that taking the expectation along the path up to time j and then taking the expectation along the path up to an earlier time i is equivalent to taking the expectation from the beginning along the path going back to time i.

Using the Tower property, we can easily check if a process is a $\mathbf{P}$-martingale. That is, S_i is a $\mathbf{P}$-martingale if and only if the process S_i itself is equal to the conditional expectation of the maturity T under the filtration F_i, $E_{\mathbf{P}}(S_T|F_i)$.

$$S_i = E_{\mathbf{P}}(S_T|F_i) \quad \Longleftrightarrow \quad S_i \text{ is a } \mathbf{P}\text{-martingale.}$$

It's important to note that a process that is a martingale with respect to one probability measure may not be a martingale with respect to a different probability measure.

7.14 Exercise 2.3

Prove that the process S_i is a $\mathbf{Q}$-martingale.
We just need to show that $E_{\mathbf{Q}}(S_2|F_i) = S_i$.

$$E_{\mathbf{Q}}(S_2|F_0)$$
$$= \frac{1}{3} \times \frac{2}{5} \times 180 + \frac{1}{3} \times \frac{3}{5} \times 80 + \frac{2}{3} \times \frac{2}{3} \times 72 + \frac{2}{3} \times \frac{1}{3} \times 36$$
$$= 24 + 16 + 32 + 8$$
$$= 80$$
$$= S_0$$

$$E_{\mathbf{Q}}(S_2|F_1 = (1,3)) = \frac{2}{5} \times 180 + \frac{3}{5} \times 80$$
$$= 72 + 48$$
$$= 120$$
$$= S_1 \quad (\text{for } F_1 = (1,3))$$

$$E_{\mathbf{Q}}(S_2|F_1 = (1,2)) = \frac{2}{3} \times 72 + \frac{1}{3} \times 36$$
$$= 48 + 12$$
$$= 60$$
$$= S_1 \quad (\text{for } F_1 = (1,2))$$

Since $S_2 = E_{\mathbf{Q}}(S_2|F_2) = E_{\mathbf{Q}}(S_T|F_T) = S_T$, we have confirmed that $E_{\mathbf{Q}}(S_2|F_i) = S_i$ for all $i(= 0, 1, 2)$. Hence, the process S_i is a $\mathbf{Q}$-martingale.

7.15   The Binomial Martingale Representation Theorem

Let the binomial processes S and N each be a $\mathbf{Q}$-martingale. Then, by setting $\Delta S_i = S_i - S_{i-1}$, there exists a predictable process ϕ such that

$$N_i = N_0 + \sum_{k=1}^{i} \phi_k \Delta S_k$$

7.16   Explanation

The increments of processes S and N from time $i-1$ to time i are $\Delta S_i = S_i - S_{i-1}$ and $\Delta N_i = N_i - N_{i-1}$, respectively. There are two possible values for the next time step. One process can be expressed using the scaling (increasing/decreasing the width) and shifting (matching the offset) of the other.

Let's start with scaling. The magnitude of the difference between the upward and downward jumps is $\delta s_i = s_{\text{up}} - s_{\text{down}}$ for S, and similarly, $\delta n_i = n_{\text{up}} - n_{\text{down}}$ for N. Both of these are determined only by the filtration F_{i-1}.

A process that is determined only by the filtration F_{i-1} is, by definition, a predictable process. So, we define the ratio of these processes by a predictable process ϕ_i:

$$\phi_i = \frac{\delta n_i}{\delta s_i}$$

Next, let's consider the shift. The respective increments $\Delta S_i, \Delta N_i$ can be made to match by multiplying by the scaling factor and adding an offset, k_i. That is,

$$\Delta N_i = \phi_i \Delta S_i + k_i$$

This k_i is determined by F_{i-1} just like ϕ (it's predictable).

On the other hand, S and N are each a **Q**-martingale. This means that

$$E_Q(\Delta N_i | F_{i-1}) = E_Q(\Delta S_i | F_{i-1}) = 0$$

because, for example,

$$
\begin{aligned}
&E_Q(\Delta N_i | F_{i-1}) \\
&= E_Q(N_i - N_{i-1} | F_{i-1}) \\
&= E_Q(N_i | F_{i-1}) - E_Q(N_{i-1} | F_{i-1}) \\
&= E_Q(N_T | F_{i-1}) - E_Q(N_T | F_{i-1}) \\
&= 0
\end{aligned}
$$

The same can be shown for S. Thus,

$$
\begin{aligned}
0 &= E_Q(\Delta N_i | F_{i-1}) \\
&= E_Q(\phi_i \Delta S_i + k_i | F_{i-1}) \\
&= 0 + k_i
\end{aligned}
$$

Therefore, $k_i = 0$.

From the above, we find that when the binomial processes S and N are each a **Q**-martingale, if we consider a general scale and shift for each process, the shift becomes 0, and only the scaling ϕ remains. This can be summarized as:

$$\Delta N_i = \phi_i \Delta S_i$$

This can then be proven by induction. It holds for $n = 1$:

$$N_1 - N_0 = \phi_1(S_1 - S_0)$$

Assuming the above equation holds for $n = k - 1$:

$$N_{k-1} = N_0 + \sum_{i=1}^{k-1} \phi_i \Delta S_i$$

we can say that

$$N_k = N_0 + \sum_{i=1}^{k} \phi_i \Delta S_i$$

holds, which proves the equation holds for all natural numbers k.

7.17   Application to Finance

Up to this point, we have understood a mathematical theorem about binomial tree processes and probability measures. The proof did not include any financial concepts like stocks or bonds, nor did it touch on arbitrage or market implications. In this next part, we'll see how the Binomial Martingale Representation Theorem can be applied to pricing.

The binomial tree model represents a market. We'll assume the stock price fluctuates as a binomial process S. If a probability measure $\mathbf{Q}$ exists that makes this process S a martingale, we can apply this theorem to represent any other $\mathbf{Q}$ martingale N_i in terms of the stock price.

From the Binomial Martingale Representation Theorem, the predictable ϕ_i can be used in an investment strategy. If we buy the appropriate amount of stock, ϕ_i, at each time point, the profit and loss will follow the $\mathbf{Q}$ martingale N_i (cash flow replication).

Now, note that a contract X is not a stochastic process but a random variable at maturity. Since this is difficult to handle mathematically, we'll take the conditional expectation $E_i = \mathbf{E_Q}(X|F_i)$ under the probability measure $\mathbf{Q}$ and consider the $\mathbf{Q}$-martingale E_i.

The bond price is a stochastic process B_i that is predictable and positive. Without loss of generality, we can set $B_0 = 1$.

Its inverse process B_i^{-1} is also predictable and positive. We call B_i^{-1} the discounting process.

We define a discounted stock price process $Z_i = B_i^{-1}S_i$. We call this Z_i the discounted stock price process.

A contract X discounted by the maturity discounting process, $B_T^{-1}X$, is also a contract. We call this $B_T^{-1}X$ the discounted contract.

7.18   The Replicating Strategy

For a probability measure $\mathbf{Q}$ that makes $Z = B_i^{-1}S_i$ a $\mathbf{Q}$-martingale, we take the conditional expectation $E_i = \mathbf{E_Q}(B_T^{-1}X|F_i)$ for a contract X. This E_i is a $\mathbf{Q}$-martingale constructed from Z.

Now, both E_i and Z are $\mathbf{Q}$-martingales. Therefore, by the Binomial Martingale Representation Theorem, there exists a predictable process ϕ_i such that

$$E_k = E_0 + \sum_{i=1}^{k} \phi_i \Delta Z_i$$

Using the above, we'll think about a replicating strategy.

At time i, we hold ϕ_{i+1} units of stock S.

At time i, we hold $\psi_{i+1} = (E_i - \phi_{i+1}B_i^{-1}S_i)$ units of the bond.

Based on these two points, we purchase the portfolio Π_i.

That is, at time i, we purchase the following portfolio:

$$\Pi_i = \phi_{i+1}S_i + \psi_{i+1}B_i$$

At $t = 0$, this portfolio Π_0 has the following value:

$$\begin{aligned}
\Pi_0 &= \phi_1 S_0 + \psi_1 B_0 \\
&= E_0 \\
&= \mathbf{E_Q}(B_T^{-1}X) \\
&= \mathbf{E_Q}(B_T^{-1}X|F_0)
\end{aligned}$$

This value is the cost required to acquire the portfolio from scratch at the beginning of the period. Since ϕ_1 and ψ_1 are predictable, they can be determined.

At the next time point, $t = 1$, the price of Π_0 changes by the following difference in price. (Note that it's Π_0, not Π_1. We're thinking about the price of the portfolio Π_0 which was held at $t = 0$ and has simply transitioned to the next state without any buying or selling.) (That is, we're considering

the case where the holdings ϕ_{i+1}, ψ_{i+1} for constructing Π_i are not changed, and only the processes S_i, B_i change from $i \to i+1$.)

$$
\begin{aligned}
& \phi_1 S_1 + \psi_1 B_1 \\
=\ & \phi_1 S_1 + \left(E_0 - \phi_1 B_0^{-1} S_0\right) B_1 \\
=\ & B_1 \left\{ E_0 + \phi_1 (B_1^{-1} S_1 - B_0^{-1} S_0) \right\}
\end{aligned}
$$

Here, if we set the discounted stock price increment $Z_1 - Z_0$ as $\Delta Z_1 = B_1^{-1} S_1 - B_0^{-1} S_0$, we can use the Binomial Martingale Representation Theorem ($k=1$), $E_1 = E_0 + \phi_1 \Delta Z_1$ to write:

$$
\phi_1 S_1 + \psi_1 B_1 = B_1 E_1
$$

(Note that this is not the price of Π_1. The value of Π_1 at $i=1$ is $\phi_2 S_1 + \psi_2 B_1$.)

In the replicating strategy, at time i, we purchase the portfolio

$$
\Pi_i = \phi_{i+1} S_i + \psi_{i+1} B_i
$$

The value $B_1 E_1$ is the cost required to purchase the replicating portfolio Π_0 when transitioning from time 0 to time 1.

In general, at any time i, it costs $B_i E_i$ to purchase the portfolio Π_i.

Next, if the portfolio Π_1 held at time 1 moves to $t=2$, we have similarly:

$$
\begin{aligned}
\Pi_1 \to\ & \phi_2 S_2 + \psi_2 B_2 \\
=\ & \phi_2 S_2 + \left(E_1 - \phi_2 B_1^{-1} S_1\right) B_2 \\
=\ & B_2 \left\{ E_1 + \phi_2 (B_2^{-1} S_2 - B_1^{-1} S_1) \right\} \\
=\ & B_2 \left(E_1 + \phi_2 \Delta Z_1\right) \\
=\ & B_2 E_2
\end{aligned}
$$

This shows that at the moment $t=2$, the cost to construct the replicating portfolio Π_2 is $B_2 E_2$.

References

[1] Financial Calculus - An Introduction to Derivative Pricing - Martin Baxter, Andrew Rennie