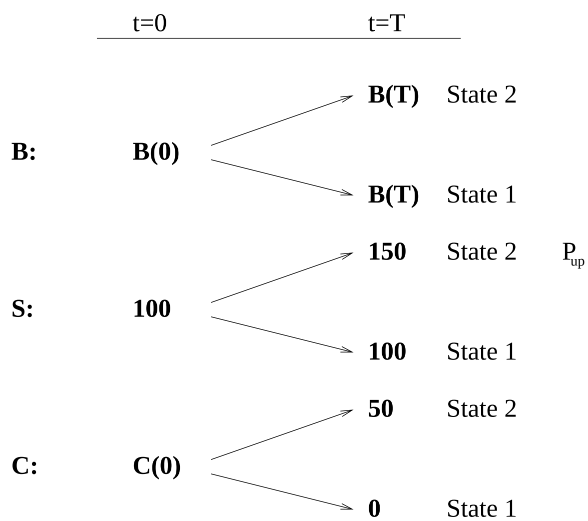
Computational Finance – The Martingale Measure and Pricing of Derivatives

1 The Martingale Measure

The Martingale measure or the Risk Neutral probabilities are a fundamental concept in the no-arbitrage pricing of instruments which links prices to expectations. This will be a very useful later on, because as we will see, there are very good randomized algorithms (Monte Carlo) for estimating expectations.

1.1 Two Period Stock–Bond Economy

To begin we revisit the previous an example economy we already studied, as illustrated in the following figure.



At $t = 0$, two instruments are available. Any amount (even fractional) of either instrument may be sold or puschased at the specified market price - i.e., arbitrary short or long positions are allowed. A *risk free* asset or *bond*, B, and a *stock*, S. At $t = 0$ (the first period), the bond is worth $B(0)$, and, the stock is worth $S(0) = 100$. At $t = T$ (the second period), the economy can be in one of two states. In both states, the bond is valued at *B(T)* and hence is risk free. In the first state the stock is valued at $S(T) = 100$ and in the second state, the stock is valued at $S(T) = 150$. Suppose that P_{up} is the probability that the market goes up. We have thus completely specified the market dynamics for our simple economy. In this simplified economy, it is clear that one can guarantee an amount $B(T)$ at $t = T$ by investing $B(0)$ in the bond at $t = 0$. The *risk free discount factor* is defined by

$$D(T) = \frac{B(0)}{B(T)} = e^{-rT}$$

where $r > 0$ is the risk free interest rate for continuous compounding. There is also a *European call option* on the stock with a strike price of 100, expiring at time T. At $t = T$, the call option will either be worth 50, if the market went up, or 0 if not.

Previously we have shown that if there is no possibility of "making money out of nothing", then the following two conditions must hold,

1. $2/3 < D(T) < 1$
2. $C(0) = 100(1 - D(T))$

independent of P_{up}. The restriction on the discount factor prevents one making guaranteed money from buying stock and shorting bond (or vice versa). The price of the call option was determined by constructing a portfolio with a stock and short one option, then the portfolio value at time T is independent of the state and hence acts like a bond. The option *hedges* against the risk in the stock.
The important things to note are:

1. The exact nature of the instruments that constitute the market are not relevant, only that they follow the specified market dynamics.
2. To price the third instrument (given $D(T)$, or alternatively, the interest rate), it was not necessary to know P_{up} (the probabilities of of the various states). One only needs to know what states are *possible*. Among other things, this implies the counter intuitive fact that the price of the option would be the same whether $P_{up} = 1 - 10^{-100}$ or $P_{up} = 10^{-100}$.

That the actual probability P_{up} plays no role in the pricing of the instruments is not a quirk of the particular economy that we happen to have chosen. We will now develop the general framework (the Martingale Measure framework) which essentially shows that the real world probabilities have very little to say about pricing derivatives. We need to formalize the notion of the economy, and some of the notions described already, such as the phrase "not being able to make money out of nothing".

1.2 Discrete Two Period Economies

First, we need to set up the notation to describe the economy/market. Suppose that there are N instruments, whose prices at time t are given by $S_i(t),\ i = 1, \dots, N$ which we will denote by the column vector $\mathbf{S}(t)$, that represents the state of the economy at time t. For the moment, we focus on the two times $t = 0$ and $t = T$. Suppose that at $t = T$, there are K possible states of the world, in other words, the state $\mathbf{S}(T)$ can be one of the K possible vectors $\mathbf{S}_1, \dots, \mathbf{S}_K$. Let the probability of state j occurring at $t = T$ be $P_j,\ j = 1 \dots K$, represented by the column vector $\mathbf{P}$. We define the payoff matrix at $t = T$ as the matrix $\mathbf{Z}(T) = [\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_K]$. The element $\mathbf{Z}_{ij}$ represents the value of one unit of instrument i in state j for $i = 1, \dots, N$ and $j = 1, \dots, K$. A particular row of $\mathbf{Z}$ represents the possible values that a particular instrument can take at time T. An instrument is risk free if its value in every state is a constant (i.e., if the row is a constant). A particular column represents the values of all the instruments in a particular state. Having specified $\mathbf{S}(t)$, $\mathbf{Z}(T)$ and $\mathbf{P}$, we have fully specified our economy, including its dynamics. For the example in the previous section, we have

$$\mathbf{S}(0) = \begin{bmatrix} B(0) \\ 100 \\ C(0) \end{bmatrix} \qquad \mathbf{Z}(T) = \begin{bmatrix} B(T) & B(T) \\ 100 & 150 \\ 0 & 50 \end{bmatrix} \qquad \mathbf{P}(T) = \begin{bmatrix} 1 - P_{up} \\ P_{up} \end{bmatrix}$$

A portfolio $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]$ is a column vector of N components that indicates how many units of each instrument is held. The value of such a portfolio at time t is given by

$$\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S}(t).$$

We use the notation $(\cdot)^{\mathrm{T}}$ to denote the transpose. Thus, the value of the portfolio at $t = 0$ is $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S}(0)$, the possible values that the portfolio can take on at time T are given by the row vector $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{Z}(T)$, and the expected value of the portfolio at time T is given by $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{Z}(T)\mathbf{P}(T)$.

Using this notation, let's re-examine the derivation of the result that the price of the call option in the very simple stock bond economy must be $100(1 - D(T))$. Suppose that $C(0) = 100(1 - D(T))(1 + \rho)$ where $\rho > 0$. Construct the portfolio

$$\boldsymbol{\theta} = \begin{bmatrix} -100/B(T) \\ 1 \\ -1/(1+\rho) \end{bmatrix}$$

At $t = 0$, the value of this portfolio is

$$\begin{bmatrix} \frac{-100}{B(T)} & 1 & \frac{-1}{(1+\rho)} \end{bmatrix} \begin{bmatrix} B(0) \\ 100 \\ 100(1 - D(T))(1+\rho) \end{bmatrix} = 0$$

At $t = T$, the possible values of the portfolio are[1]

$$\begin{bmatrix} \frac{-100}{B(T)} & 1 & \frac{-1}{(1+\rho)} \end{bmatrix} \begin{bmatrix} B(T) & B(T) \\ 100 & 150 \\ 0 & 50 \end{bmatrix} = \begin{bmatrix} 0 & \frac{50\rho}{1+\rho} \end{bmatrix} \geq \mathbf{0}$$

In every state of the world, the portfolio is worth at least zero, and in at least one state of the world, the portfolio is worth a strictly positive amount. Certainly, any rational person would want to consume an infinite amount of this portfolio since there is no downside, and only potential upside. Thus the economy cannot be in equilibrium. Hence we rule out the

[1]For vectors, we use the notation $\mathbf{x} \leq 0$ to signify a vector that has every component ≤ 0 and at least one component < 0. $\mathbf{x} \geq 0$ if and only if $-\mathbf{x} \leq 0$. We use the notation $\mathbf{x} \leqq 0$ to signify a vector with every component ≤ 0, allowing the possibility that every component $= 0$. $\mathbf{x} \geqq 0$ if and only if $-\mathbf{x} \leqq 0$.

existence of such portfolios portfolios because any "sane" individual would want to hold an unlimited amount of such portfolios, leading to disequilibrium, and so we rule out the possibility of $\rho > 0$.

Exercise 1.1

Suppose that $\rho < 0$. Construct a similar portfolio and hence argue that if the market is in equilibrium, then we can rule out the porribility of $\rho < 0$.

Thus the only remaining possibility is that $\rho = 0$, which it turns out does not lead to any possibility of such portfolios. The proof of this in this specific case requires some insight since it requires to show that no such portfolio exists. It will follow from a more general result we will prove later. The next step is to formulate precisely what these types of portfolios are that we wish to rule out.

1.3 The No-Arbitrage Axiom

Informally, arbitrage is the ability to make money from nothing. In an earlier discussion we defined arbitrage as the ability to access a non-negative stream of cash flows at least one of which was positive. This definition can be viewed as deterministic arbitrage. The example of arbitrage presented in the discussion above is not of this form, because one cannot guarantee a at least one positive cash flow. In one state of the world, all the cashflows were negative. In the other, they were all non-negative with at least one positive. It turns out that one construct a portfolio in the above example which guarantees at least one positive cash flow with all others being non-negative, and hence we have essentially deterministic arbitrage.

Exercise 1.2

Construct a portfolio in the above example in which all states of the world lead to a stream of cash flows in which all non-negative and at least one of which is positive.

In general one cannot guarantee the existence of a portfolio as in the previous exercise and have there are two different types of probabilistic arbitrage that we would like to define.

Definition 1.1 (Type I Arbitrage) *An* arbitrage opportunity *of type I exists if and only if there exists a portfolio* Θ *such that*

$$\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S}(0) \leq 0 \textit{ and } \boldsymbol{\theta}^{\mathrm{T}}\mathbf{Z}(T) \geq 0$$

Type I arbitrage is probabilistic notion of arbitrage in the sense that one is not guaranteed free money. The cost of the portfolio at time 0 is $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S}(0) \leq 0$. If this is 0, then it costs you nothing to own this portfolio. On the other hand the value of the portfolio in all states of the world is at least 0, with *at least* one state having value strictly positive. Thus this type of arbitrage gives the owner of the arbitrage portfolio $\boldsymbol{\theta}$ the *possibility* of money at no or negative cost. The mere possibility of free money alone should create disequilibrium, i.e. any rational individual should want to consume an infinite amount of such a portfolio.

Definition 1.2 (Type II Arbitrage) *An* arbitrage opportunity *of type II exists if and only if there exists a portfolio* Θ *such that*

$$\Theta^{\mathrm{T}}\mathbf{S}(0) < 0 \textit{ and } \Theta^{\mathrm{T}}\mathbf{Z}(T) \geqq 0$$

This kind of arbitrage portfolio has strictly negative cost, i.e., to own this portfolio, you get money (positive cashflow). And in every future state of the world, your cash flow is non-negative. This type of arbitrage portfolio is the generalization of our previous notion of deterministic arbitrage and is the *guarantee* of free money. Certainly any rational individual would want to consume an unlimited amount of such a portfolio.

An economy in equilibrium cannot sustain either type I or type II arbitrage opportunities. In words, an investment that guarantees a future non-negative return, with a possible positive return, must have a positive cost today, and, an investment that guarantees a non-negative return in the future must cost a non-negative amount today. Note that these two types of arbitrage opportunities are distinct in that the existence of type I arbitrage does not in general imply the existence of type II, and the existence of type II in general does not imply the existence of type I.

Exercise 1.3

Give an example economy in which there is a Type I arbitrage opportunity but no Type II arbitrage opportunity. Similarily, give an economy in which there is a Type II arbitrage opportunity but no Type I arbitrage opportunity.

Exercise 1.4

If there exists a risk free asset with positive value at $t = 0$ and $t = T$, show that any Type II arbitrage portfolio can be converted into a Type I arbitrage portfolio.

Hint: consider adding a suitable amount of the risk free asset to the portfolio]

The previous exercise shows that when there is a risk free asset, the existence of Type II arbitrage implies the existence of Type I arbitrage.

Axiom 1.3 (The No Arbitrage Axiom) *An economy in equilibrium cannot have arbitrage of Type I or Type II.*

1.4 The Positive Supporting Price Theorem

We are now ready for the main theorem, perhaps one of the most fundamental theorems in the arbitrage pricing theory. The statement of the theorem may appear a little abstract, but as we shall soon see, it essentially tells us that we can price an instrument at $t = 0$ (i.e., today) by taking an expectation over its future values at $t = T$ (i.e., tomorrow).

Theorem 1.4 (Positive supporting price) *There do not exist arbitrage opportunities of type I or type II* if and only if *There exists $\psi > 0$ such that $\mathbf{S}(0) = \mathbf{Z}(T)\psi$, or more explicitly, $\exists\psi > 0$ such that*

$$\begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_N \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} & \dots & Z_{1K} \\ Z_{21} & Z_{22} & \dots & Z_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & Z_{N2} & \dots & Z_{NK} \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_K \end{bmatrix} \qquad \psi_i > 0$$

In the theorem, S refers to $t = 0$ and Z refers to $t = T$. Since this is the central theorem, we provide an elementary proof using some fundamental tools from separating plane theory for convex sets. Before we move on to the proof, we elaborate on the important points in the theorem, and examine some of the consequences of this theorem. First note that the statement of the theorem is an *if and only if* statement, thus if we find a $\boldsymbol{\psi} > 0$ such that $\mathbf{S} = \mathbf{Z}\boldsymbol{\psi}$, then there is no arbitrage, and no arbitrage implies the existence of such a $\psi > 0$. Second note that every component of ψ should be strictly positive. Third, and most important, note that in the statement of the theorem, no where do the real world probabilities $\mathbf{P}$ come in.

The beauty of this theorem is that it induces a linear relationship (with positive coefficients) between the current prices and the possible future prices. In particular, if some subset of todays prices are not known, this linear relationship may be sufficient to determine them assuming there is no arbitrage. Lets see the application of this theorem to determine such prices in our simple stock-bond economy. it states that if there is no arbitrage in our little economy, then there exists $\boldsymbol{\psi} > 0$ such that

$$\begin{bmatrix} B(0) \\ 100 \\ C(0) \end{bmatrix} \begin{bmatrix} B(T) & B(T) \\ 100 & 150 \\ 0 & 50 \end{bmatrix} \begin{bmatrix} \psi_1 \\ \psi_2 \end{bmatrix},$$

or that

$$\begin{aligned} D(T) &= (\psi_1 + \psi_2), \\ 100 &= 100\psi_1 + 150\psi_2, \\ C(0) &= 50\psi_2. \end{aligned}$$

Notice that the first two equations give us two equations in two unknowns, ψ_1, ψ_2. Solving for ψ_1, ψ_2 we find that

$$\begin{aligned} \psi_1 &= 3(D(T) - 2/3), \\ \psi_2 &= 2(1 - D(T)). \end{aligned}$$

For $\boldsymbol{\psi} > 0$, it must therefore be that $\frac{2}{3} < D(T) < 1$. In this case $C(0) = 100(1 - D(T))$, and we have effortlessly reproduced all the results we painstakingly derived before. In particular, because the theorem is an if and only if theorem, we have that if $\frac{2}{3} < D(T) < 1$ and $C(0) = 100(1 - D(T))$, then for $\boldsymbol{\psi} > 0$ as specified above, $\mathbf{S} = \mathbf{Z}\boldsymbol{\psi}$ and so there is no arbitrage.

In general, if $N > K$, then there are fewer future states of the world than there are instruments. In that case, if K of the instruments have known prices today, then these K instruments can be used to extract the supporting vector ψ as follows. For concreteness suppose it is the first K instruments whose prices today are known, and let $\mathbf{S}_K$ be the vector consisting of the first K prices today. Similarly let $\mathbf{Z}_K$ be the first K rows of the $\mathbf{Z}$ matrix, which are the known prices of the instruments in the future. Then,

$$\boldsymbol{\psi} = \mathbf{Z}_K^{-1}\mathbf{S}_K$$

is the only set of prices which can satisfy the requirements of the positive supporting price theorem (providing $\mathbf{Z}_K$ is invertible). The first thing to check is that $\boldsymbol{\psi} > 0$; if not then there is arbitrage, as there is no $\boldsymbol{\psi} > 0$ which can satisfy $\mathbf{S}_K = \mathbf{Z}_K\boldsymbol{\psi}$, let alone $\mathbf{S} = \mathbf{Z}\boldsymbol{\psi}$. If $\boldsymbol{\psi} > 0$, then

$$\mathbf{S} = \mathbf{Z}\mathbf{Z}_K^{-1}\mathbf{S}_K$$

prices all the instruments today. If some of those instruments have prices already, they must agree with these prices or else there is arbitrage. If some of those instruments do not have prices, then these are the only possible arbitrage-free prices. This formula represents all the prices today as a function of the current and future prices of the first K instruments.

Exercise 1.5

One problem can arise in the above prescription if the matrix $\mathbf{Z}_K$ is not invertible. In this case some row of $\mathbf{Z}_K$ is a linear combination of the other rows. Show that in this event, the current price of that instrument should be this exact linear combination of the current prices of the other instruments, or else there is arbitrage. Thus, this instrument is redundant and can be removed from the economy. Hence we are justified in assuming that $\mathbf{Z}_K$ is invertible.

Exercise 1.6

[The Two Stock, Bond Economy]

1.4.1 Proof of the Positive Supporting Price Theorem

We will now give the formal proof of the main theorem. First, suppose that there is a set of positive supporting prices $\boldsymbol{\psi} > 0$ such that $\mathbf{S} = \mathbf{Z}\boldsymbol{\psi}$. We will show that there is no arbitrage. Suppose that there is a portfolio $\boldsymbol{\theta}$ such that $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S} \leq 0$ with $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{Z} \geq 0$. In this case $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\psi} > 0$ since $\boldsymbol{\psi} > 0$; but since $\mathbf{S} = \mathbf{Z}\boldsymbol{\psi}$, this means that $\boldsymbol{\theta}\mathbf{S} > 0$, contradicting $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S} \leq 0$, and so there cannot be a Type I arbitrage. Suppose that there is a portfolio $\boldsymbol{\theta}$ such that $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S} < 0$ with $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{Z} \geq 0$. In this case $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\psi} \geq 0$ since $\boldsymbol{\psi} > 0$; but since $\mathbf{S} = \mathbf{Z}\boldsymbol{\psi}$, this means that $\boldsymbol{\theta}\mathbf{S} \geq 0$, contradicting $\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S} < 0$, and so there cannot be a Type II arbitrage. Thus, if $\mathbf{S} = \mathbf{Z}\boldsymbol{\psi}$ for $\boldsymbol{\psi} > 0$, then there is no arbitrage.

We now consider the converse, which is the hard part of the proof. Suppose that there is no arbitrage of Type I or II. We will introduce two sets M and C as follows.

Consider all possible portfolios, $\boldsymbol{\theta}$ and to each portfolio define the $K + 1$ dimensional vector $\mathbf{x}(\boldsymbol{\theta}) \in \mathbb{R}^{K+1}$ by

$$\mathbf{x}(\boldsymbol{\theta}) = \begin{bmatrix} -\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S} \\ \mathbf{Z}^{\mathrm{T}}\boldsymbol{\theta} \end{bmatrix}.$$

The first component is the value of the portfolio at $t = 0$, and the remaining components are the possible future values of the portfolio in the K possible future states of the world. The set M is the set of all possible vectors $\mathbf{x}(\boldsymbol{\theta})$ induced by all possible $\boldsymbol{\theta} \in \mathbb{R}^N$,

$$M = \left\{ \mathbf{x}(\boldsymbol{\theta}) \in \mathbb{R}^{K+1} : \boldsymbol{\theta} \in \mathbb{R}^N, \mathbf{x}(\boldsymbol{\theta}) = \begin{bmatrix} -\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S} \\ \mathbf{Z}^{\mathrm{T}}\boldsymbol{\theta} \end{bmatrix} \right\}.$$

Since the zero portfolio $\boldsymbol{\theta} = \mathbf{0}$ is a perfectly valid portfolio, for which $\mathbf{x}(\mathbf{0}) = \mathbf{0}$, and so $\mathbf{0} \in M$. Note that $\mathbf{x}(-\boldsymbol{\theta}) = -\mathbf{x}(\boldsymbol{\theta})$, and so if $\mathbf{x} \in M$, then $-\mathbf{x} \in M$. Two important observations about M are that it is closed and convex.

Exercise 1.7

Prove that M is closed and convex.

A set is closed if its complement is open. A set is open U is open if for every $\mathbf{x} \in U$ there is a ball B small enough and centered at $\mathbf{x}$ such that the entire ball is in U, i.e. $B \subset U$.

A set U is convex if for every $\mathbf{x}, \mathbf{y} \in U$ and any $0 \leq \alpha \leq 1$, $\alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in U$.

We define the positive orthant cone C to be the set of vectors with all coordinates non-negative,

$$U = \left\{\mathbf{y} \in \mathbb{R}^{K+1} : y_i \geq 0\right\}.$$

Note that $\mathbf{0} \in C$, and so $\mathbf{0} \in M \cap C$. Just like M, C is also closed and convex.

Exercise 1.8

Prove that C is closed and convex.

The no arbitrage condition is equivalent to the condition that the only intersection point of C and M is $\mathbf{0}$.

Lemma 1.5 *There is no arbitrage of Types I and II if and only if the only intersection point of C and M is $\mathbf{0}$.*

The proof of this lemma is relatively straightforward, and we leave it as an exercise. Essentially the points in M which are in the positive orthant are arbitrage opportunities, and since there are no such opportunities, there must be no points in M which are in $C \setminus \{\mathbf{0}\}$.

Exercise 1.9

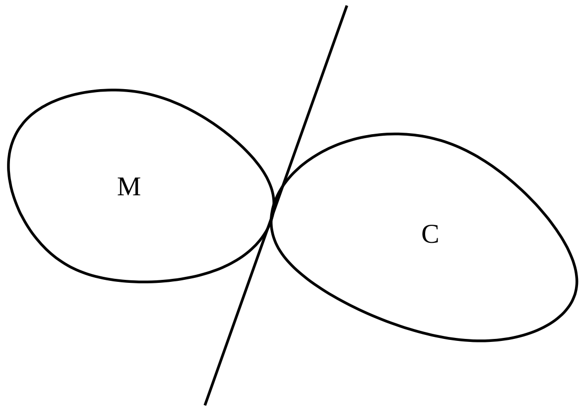
Prove Lemma 1.5.

Thus, M and C are closed and convex and intersect only at the point $\mathbf{0}$. We now invoke a very powerful theorem from linear programming, a separating hyperplane theorem for convex sets, [?].

Lemma 1.6 (Separating Hyperplane Lemma) *Let M, C be two closed convex sets having a single point of intersection $\mathbf{x}_0$. Then there is a separating hyperplane $(w_0, \mathbf{w})$ passing through p which separates M and C. Specifically,*

$$\begin{aligned} \forall\mathbf{x} \in M \qquad & w_0 + \mathbf{w}^T\mathbf{x} \leq 0, \\ \forall\mathbf{y} \in C \setminus \mathbf{x}_0 \qquad & w_0 + \mathbf{w}^T\mathbf{y} > 0. \end{aligned}$$

The proof of this lemma is beyond the scope of our discussion, however, we give a pictorial proof of the lemma below,



Since $\mathbf{0} \in M$, we conclude that $w_0 \leq 0$. Thus, $\mathbf{w} \neq \mathbf{0}$ because otherwise $w_0 + \mathbf{w}^{\mathrm{T}}\mathbf{y} \leq 0$ for all $\mathbf{y} \in C$, which is a contradiction. Suppose that $w_0 < 0$, then choose a $\mathbf{y}^* \in C \setminus \mathbf{0}$ such that $\| \mathbf{y}^* \| < -\frac{w_0}{\| \mathbf{w} \|}$. In this case $|\mathbf{w}^{\mathrm{T}}\mathbf{y}^*| \leq \| \mathbf{w} \| \| \mathbf{y}^* \| < -w_0$ from which we deduce that $\mathbf{w}^{\mathrm{T}}\mathbf{y}^* < -w_0$. Hence $w_0 + \mathbf{w}^{\mathrm{T}}\mathbf{y}^* < 0$ which is a contradiction, hence the only possibility is that $w_0 = 0$, from which we have that for some $\mathbf{w} \in \mathbb{R}^{K+1}$,

$$\begin{aligned} \forall \mathbf{x} \in M \qquad & \mathbf{w}^{\mathrm{T}}\mathbf{x} \leq 0, \\ \forall \mathbf{y} \in C \setminus \mathbf{x}_0 \qquad & \mathbf{w}^{\mathrm{T}}\mathbf{y} > 0. \end{aligned}$$

We now argue that $\mathbf{w} > 0$. Suppose, to the contrary, that some $w_i \leq 0$, and select y such that $y_i > 0$ and $y_j = 0$ for $j \neq i$. Then $\mathbf{y} \in C \setminus \mathbf{0}$ and $\mathbf{w}^{\mathrm{T}}\mathbf{y} = 0$ which is a contradiction, hence $\mathbf{w} > 0$.

We now claim that for all $\mathbf{x} \in M$,$\mathbf{w}^{\mathrm{T}}\mathbf{x} = 0$. Indeed, suppose to the contrary that for some $\mathbf{x}$, $\mathbf{w}^{\mathrm{T}}\mathbf{x} < 0$. Then, since $-\mathbf{x} \in M$, and $-\mathbf{w}^{\mathrm{T}}\mathbf{x} > 0$ which is a contradiction. Hence $\forall \mathbf{x} \in M$, $\mathbf{w}^{\mathrm{T}}\mathbf{x} = 0$. Let $\mathbf{w} = [\beta_0, \boldsymbol{\beta}]^{\mathrm{T}}$. Then the statement $\mathbf{w}^{\mathrm{T}}\mathbf{x} = 0$ for all $\mathbf{x} \in M$ is equivalent to the statement

$$\forall \boldsymbol{\theta} \in \mathbb{R}^N \qquad -\boldsymbol{\theta}^{\mathrm{T}}\mathbf{S}\beta_0 + \boldsymbol{\theta}^{\mathrm{T}}\mathbf{Z}\boldsymbol{\beta} = 0.$$

Let $\boldsymbol{\psi} = \boldsymbol{\beta}/\beta_0$, then since $\mathbf{w} > 0$, $\boldsymbol{\psi} > 0$, and

$$\forall \boldsymbol{\theta} \in \mathbb{R}^N \qquad \boldsymbol{\theta}^{\mathrm{T}}(\mathbf{Z}\boldsymbol{\psi} - \mathbf{S}) = 0.$$

Since the above equation holds for all $\boldsymbol{\theta} \in \mathbb{R}^N$, i.e., $\mathbf{Z}\boldsymbol{\psi} - \mathbf{S}$ is orthogonal to every vector in $\mathbb{R}^N$, it must be that $\mathbf{Z}\boldsymbol{\psi} - \mathbf{S} = \mathbf{0}$, concluding the proof of the theorem.

1.5 The Equivalent Martingale (Risk Neutral) Measure

The essential content of the positive supporting price theorem is that if there is no arbitrage, then a positive price vector ψ, exists such that the prices today are given by a weighted sum of the prices tomorrow, where the weightings are given by the price vector ψ. The weightings are the *same* for each instrument. This looks suspiciously like an expectation over the possible prices tomorrow taken over the distribution ψ. To fully appreciate this, we need to massage the theorem into a more suitable form. First, notice that only the relative prices should matter. To this end, we pick an instrument with respect to which we price all other instruments. This instrument is typically called the *numeraire*, which for our purposes is just a fancy word for reference instrument.

Often, one chooses the numeraire to be the risk free instrument if one exists[2]. Without loss of generality, assume that the first instrument is chosen as numeraire. Rewriting (1.4), we have,

$$\begin{bmatrix} 1 \\ S_2/S_1 \\ \vdots \\ S_N/S_1 \end{bmatrix} = \left[\begin{pmatrix} 1 \\ Z_{21}/Z_{11} \\ \vdots \\ Z_{N1}/Z_{11} \end{pmatrix} \begin{pmatrix} 1 \\ Z_{22}/Z_{12} \\ \vdots \\ Z_{N2}/Z_{12} \end{pmatrix} \cdots \begin{pmatrix} 1 \\ Z_{2K}/Z_{1K} \\ \vdots \\ Z_{NK}/Z_{1K} \end{pmatrix} \right] \begin{bmatrix} \psi_1 Z_{11}/S_1 \\ \psi_2 Z_{12}/S_1 \\ \vdots \\ \psi_K Z_{1K}/S_1 \end{bmatrix} \tag{1}$$

Notice that for each instrument, we are measuring its price in a given state relative to the price of the numeraire in that state. Define a vector $\tilde{\mathbf{P}}$ by

$$\tilde{p}_i = \frac{Z_{1i}}{S_1}\psi_i$$

Then, since $\boldsymbol{\psi} > 0$, we see that $\tilde{\mathbf{P}} > 0$. Further, from the first row in (1), $\sum_i \tilde{P}_i = 1$, thus $\tilde{\mathbf{P}}$ is a probability vector. Further, the remaining rows in (1) give that

$$\frac{S_i}{S_1} = \sum_{j=1}^{K} \tilde{P}_j \frac{Z_{ij}}{Z_{1j}}$$

We have now arrived at an equivalent formulation of the positive supporting price theorem,

Theorem 1.7 (Equivalent Martingale Measure) *There do not exist arbitrage opportunities of type I or II if and only if there exists a probability vector* $\tilde{\mathbf{P}}$*, called an* equivalent martingale measure *such that*

$$\frac{S_i(0)}{S_1(0)} = E_{\tilde{\mathbf{P}}}\left[\frac{S_i(T)}{S_1(T)}\right]$$

In other words, S_i/S_1 *is a martingale*[3] *under the measure* $\tilde{\mathbf{P}}$*.*

The measure $\tilde{\mathbf{P}}$ is also often referred to as the *risk-neutral measure*, or the *risk-adjusted probabilities*. If there is a risk free asset, one usually chooses this as the numeraire, in which case

$$S_i(0) = E_{\tilde{\mathbf{P}}}\left[\frac{S_1(0)}{S_1(T)}S_i(T)\right] = E_{\tilde{\mathbf{P}}}\left[D(T)S_i(T)\right] = D(T)E_{\tilde{\mathbf{P}}}\left[S_i(T)\right]$$

[2]It should be clear that there cannot exist two risk free instruments offering different returns.

[3]X is a martingale if $X(0) = E[X(T)]$.

where $D(T)$ factors out of the expectation because the numeraire is a risk free asset and thus is constant in every state. If the numeraire is not a risk free asset, then this step is not valid and the "discount factor" must remain inside the expectation. Thus, *the price of an instrument today is given by the discounted expectation of its future price, where the expectation is with respect to the risk neutral measure*,

$$S_i(0) = PV(E_{\tilde{\mathbf{P}}}[S_i(T)]).$$

We emphasize that this is a completely natural process, and had we started the discussion by saying that to compute the price of an instrument you take its expected future value and compute the present value of this expected future value one might have probably dismissed it as obvious. However, this is *not* what we are saying. What we are saying is that the real world probabilities are irrelevant. There is no mention of the real world probabilities $\mathbf{P}$ in the theorem. The theorem says that doing the intuitive thing (present value of expected future values) is correct for *some* probability measure $\tilde{\mathbf{P}}$. This probability measure $\tilde{\mathbf{P}}$ in general has no relationship with the real world probabilities. Also we emphasize that the prices of *all* instruments are obtained by takine the present value of expected future values with respect to the same Martingale Measure $\tilde{\mathbf{P}}$. One can view the Martingale Measure as defining some other "fictitious" economy in which the probabilities are $\tilde{\mathbf{P}}$ in which prices are given by present value of future expected values. This fictitious world is usually denoted the *risk neutral world* for the obvious reason that risk does not seem to affect the price.

To illustrate how this theorem is applied, we continue with the example that was presented earlier. By theorem 1.7, there is a probability vector $\tilde{\mathbf{P}} > 0$ such that

$$\begin{bmatrix} 1 \\ 100/B(0) \\ C(0)/B(0) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 100/B(T) & 150/B(T) \\ 0 & 50/B(T) \end{bmatrix} \begin{bmatrix} \tilde{P}_{down} \\ \tilde{P}_{up} \end{bmatrix}$$

The first two rows yield that $\tilde{P}_{up} = 2(1/D(T) - 1)$ and $\tilde{P}_{down} = 1 - \tilde{P}_{up}$, hence, because $0 < \tilde{P}_{up} < 1$, we see that $2/3 < D(T) < 1$. The third row then yields $C(0) = 100(1 - D(T))$. We have thus (mechanically) reproduced this result for third time now.

In general, if the number of states in the economy is large, the risk neutral measure will not be unique. More generally, suppose that the rank of $\mathbf{Z}$ is k, which represents the number of independent instruments. Then, we see that if the number of states (K) in the next period is less than or equal to k, we can choose K linearly independent rows of $\mathbf{Z}$ to solve uniquely for the risk neutral probabilities (by matrix inversion). If all of these risk neutral probabilities are not real probabilities, then an arbitrage opportunity exists. If the risk neutral probabilities exist, then they can be used to price *all* the instruments. If, on the other hand, $k < K$, then there are either infinitely many possible risk neutral probabilities (in which case, there are many pricings that are consistent with no arbitrage) or there are none (in which case, there is an arbitrage opportunity). As a further note, the the risk neutral probabilities can be obtained without paying any attention to the actual probabilities with which the future states occur. One simply needs to agree on what states are *possible*.

Exercise 1.10

Consider the following two period economies. Determine (to within reasonable precision) which ones have arbitrage opportunities. If you think that an economy has an arbitrage opportunity, give the portfolio that results in the arbitrage opportunity. If you believe that there is no arbitrage opportunity, then give the risk neutral probabilities (or the Martingale measure). Note, sometimes the risk neutral probabilities may not be unique. In all cases, justify your answer.

We use $\mathbf{S}$ to denote the current instrument prices, and $\mathbf{Z}$ to denote the future instrument price matrix.

(a)

$$(i) \qquad \mathbf{S}_1 = \begin{bmatrix} 5 \\ 14 \\ 20 \\ 33.5 \end{bmatrix} \qquad \mathbf{Z}_1 = \begin{bmatrix} 14 & 5 & 5 \\ 10 & 12 & 17 \\ 10 & 17 & 25 \\ 10 & 9 & 8 \end{bmatrix}$$

$$(ii) \qquad \mathbf{S}_2 = \begin{bmatrix} 5.47 \\ 14 \\ 20 \\ 8.825 \end{bmatrix} \qquad \mathbf{Z}_2 = \begin{bmatrix} 14 & 5 & 5 \\ 10 & 12 & 17 \\ 10 & 17 & 25 \\ 10 & 9 & 8 \end{bmatrix}$$

(b)

$$(i) \qquad \mathbf{S}_1 = \begin{bmatrix} 15 \\ 10 \end{bmatrix} \qquad \mathbf{Z}_1 = \begin{bmatrix} 14 & 5 & 5 \\ 10 & 12 & 17 \end{bmatrix}$$

$$(ii) \qquad \mathbf{S}_2 = \begin{bmatrix} 8 \\ 13 \end{bmatrix} \qquad \mathbf{Z}_2 = \begin{bmatrix} 14 & 5 & 5 \\ 10 & 12 & 17 \end{bmatrix}$$

Change of Numeraire

In an economy where the risk neutral probabilities are well defined, they still need not be unique. It is possible for different people to disagree as to the choice of numeraire. The choice of numeraire is arbitrary, and a particular numeraire is chosen in order to make subsequent calculations easier (for example it may be easier to derive the risk neutral measure for a particular numeraire). However, independent of the choice of numeraire, provided that the correct expectation is computed, the prices obtained for the assets at $t = 0$ are unique. It is possible that the risk neutral measure is more easily obtained with a particular numeraire, but that the expectations are more conveniently computed with respect to another numeraire.

If this is the case, then one needs to change the risk neutral probabilities to the relevant numeraire. Suppose that with numeraire S_i one has computed the martingale measure $\tilde{\mathbf{P}}$, but that one wishes to compute the expectations with S_j as numeraire. By direct substitution into (1.7) it can be verified that the new martingale measure $\tilde{\mathbf{P}}'$ that reflects this *change of numeraire* is given by

$$\tilde{p}'_\alpha = \frac{S_i/Z_{i\alpha}}{S_j/Z_{j\alpha}}\tilde{p}_\alpha$$

In particular, all one needs to know are the discount factors for the two instruments in the various future states. Such issues are discussed in more detail in [**?**]. In general, there is a risk free asset and it is usually most convenient to choose it as the numeraire.

1.6 Continuous State Economies

We have restricted our attention to a finite state economy. We will here present a very brief discussion of the continuous state (but finite instrument) economy. More details can be found in some of the references.

The future state can be indexed by a random variable $\tilde{s}$ which we assume to have a probability density function $\pi(s)$. Then the price matrix at $t = T$ becomes $\tilde{\mathbf{Z}}(\tilde{s})$, a random vector. Given a portfolio Θ, its value at $t = T$ is a random variable $\tilde{V}_\Theta(s) = \Theta^{\mathrm{T}}\tilde{\mathbf{Z}}(s)$. Then, type I arbitrage would require the existence of a portfolio that satisfies

$$\Theta^{\mathrm{T}}\mathbf{S}(0) \leq 0, \qquad \tilde{V}_\Theta \geq 0, \qquad P[\tilde{V}_\Theta > 0] > 0$$

and type II arbitrage requires the existence of a portfolio that satisfies

$$\Theta^{\mathrm{T}}\mathbf{S}(0) < 0, \qquad \tilde{V}_\Theta \geq 0$$

The analog of theorem 1.4 is that under suitable regularity conditions, the absence of arbitrage opportunities is equivalent to the existence of a positive integrable function $\psi(s)$ such that,

$$\mathbf{S}(0) = \int ds\ \tilde{\mathbf{Z}}(s)\psi(s), \qquad \psi(s) > 0$$

Further, choosing a numeraire (instrument 1) and redefining $\tilde{\pi}(s) = \psi(s)\tilde{\mathbf{Z}}_1(s)/S_1(0)$, we have that

$$S_i(0) = \int \frac{S_1(0)}{\tilde{\mathbf{Z}}_1(s)}\tilde{\mathbf{Z}}_i(s)\tilde{\pi}(s) = E_{\tilde{\pi}}[D(T)(s)\tilde{\mathbf{Z}}_i]$$

The previous section suggests an approach to pricing which we summarize here.

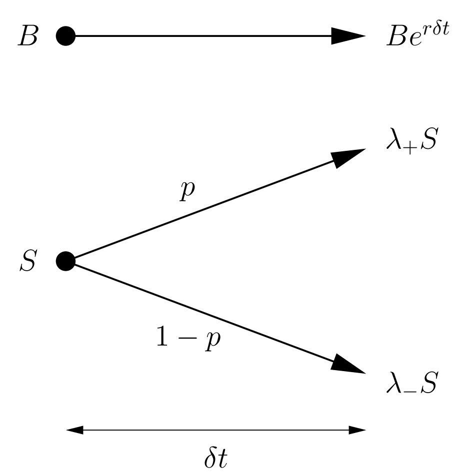
> Let K be the number of possible future states for the economy, and let the number of independent instruments be $N = K' + n$. Without loss of generality, we assume that the dynamics of the first K' (underlying) instruments are completely specified, including their current prices. The remaining n instruments are derivatives of the first K' instruments, in that their values in the possible future states are known, given the values of the underlying instruments. It is necessary to obtain the correct prices of the n derivatives at $t = 0$.

1. Choose a numeraire (reference instrument), and call this S_1.
2. Solve the set of K' simultaneous linear equations in the K unknowns $\tilde{p}_i > 0$, $i = 1 \ldots K$,
$$\frac{S_i}{S_1} = \sum_{j=1}^{K} \frac{Z_{ij}}{Z_{1j}} \tilde{p}_j \qquad \text{for } i = 1, \ldots, K'$$
to obtain the martingale measure $\tilde{\mathbf{P}}$.
3. Obtain the risk neutral prices of all the instruments as follows
$$S_i(0) = S_1(0) E_{\tilde{\mathbf{P}}} \left[\frac{S_i(T)}{S_1(T)} \right] = E_{\tilde{\mathbf{P}}} \left[D(T) S_i(T) \right] \stackrel{(a)}{=} D(T) E_{\tilde{\mathbf{P}}} \left[S_i(T) \right]$$
where $D(T) = S_1(0)/S_1(T)$ and (a) only follows if S_1 is a risk free instrument.
4. These prices (if the measure exists) are unique if $K' > K$. If such a measure cannot be found, then there exist arbitrage opportunities.

The connection with Monte Carlo techniques should now be clear. It is often possible to obtain the risk neutral measure, but due to the complexity of derivatives, it is often not possible to compute the desired expectation analytically. One still needs to price the derivatives efficiently, and an ideal tool for obtaining the necessary expectation is a Monte Carlo simulation. The following examples will illustrate the technique and how Monte Carlo simulations can be useful.

2 Stock and Bond Binomial Tree Model

Our little stock-bond economy which we have used all along is very simple, however a suitable generalization of it leads to a very powerful model of a stock and bond economy. Suppose the economy contains a stock and a risk free asset whose value grows according to a compounding interest rate in the one time period model



This is a two state two instrument two period economy with a risk free instrument, pretty much the simplest non-trivial economy that we could conceive of. This economy (and its dynamics) can be fully specified by giving $\mathbf{S}(0)$, $\mathbf{Z}(\delta t)$ and $\mathbf{P}$. In this dynamics, the stock, at time δt can exist in an up state, $\lambda_+ S$, or a down state, $\lambda_- S$, where we assume without loss of generality that $\lambda_+ > \lambda_-$,

$$\mathbf{S}(0) = \begin{bmatrix} B \\ S \end{bmatrix} \qquad \mathbf{Z}(\delta t) = \begin{bmatrix} Be^{r\delta t} & Be^{r\delta t} \\ S\lambda_+ & S\lambda_- \end{bmatrix} \qquad \mathbf{P} = \begin{bmatrix} p \\ 1-p \end{bmatrix}$$

Setting up and solving (1) for $\tilde{p}_u$, we have

$$\begin{bmatrix} 1 \\ \dfrac{S(0)}{B(0)} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \dfrac{\lambda_+ S}{Be^{r\delta t}} & \dfrac{\lambda_- S}{Be^{r\delta t}} \end{bmatrix} \begin{bmatrix} \tilde{p} \\ 1-\tilde{p} \end{bmatrix},$$

where $\tilde{p}$ is the Martingale Measure. We have one non-trivial equation in one unknown $\tilde{p}$,

$$S = e^{-r\delta t}(\tilde{p}\lambda_+ S + (1-\tilde{p})\lambda_- S),$$

which implies that

$$\tilde{p} = \frac{e^{r\delta t} - \lambda_-}{\lambda_+ - \lambda_-} \tag{2}$$

The requirement of no arbitrage gives that the stock price today is the linear combination of two different prices tomorrow. This in an of itself is a vacuous statement until we add the additional constraint that $0 < \tilde{p} < 1$, which means that $\lambda_- < e^{r\delta t}$.

This is a very simplistic model of the stock price, and it is only a good aproximation in the limit $\delta t \to 0$. We will see later that this limit is ultimately equivalent to a very popular model for the stock price, namely the *Geometric Brownian Motion*.

Pricing Derivatives The most important aspect of the "fictitious" risk neutral (Martingale) world is that this is where we can now price any instrument, specifically any derivative of S, as long as its values in the next time step are known. Pricing is done by taking the present value of the expected future value where the expectation is with respect to the risk neutral (Martingale) probability $\tilde{p}$.

In particular, lets consider as an example the call option with strike K with maturity δt. It's value at time δt is $\max\{0, S(\delta t) - K\} = [S(\delta t) - K]^+$. (We use the notation $[\cdot]^+$ as shorthand for taking the positive part.) Let $C(K)$ be the price of this option, then taking the present value of the expected future value, we have for the price

$$C(K) = e^{-r\delta t}\left(\tilde{p}[\lambda_+ S - K]^+ + (1-\tilde{p})[\lambda_- S - K]^+\right).$$

In general, a derivative f is a function which specifies the cashflow at the future time δt. In our case there are two possible cashflows to specify, $f(\lambda_+ S)$ and $f(\lambda_- S)$. Let F be the price of this derivative, then

$$F == e^{-r\delta t}\left(\tilde{p}f(\lambda_+ S) + (1-\tilde{p})f(\lambda_- S)\right).$$

Exercise 2.1

What is the price $P(K)$ at time 0 of the put option with strike K and maturity δt.

Exercise 2.2

[Three State Stock Dynamics.]

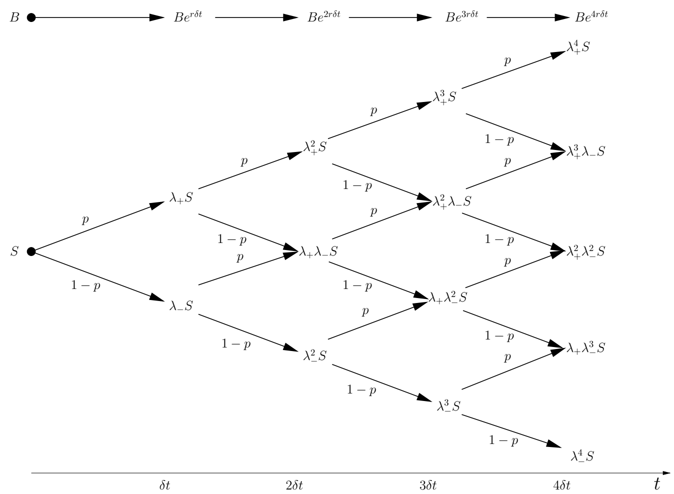
In this dynamics, the stock can exist in an "unchanged", up or down state in the next time period.

$$\mathbf{S}(0) = \begin{bmatrix} B \\ S \end{bmatrix} \qquad \mathbf{Z}(\delta t) = \begin{bmatrix} Be^{r\delta t} & Be^{r\delta t} & Be^{r\delta t} \\ S\lambda_+ & S\lambda_0 & S\lambda_- \end{bmatrix} \qquad \mathbf{P} = \begin{bmatrix} P_u \\ P_0 \\ P_d \end{bmatrix}$$

For this economy, obtain the risk neutral probabilities (how many possibilities are there?)

2.1 Multi-period Binomial Tree Dynamics

We now extend our analysis to the multiperiod setting, in which we move forward in time using an independent binomial two period model at each time step. We will assume that the model is stationary, so that at each time step, the stock price can go up by a factor λ_+ or down by a factor λ_-. This results in a recombining binomial tree model, which has very important consequences on the efficiency of pricing algorithms. Each time step is independent in the real world, and so we have the following picture for the dynamics, which we illustrate to time $4\delta t$.



One important property of the recombining nature of the tree (more appropriately a lattice[4]) is that its size does not grow exponentially.

Exercise 2.3

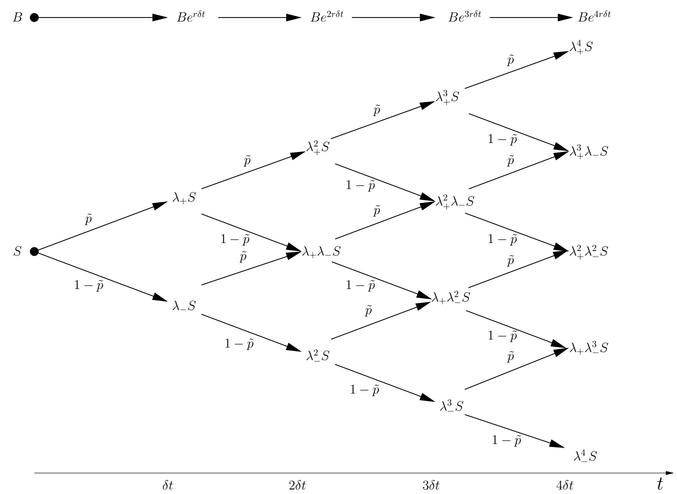
For the n-step tree, how many possible final stock prices are there?

What is the probability of a particular final stock price?

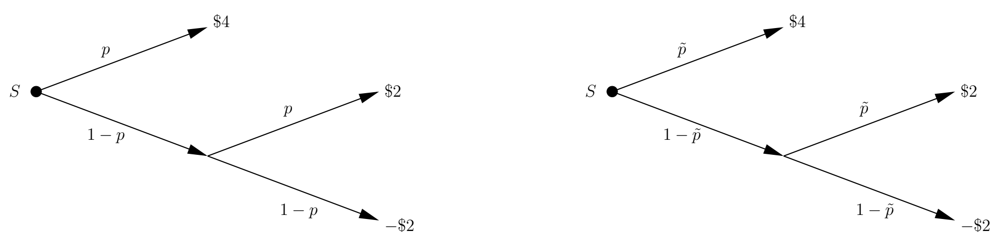
As we will prove later, the risk neutral (Martingale) world in this multiperiod setting is also an *independent* binomial tree dynamics, with the probability $\tilde{p}$ which is exactly the same $\tilde{p}$ which we obtained int he two period binomial model. The Martingale world, illustrated below, is the exact analogue of the real world with p replaced by $\tilde{p}$, where

$$\tilde{p} = \frac{e^{r\delta t} - \lambda_-}{\lambda_+ - \lambda_-}.$$

[4]Such binary lattices are often used to model economies (see for example [**?**]).



The general statement of the Martingale Measure theorem for pricing is that the price of an unpriced derivative can be obtained by obtaining the expected present value of future cashflows in this "fictitious" risk neutral world. This fictitious risk neutral world is a binomial multiperiod model with probability $\tilde{p}$. To illustrate, we consider a two period *derivative* which pays \$4 if the stock goes up. If the stock goes down, and then up, the cashflow is \$2 and if it goes down twice the cashflow is $-\$2$ (i.e., you owe \$2). The possible cashflows are illustrated below,



Real World Risk Neutral Pricing World

The real world dynamics of this derivative are not relevant to the pricing. All the pricing goes on in the risk neutral world. We can compute the price of this derivative by computing the expected present values of the cash flows in the risk neutral world. We obtain

$$\text{Price} = \tilde{p}e^{-r\delta t} \cdot 4 + (1-\tilde{p})\tilde{p}e^{-2d\delta t} \cdot 2 - (1-\tilde{p})^2 e^{-2d\delta t} \cdot 2.$$

The prescription for pricing is thus relatively straightforward in the risk neutral world. In fact it is as simple as taking an expectation over discounted cash flows. There is of course a gaping hole in that we have not proved that this is the right thing to do, and that it results in the arbitrage free price. For the moment we will just take for granted that this is the right thing to do, and explore more generally the process of pricing derivatives.

2.2 Pricing Derivatives in the Multiperiod Risk Neutral World

Assume that we have n periods, $\delta t, 2\delta t, \ldots, n\delta t$. A price path for the stock can be represented by a binary vector, $\mathbf{q} \in \{0,1\}^n$. Each 1 represents an increase in the stock price by a factor

λ_+ and each 0 in the path a decrease by a factor λ_-. For example, the vector $\mathbf{q} = [1, 0, 0, 1]$ correspondsto the stock path

$$S \to \lambda_+ S \to \lambda_+\lambda_- S \to \lambda_+\lambda_-^2 S \to \lambda_+^2\lambda_-^2 S.$$

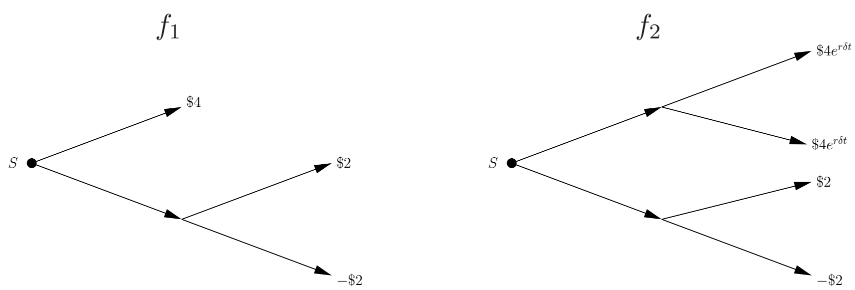
Exercise 2.4

As a function of n, how many possible paths are there?

We may now define a derivative formally as a function which assigns each path (there are 2^n paths) to a cashflow at time $n\delta t$. We denote a derivative function by $f : \{0, 1\}^n \mapsto \mathbb{R}$. Technically speaking, the derivative we considered in the previous section is not defined in this way. For example there was a cash flow which was received at time δt if the stock went up. Thus, in general one might expect a derivative to release cashflows along a stock price path, as opposed to only at the end. However, thanks to the existence of a risk free instrument, we can convert any cashflows along a path to an equivalent single cashflow at time $n\delta t$.

Exercise 2.5

Consider the following two period derivatives, and assume the bond grows in price at every time step by $e^{r\delta t}$.



Show using an arbitrage argument that the price of these two derivatives must be the same. Hence f_1 and f_2 are essentially the same derivative, except that in f_2 all cashflows come on the last timestep.

In the previous exercise, f_2 is an example of a *path dependent* derivative – different stock price paths leading to the same final price have different payoffs. This is the most general type of derivative that we will consider. The Asian option, the barrier option, the minimum option, the American option, etc. are all examples of path dependent options. In general, these are the hardest types of options to price, as we will see below (in terms of computational complexity).

The previous exercise shows that the derivative f_1 is equivalent to a derivative f_2 in which all cash flows occur at the last time step. In general this is the case, and so our notion of a derivative is general enough to accomodate intermediate cashflows.

Exercise 2.6

Show that in an economy with a risk free instrument, a derivative f_1 with cashflows occuring at intermediate times during a path is equivalent to some derivative f_2 with all cashflows occuring at the last time step.

There is one subtlety not addressed in our notion of a derivative, which is the notion of exercise, in the sense that the definition of a derivative assumes that to every path $\mathbf{q}$, a well defined predetermined cash flow $f(\mathbf{q})$ can be computed. However for an option like an American option, the cashflow along a particular path will depend on when the option is exercised, i.e., it will depend on the particular *exercise strategy* which is employed. Thus, technically speaking, an option with an exercise aspect is only a well specified derivative given a particular exercise strategy. To obtain the price of such an option, we may assume that the holder of the option will exercise optimally, and so to price something like an American option, one needs to compute the price of the resulting cashflow stream from optimal exercise. In such a case we have to simultaneously compute what the optimal exercise function is in order to obtain the cash flow stream. We will discuss the American option in depth later. For the moment, we will simply focus on derivatives with predetermined cashflows.

In the risk neutram framework, the pricing of an arbitrary derivative is conceptually simple, given by the expectation of the discounted cashflow. specifically, for a path $\mathbf{q}$ let $|\mathbf{q}| = \sum_i q_i$ be the number of up moves in the path. Then the probability of a path in the risk neutral world is given by $\tilde{P}[\mathbf{q}] = \tilde{p}^{|\mathbf{q}|}(1-\tilde{p})^{n-|\mathbf{q}|}$. Let $C(f)$ be the price of the derivative f. Then

$$\begin{aligned} C(f) &= e^{-nr\delta t} \sum_{\mathbf{q}} \tilde{P}[\mathbf{q}] f(\mathbf{q}), \\ &= e^{-nr\delta t} \sum_{\mathbf{q}} \tilde{p}^{|\mathbf{q}|}(1-\tilde{p})^{n-|\mathbf{q}|} f(\mathbf{q}). \end{aligned}$$

There are several problems with this formula, starting with the specification of the derivative function f itself. It must be specified for every possible path, hence there are 2^n numbers to be specified which is $O(2^n)$ in memory requirement. Even if the function f could be specified compactly as a function of the n inputs $\mathbf{q}$, the computation in the summation involves 2^n

terms which is an $O(2^n)$ computational complexity. For the general derivative, there is no way around this computational hurdle, and so there are only two ways to proceed. The first is to restrict the class of derivatives, the second is to consider approximate computations of this price. This price can be estimated to a very high accuracy very accurately using Monte Carlo techniques based on the fact that it is the expectation of some quantity according to some well defined probability distribution. We will discuss such issues later. For the moment, we will consider simpler classes of derivatives.

State Dependent Derivatives. Aside from the Asian option which is essentially path dependent, all the other common derivatives are state dependent. For this reason, Asian options are notoriously hard to price.

A state dependent derivative is one whose cashflow function f is a function of only the states of the stock visited along the path. Thus, a state dependent derivative is fully specified by a cash flow function $f(s,t)$, which specifies a cashflow $f(s,t)$ at the time t if the stock is at price s. The cashflow for a path $\mathbf{q}$ is the sum of all the cashflows accrued over the path.

$$f(\mathbf{q}) = \sum_{i=1}^{n} e^{(n-i)r\delta t} f(s_i(\mathbf{q}), i\delta t),$$

where $s_i(\mathbf{q})$ is the price at time $i\delta t$. The term $e^{(n-i)r\delta t}$ is the term required to convert a cashflow at time $i\delta t$ to one at time $n\delta t$. At time step i, the possible prices are

$$s_i \in \{\lambda_-^i S, \lambda_+\lambda_-^{i-1} S, \lambda_+^2\lambda_-^{i-2} S, \ldots, \lambda_+^j\lambda_-^{i-j} S, \ldots, \lambda_+^i S\}.$$

In other words, $s_i = \lambda_+^j \lambda_-^{i-j} S$, where j is the number of up moves in the path till time i. The state of the stock is summarized in the value j, and so we will use the more compact notation $f(j,i)$ to denote $f(s(j,i), i\delta t)$ where $s(j,i) = \lambda_+^j\lambda_-^{i-j} S$. The state of the process can be represented by the pair (j,i). Let $\mathbf{q}_i$ be the prefix of the path $\mathbf{q}$ up to time i, then $j = |\mathbf{q}_i|$. In this notation, we have that

$$f(\mathbf{q}) = \sum_{i=1}^{n} e^{(n-i)r\delta t} f(|\mathbf{q}_i|, i),$$

and so for the price of the derivative, we get that

$$\begin{aligned}
C(f) &= e^{-nr\delta t}\sum_{\mathbf{q}} \tilde{p}^{|\mathbf{q}|}(1-\tilde{p})^{n-|\mathbf{q}|} f(\mathbf{q}), \\
&= e^{-nr\delta t}\sum_{\mathbf{q}} \tilde{p}^{|\mathbf{q}|}(1-\tilde{p})^{n-|\mathbf{q}|} \sum_{i=1}^{n} e^{(n-i)r\delta t} f(|\mathbf{q}_i|, i), \\
&= \sum_{\mathbf{q}} \tilde{p}^{|\mathbf{q}|}(1-\tilde{p})^{n-|\mathbf{q}|} \sum_{i=1}^{n} e^{-ir\delta t} f(|\mathbf{q}_i|, i), \\
&= \sum_{i=1}^{n} e^{-ir\delta t} \sum_{\mathbf{q}} \tilde{p}^{|\mathbf{q}|}(1-\tilde{p})^{n-|\mathbf{q}|} f(|\mathbf{q}_i|, i), \\
&\overset{(a)}{=} \sum_{i=1}^{n} e^{-ir\delta t} \sum_{j=0}^{i} \binom{i}{j} \tilde{p}^{j}(1-\tilde{p})^{i-j} f(j, i),
\end{aligned}$$

where (a) follows from next exercise.

Exercise 2.7

Show that

$$\sum_{\mathbf{q}\in\{0,1\}^n} \tilde{p}^{|\mathbf{q}|}(1-\tilde{p})^{n-|\mathbf{q}|} g(|\mathbf{q}_i|) = \sum_{j=0}^{i} \binom{i}{j} \tilde{p}^{j}(1-\tilde{p})^{i-j} g(j),$$

for any function g defined on $\mathbb{R}$.

[Hint: Break a path into two parts, $\mathbf{q} = [\mathbf{q}_i, \mathbf{q}_l]$ and break up the sum over all paths into $i+1$ sums where in each sum you only consider those paths in which $|\mathbf{q}_i| = j$ for $j = 0, 1, \ldots, i$. Show that

$$\sum_{\mathbf{q}\in\{0,1\}^n : |\mathbf{q}_i| = j} \tilde{p}^{|\mathbf{q}|}(1-\tilde{p})^{n-|\mathbf{q}|} = \binom{i}{j} \tilde{p}^{j}(1-\tilde{p})^{i-j}.$$

In terms of efficiency, we see that state dependent derivatives are easy to represent, requiring only the specification of $f(j, i)$ for $i = 1 \ldots, n$ and $j = 0, \ldots i$ which is $O(n^2)$ memory. We also see that pricing a state dependent derivative only involves the computation of a double summation which is also $O(n^2)$.

Exercise 2.8

Compute exactly how many values of f are required to represent a state dependent derivative.

Compute exactly how many terms are in the summation for the price of a state dependent derivative.

Exercise 2.9

Assuming that basic arithmetic operations (multiplication, division, exponentiation) are constant time operations (which is not quite true) and that $f(j,i)$ can be accessed in constant time, construct an efficient quadratic time algorithm to compute $C(f)$ for a state dependent derivative.

Exercise 2.10

We consider an alternative derivation of the pricing formula for a state dependent derivative based on the facts that the price is an expectation of the discounted cashflow, and that the cashflow for a path is the sum of the cashflows along the path.

1. Show that the price $C(f) = \sum_{i=1:n} e^{-ir\delta t} c_i$, where c_i is the expected cashflow at timestep i.

   [Hint: Write $f(\mathbf{q})$ as $\sum_i f_i(\mathbf{q})$, where f_i is the cashflow for path $\mathbf{q}$ at time step i, and use the fact that the expectation of the sum of the cashflows is the sum of the expected cashflows.]
2. Show that

$$c_i = \sum_{j=0}^{i} \binom{i}{j} \tilde{p}^j (1-\tilde{p})^{i-j} f(j,i).$$

Exercise 2.11

We consider the European call option with maturity $n\delta t$ and strike K

(a) Show that this option is a state dependent derivative with

$$f(j,i) = \begin{cases} 0 & i < n, \\ (\lambda_+^j \lambda_-^{i-j} S - K)^+ & i = n. \end{cases}$$

[Hint: Assume optimal exercise.]

(b) Show that the price of this derivative is given by

$$C(S,K,n) = e^{-nr\delta t} \sum_{i=0}^{n} \binom{n}{i} \tilde{p}^i (1-\tilde{p})^{n-i} (\lambda_+^i \lambda_-^{n-i} S - K)^+.$$

One-Shot State Dependent Derivatives By and large most derivatives are one-shot in the sense that over the path of a stock's price, either there will be no cash flow (for example if an option is never exercised), or ther will be only one cash flow (for example after an option is exercised, there is an immediate cashflow, and no further cashflow).

For each path, there is only one point at which a cash flow will be realized and it is only state dependent, i.e., will equal $f(j,i)$. This will be the first state along the path for which the cashflow is non-zero. All paths which have the state (j,i) as their first non-zero cashflow state will realize the cashflow $f(j,i)$ at time $i\delta t$ (present value $e^{-ir\delta t} f(j,i)$) and no further cashflow.

In order to specify such a derivative, it is only necessary to specify the cashflows for the set of nodes which comprise the possible first points in paths for which there is a non-zero cashflow. One can show that such sets of nodes are minimal cuts in the lattice and are hence of linear size, hence to specify such a derivative, one only needs to specify $O(n)$ non-zero cashflows – and example is the European option. It is then possible to compute the price of such an option in linear time.

We will not elaborate more generally on algorithms for one-shot derivatives except to consider some examples.

Example: The European Call Option As described in an earlier exercise, the price of the European call option is given by

$$C(S,K,n) = e^{-nr\delta t} \sum_{i=0}^{n} \binom{n}{i} \tilde{p}^i (1-\tilde{p})^{n-i} (\lambda_+^i \lambda_-^{n-i} S - K)^+.$$

Clearly there are only a linear number of terms in the summation. Denote each summand by $x_i = p_i(s_i - K)^+$, where $p_i = \binom{n}{i} \tilde{p}^i (1-\tilde{p})^{n-i}$, $s_i = \lambda_+^i \lambda_-^{n-i} S$; The first summand

$x_0 = p_0(s_0 - K)^+$, with $p_0 = (1-\tilde{p})^n$ and $s_0 = \lambda_-^n S$. Since exponentiation can be done in logarithmic time, s_0, p_0 and hence x_0 can be computed in $O(\log n)$ time. We now show how to get (s_{i+1}, p_{i+1}) from (s_i, p_i) in constant time. The following simple recursion which we leave as an exercise establishes this fact,

---

Exercise 2.12

Show that

$$\begin{aligned} s_{i+1} &= \frac{\lambda_+}{\lambda_-} s_i, \\ p_{i+1} &= \left(\frac{n-i}{i+1}\right)\left(\frac{\tilde{p}}{1-\tilde{p}}\right) p_i, \end{aligned}$$

Exercise 2.13

Explain why the recursion in the previous exercise implies that the price of the European call option can be computed in linear time.

Exercise 2.14

The price for the algorithm obtained in the previous section results in a numerically very unstable algorithm because it invloves taking very large exponents. For example $p_0 = (1-\tilde{p})^n$ which for large n will be zero to machine precision. Thus it is numerically more stable to compute $\log p_i$ and $\log s_i$ and use these to construct the summands x_i.

(a) Give the appropriate recursions for $\log p_i$ and $\log s_i$.

(b) Using the answer to the previous part, give a linear time algorithm to price the Eurpoean call option.

Example: The Barrier Option We give one final example to illustrate some techniques for constructing linear time algorithms to price one-shot options.

2.2.1 The Multiperiod Risk Neutral Pricing Framework

We will now prove that the multiperiod risk neutral world in which prices of derivatives are obtained is exactly the independently iterated binomial model with the probability $\tilde{p}$ exactly as calculated in the two period case. We would like to argue that the risk neutral probabilities may be treated like ordinary probabilities. As such, they may be multiplied to obtain the probability of independent events. We illustrate with an example using the three period economy. Applying the risk neutral pricing at $t = 0$ we see that

$$S = e^{-r\delta t}(S_u\tilde{p} + S_d(1-\tilde{p}))$$

Applying the risk neutral pricing at $t = \delta t$ we see that $S_u = e^{-r\delta t}(S_{uu}\tilde{p} + S_{ud}(1-\tilde{p}))$ and $S_d = e^{-r\delta t}(S_{ud}\tilde{p} + S_{dd}(1-\tilde{p}))$ giving that

$$S = e^{-2r\delta t}(S_{uu}\tilde{p}^2 + 2S_{ud}\tilde{p}(1-\tilde{p}) + S_{dd}(1-\tilde{p})^2)$$

Further, using $S_{uu} = e^{-r\delta t}(S_{uuu}\tilde{p} + S_{uud}(1-\tilde{p}))$, $S_{ud} = e^{-r\delta t}(S_{uud}\tilde{p} + S_{udd}(1-\tilde{p}))$ and $S_{dd} = e^{-r\delta t}(S_{udd}\tilde{p} + S_{ddd}(1-\tilde{p}))$, we find that

$$S = e^{-3r\delta t}(S_{uuu}\tilde{p}^3 + 3S_{uud}\tilde{p}^2(1-\tilde{p}) + 3S_{udd}\tilde{p}(1-\tilde{p})^2 + S_{ddd}(1-\tilde{p})^3)$$

By inspecting this formula, we can pick out that $\tilde{p}_{u^3} = \tilde{p}^3$, $\tilde{p}_{u^2d} = 3\tilde{p}^2(1-\tilde{p})$, $\tilde{p}_{ud^2} = 3\tilde{p}(1-\tilde{p})^2$ and $\tilde{p}_{d^3} = (1-\tilde{p})^3$. It is apparent that the risk neutral probabilities at the N^{th} time period will be given by a binomial distribution of order N

$$\tilde{\mathbf{P}}\left[S(N) = S(0)u^k d^{N-k}\right] = \binom{N}{k}\tilde{p}^k(1-\tilde{p})^{N-K}$$

We do not prove this here, but a proof induction should be plausible at this point.

The *European call option* with expiry at time T and strike at K gives the holder the right (but not the obligation) to by the stock at time $t = T$ at the strike price K. Clearly the value of this option (and hence its price) at time T is given by $(S(T)-K)\Theta(S(T)-K)$, where $\Theta(x) = 0$ for $x < 0$ and $\Theta(x) = 1$ otherwise. Suppose that we consider the time interval $[0, T]$ broken into N smaller intervals of size $\delta t = T/N$. If δt is small enough, it is reasonable to assume that the stock could only evolve into an up or down state during the time interval. In which case, the multi-period binary lattice dynamics would apply and the price of the call option at time $t = 0$, $C(0)$, is given by taking the discounted expectation with respect to the risk neutral measure given in (2.2.1). Thus,

$$\begin{aligned} C(0) &= C(S(0), K, r, \sigma, \mu, T) = e^{-rT}E_{\tilde{\mathbf{P}}}[(S(T)-K)\Theta(S(T)-K)] \\ &= e^{-rT}\sum_{k=0}^{N}\binom{N}{k}\left(S(0)\lambda_+^k\lambda_-^{N-k} - K\right)\tilde{p}^k(1-\tilde{p})^{N-K}\Theta\left(S(0)\lambda_+^k\lambda_-^{N-k} - K\right) \end{aligned}$$

If we insist on using a discrete time representation of our market, then there is no option but to compute this expectation numerically. The natural approach would be to use Monte Carlo simulations for more complex dynamics. A key issue is the efficiency of the computation. Given μ, σ and assuming that a proxy for the interest rate r exists, one only needs to compute a single expectation. Often μ and σ are not known. In fact, one needs to *calibrate* the model to the observed option data to extract μ and σ. This usually involves an optimization in the μ and σ space, which is extremely computationally intensive.

We note that the μ dependence has disappeared due to some "lucky" cancellations. We could have expected this, as the risk neutral dynamics (**??**) do not contain μ dependence. Up to now, we have not invoked any stochastic calculus. The limit of the dynamics in (**??**) can be represented by the Ito stochastic differential equation

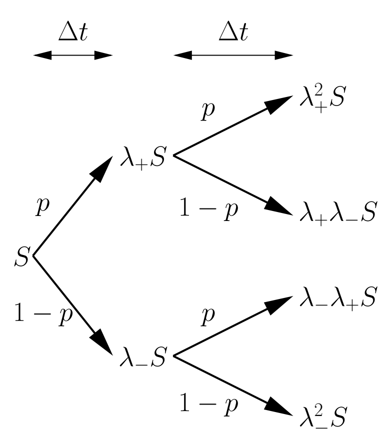
$$dS = rSdt + \sigma SdW$$

where (informally) dW is a random variable with variance dt. A standard application of Ito's lemma (see for example [**?**]) immediately yields (**??**), from which (**??**) follows by taking the risk neutral expectation.

There are many approaches to obtaining the price of the European call for the dynamics that we have chosen. For other dynamics, such closed form solutions are generally not possible, and one can resort to obtaining the prices using Monte Carlo techniques to take risk neutral expectations.

Exercise 2.15

Consider the following stock dynamics according to our *binomial* model with probability p.



A bond has values B, $Be^{r\Delta t}$ and $Be^{2r\Delta t}$ at the three times, where $e^{r\Delta t} = \frac{5}{4}$. There is a variant of the call option, instrument I, with current price C. The strike is $K = \frac{S}{4}$ and if the stock initially goes down, you *must* exercise, getting a cashflow $\lambda_- S - K$ at time Δt. If the stock initially goes up, then you *must* wait and exercise at the end of the next period.

The stock dynamics are given by $p = \frac{1}{2}$, $\lambda_+ = \frac{3}{2}$, $\lambda_- = \frac{1}{2}$. There are two possible prices $(P,\ \tilde{P})$ that one could select for instrument I. P is the expected value of the discounted future cashflows according to the real dynamics, p. $\tilde{P}$ is the expected value of the discounted future cashflows according to the risk neutral (fictitious pricing) dynamics, $\tilde{p} = \frac{e^{r\Delta t} - \lambda_-}{\lambda_+ - \lambda_-}$.

(a) Compute $P,\ \tilde{P}$.

(b) If the price is actually P, then then show that there is an arbitrage opportunity – i.e., construct the arbitrage opportunity.

(c) What happens to your arbitrage opportunity if the price were $\tilde{P}$.

(d) Show that if the price is not $\tilde{P}$, then there is an arbitrage opportunity.

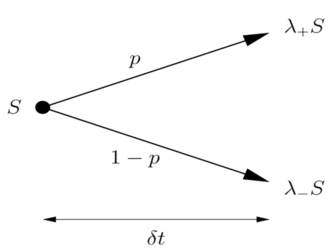
[Hint: You may want to consider what the price of I should be at time Δt if the stock went up. Use this knowledge to construct an arbitrage *strategy*.

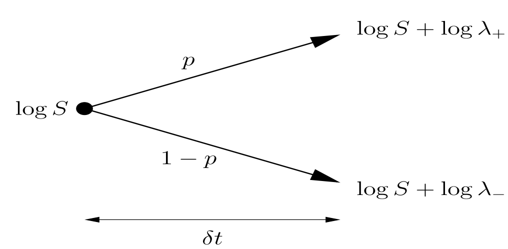
Note: that in multi-period economies, we have to generalize the concept of an arbitrage protfolio to an arbitrage strategy, where one is allowed to change the portfolio according to the stock path. As long as at any time, there is no nett investment using the strategy, and the strategy always yields, at time $2\Delta T$, non-negative and sometimes positive return, then there is arbitrage.]

2.3 The Real World, The Risk Neutral World and Geometric Brownian Motion

We now return to the binomial model with a view towards linking it to the real stock price world in which observations can be made about the real world dynamics. These measurements should in turn imply the parameters of the binomial model, which corresponds to *callibrating* the binomial world dynamics to the real world dynamics.

We begin by moving into an equivalent binomial model, namely the log-price world, illustrated below,





Note that these two models yield identical stock price dynamics. The main difference is that the log-price world is easier to analyze because it is additive. Imagine now iterating forward for n time steps to some time increment $\Delta t = n\delta t$. The idea is that δt is infinitessimal and will tend to zero, $\delta t \to 0$ with Δt fixed, and so $n = \frac{\Delta t}{\delta t} \to \infty$. At each time step i, the log price $\log S$ will change by some amount x_i, so

$$\log S(\Delta t) = \log S + \sum_{i=1}^{n} x_i,$$

where the x_i are independent identically distributed random variables with

$$x_i = \begin{cases} \log \lambda_+ & \text{w.p. } p, \\ \log \lambda_- & \text{w.p. } 1-p. \end{cases}$$

Exercise 2.16

Show that

$$\begin{aligned} E[x_i] &= p \log \lambda_+ + (1-p) \log \lambda_-, \\ Var[x_i] &= p(1-p)(\log \lambda_+ - \log \lambda_-)^2. \end{aligned}$$

Thus, $\Delta \log S = \sum_{i=1}^{n} x_i$ is the sum of a large number of iid random variables each of whose mean and variance is given by the previous exercise. We will now argue informally, assuming that the conditions for the central limit theorem hold (which in reality is true). By the central limit theorem, $\frac{1}{n}\Delta \log S = \frac{1}{n}\sum_{i=1}^{n} x_i$ converges in distribution to a Normal random variable[5] with mean $E[x_i]$ and variance $Var[x_i]$. We thus have that

$$S(\Delta t) = Se^{\eta},$$

where η has a Normal distribution,

$$\eta \sim N(\mu, \sigma^2),$$

[5] A random variable X has a normal distribution with mean μ and variance σ^2, written $X \sim N(\mu, \sigma^2)$ is its probability density function (pdf) is given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

.

with

$$\begin{aligned}
\mu &= \lim_{n\to\infty} n(p\log\lambda_+ + (1-p)\log\lambda_-), \\
&= \lim_{\delta t\to 0} \frac{\Delta t}{\delta t}(p\log\lambda_+ + (1-p)\log\lambda_-); \\
\sigma^2 &= \lim_{n\to\infty} np(1-p)(\log\lambda_+ - \log\lambda_-)^2, \\
&= \lim_{\delta t\to 0} \frac{\Delta t}{\delta t} p(1-p)(\log\lambda_+ - \log\lambda_-)^2.
\end{aligned}$$

Clearly these expressions only make sense when these limits exist, which means that for these limits to exist, $\lambda_\pm$ should have a particular dependence on δt in this limit.

Exercise 2.17

Suppose that if in the limit we require that $\mu \to \mu_R\Delta t$ and $\sigma^2 \to \sigma_R^2\Delta t$.
Show that the dependence of $\lambda_\pm$ on δt is given by

$$\begin{aligned}
\lambda_+ &= e^{\mu_R\delta t + \sigma_R\sqrt{\frac{1-p}{p}}\sqrt{\delta t}}, \\
\lambda_- &= e^{\mu_R\delta t - \sigma_R\sqrt{\frac{p}{1-p}}\sqrt{\delta t}}.
\end{aligned}$$

The stock dynamics $S(t+\Delta t) = S(t)e^\eta$, where $\eta \sim N(\mu_R\Delta t, \sigma_R^2\Delta t)$ is known as a *Geometric Brownian Motion (GBM)* with drift parameter μ_R and volatility parameter σ_R. We see that in the limit as $\delta t \to 0$ the binomial model approaches a geometric Brownian motion.

The geometric Brownian motion is a very popular model for a stock's dynamics. In particular, it implies that the stock prices at some future time are lognormally distributed. This is a reasonable approximation to reality when it comes to pricing, however it has been widely observed that prices have a fatter tail than the lognormal distribution. This has important implications for tasks such as risk analysis, however we will continue using the geometric Brownian motion as our model of a stocks dynamics for pricing purposes. Typically the GBM parameters of the real world's stock price dynamics (μ_R, σ_R) can be measured by observing the return time series of a stock. Assume that μ_R, σ_R are given. The result of the previous exercise shows how to calibrate the parameters of the "real world" binomial model so that in the $\delta t \to 0$ the real world's GBM is recovered with parameters μ_R, σ_R. In particular, we see that

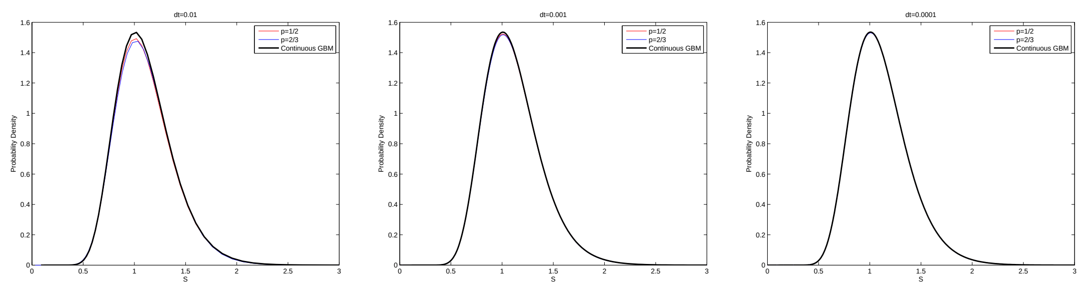
$$\begin{aligned}
\lambda_+ &= e^{\mu_R\delta t + \sigma_R\sqrt{\frac{1-p}{p}}\sqrt{\delta t}}, \\
\lambda_- &= e^{\mu_R\delta t - \sigma_R\sqrt{\frac{p}{1-p}}\sqrt{\delta t}}.
\end{aligned}$$

Note that the "real world" binomial model has three parameters (three degrees of freedom), $p, \lambda_\pm$, and as can be seen in the above formulas, in the $\delta t \to 0$ limit, the GBM has only two degrees of freedom, and so one of the binomial model's degrees of freedom is spurious. Thus we are free to choose one of the parameters more or less arbitrarily, for example we could set $p = \frac{1}{2}$.

Exercise 2.18

Let $\mu_R = 0.07$ and $\sigma_R = 0.25$ and set $\delta t = 0.0001$. Starting with an initial stock price $S_0 = 1$, simulate forward to time $T = 1$ the stock price for the binomial model in which we set $p = \frac{1}{2}$, repeating many times to obtain the distribution for the stock price at time 1. Now repeat for a choice of $p = \frac{2}{3}$ and compare the resulting two distributions.

[Answer:



]

The result of the previous exercise be convincing that one parameter, for example p is redundant and could be specified arbitrarily. We now consider the risk neutral pricing world, in which $\lambda_\pm$ are unchanged, but $\tilde{p} = (e^{r\delta t} - \lambda_-)/(\lambda_+ - \lambda_-)$ and $1 - \tilde{p} = (\lambda_+ - e^{r\delta t})/(\lambda_+ - \lambda_-)$. We now compute the geometric Brownian motion corresponding to this risk neutral world. We need to compute

$$
\begin{aligned}
\tilde{\mu} &= \lim_{\delta t \to 0} \frac{1}{\delta t}(\tilde{p} \log \lambda_+ + (1 - \tilde{p}) \log \lambda_-); \\
\tilde{\sigma}^2 &= \lim_{\delta t \to 0} \frac{1}{\delta t} \tilde{p}(1 - \tilde{p})(\log \lambda_+ - \log \lambda_-)^2.
\end{aligned}
$$

We first consider $\tilde{\mu}$.

$$
\begin{aligned}
\tilde{\mu} &= \lim_{\delta t \to 0} \frac{1}{\delta t}(\tilde{p} \log \frac{\lambda_+}{\lambda_-} + \log \lambda_-), \\
&= \lim_{\delta t \to 0} \frac{1}{\delta t} \left(\frac{\tilde{p}\sigma_R \sqrt{\delta t}}{\sqrt{p(1-p)}} + \mu_R \delta t - \sigma_R \sqrt{\frac{p}{1-p}} \sqrt{\delta t} \right).
\end{aligned}
$$

Exercise 2.19

Show that

$$\begin{aligned}
\tilde{p} &= \frac{(r-\frac{1}{2}\sigma_R^2-\mu_R)\delta t+\sigma_R\sqrt{\frac{p}{1-p}}\sqrt{\delta t}+O(\delta t^{3/2})}{\frac{\sigma_R\sqrt{\delta t}}{\sqrt{p(1-p)}}}, \\
1-\tilde{p} &= \frac{\sigma_R\sqrt{\frac{1-p}{p}}\sqrt{\delta t}-(r-\frac{1}{2}\sigma_R^2-\mu_R)\delta t+O(\delta t^{3/2})}{\frac{\sigma_R\sqrt{\delta t}}{\sqrt{p(1-p)}}}, \\
\tilde{p}(1-\tilde{p}) &= \frac{\sigma_R^2\delta t+O(\delta t^{3/2})}{\frac{\sigma_R^2\delta t}{p(1-p)}}.
\end{aligned}$$

Using the result of the previous exercise, we conclude that

$$\begin{aligned}
\tilde{\mu} &= \lim_{\delta t\to 0}\frac{1}{\delta t}\left((r-\frac{1}{2}\sigma_R^2)\delta t+O(\delta t^{3/2})\right), \\
&= r-\frac{1}{2}\sigma_R^2.
\end{aligned}$$

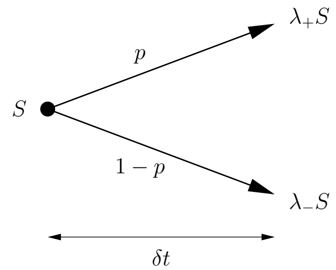
Using the result of the previous exercise, we perform a similar analysis for $\tilde{\sigma}^2$ to obtain

$$\begin{aligned}
\tilde{\sigma}^2 &= \lim_{\delta t\to 0}\frac{1}{\delta t}\tilde{p}(1-\tilde{p})\log^2\frac{\lambda_+}{\lambda_-}, \\
&= \lim_{\delta t\to 0}\frac{1}{\delta t}\tilde{p}(1-\tilde{p})\frac{\sigma_R^2\delta t}{p(1-p)}, \\
&= \lim_{\delta t\to 0}\frac{1}{\delta t}(\sigma_R^2\delta t+O(\delta t^{3/2})), \\
&= \sigma_R^2.
\end{aligned}$$

Summary We are now ready to summarize our conclusions. Assume that in the real world, the stock dynamics is given by a geometric Brownian motion,

$$S(t+\Delta t)=S(t)e^{\eta_R},$$

where η_R is Normally distributed, $\eta_R\sim N(\mu_R,\sigma_R^2)$ (μ_R,σ_R^2) are typically measured from the data. This geometric Brownian motion can be represented as the $\delta t\to 0$ limit of the real world binomial model,



in which the parameters $\lambda_\pm$ are given by

$$\begin{aligned}\lambda_+ &= e^{\mu_R \delta t + \sigma_R \sqrt{\frac{1-p}{p}}\sqrt{\delta t}},\\ \lambda_- &= e^{\mu_R \delta t - \sigma_R \sqrt{\frac{p}{1-p}}\sqrt{\delta t}}.\end{aligned}$$

and p is arbitrary.

The risk neutral world is given by a similar dynamics, with the same parameters for $\lambda_\pm$, but $\tilde{p}$ is no longer arbitrary. Instead, $\tilde{p} = (e^{r\delta t} - \lambda_-)/(\lambda_+ - \lambda_-)$. Taking the $\delta t \to 0$ limit of the binomial model for the risk neutral world, we obtain the geometric Brownian motion corresponding to the risk neutral pricing world, namely

$$S(t + \Delta t) = S(t)e^{\tilde{\eta}}$$

where $\eta \sim N(r - \frac{1}{2}\sigma_R^2, \sigma_R^2)$. Notice that the drift in the risk neutral world is independent of the drift μ_R in the real world.

Pricing in the Risk Neutral World We may use either the binomial version of the risk neutral world or the geometric Brownian motion version of the risk neutral world for pricing. In either case, we need to compute the expectation of the discounted cashflows for a derivative. In many cases, this can be done efficiently in the binomial model, especially for state dependent derivatives. Typically the algorithm will be $O(n^2)$. For one-shot derivatives this may often be improved to $O(n)$. One should take n to be as large as is numerically feasible which for state dependent derivatives in the general case is means that one cannot go much beyond $n = 10,000$ if the price is required in a matter of seconds. For one-shot derivatives for which one can construct an efficient linear time algorithm, one can take n very large, even in the million range and still get a price in a matter of seconds. For n approaching a million, one essentially has the exact price. In general, the price converges to the GBM price with an error that is $O(\frac{1}{n})$.

In general in the binomial world it is not possible to get an analytic closed form solution to the price of a derivative. However, in the GBM risk neutral world, if the option is simple enough, it may be possible to get the closed form solution. We will illustrate with two examples. Note that for purposes of closed form solutions, the error function $erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x ds\ e^{-s^2}$ or the standard normal cumulative distribution function $\phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^x ds\ e^{-\frac{1}{2}s^2}$ are considered closed form.

Example: Analytic Price of The European Call Option in the GBM Model The cashflow is given by $f(s,t) = (s-K)^+$ for $t = T$ and 0 for $t < T$. We need to compute the expected discounted cashflow in the risk neutral world, which is

$$E[e^{-rT}(S(T)-K)^+].$$

In the risk neutral world, $S(T) = Se^{\tilde{\eta}}$, with $\eta \sim N((r-\frac{1}{2}\sigma_R^2)T, \sigma_R^2 T)$, so we have that

$$\begin{aligned} C(S,K,T) &= E[e^{-rT}(S(T)-K)^+], \\ &= e^{-rT}\int_{-\infty}^{\infty} d\eta\ \frac{1}{\sqrt{2\pi\sigma_R^2 T}} e^{-\frac{(\eta-(r-\frac{1}{2}\sigma_R^2)T)^2}{2\sigma_R^2 T}} (Se^{\eta}-K)^+, \\ &= e^{-rT}\int_{\log\frac{K}{S}}^{\infty} d\eta\ \frac{1}{\sqrt{2\pi\sigma_R^2 T}} e^{-\frac{(\eta-(r-\frac{1}{2}\sigma_R^2)T)^2}{2\sigma_R^2 T}} (Se^{\eta}-K). \end{aligned}$$

After some algebraic manipulation (left as an exercise) using changes of variables in the final integration, we arrive at the celebrated *Black-Scholes* formula for the price of the European call option.

Exercise 2.20

Manipulate the integral for the price of the European call option to write it using the standard normal distribution function to obtain the Black-Scholes formula,

$$C(S,K,T;r,\sigma_R) = S\phi(d_1) - Ke^{-rT}\phi(d_2),$$

where

$$d_1 = \frac{\log\frac{S}{K} + (r+\frac{1}{2}\sigma_R^2)T}{\sqrt{\sigma_R^2 T}} \qquad d_2 = \frac{\log\frac{S}{K} + (r-\frac{1}{2}\sigma_R^2)}{\sqrt{\sigma_R^2 T}}.$$

[Hint: Use a change of variables to $u = (\eta-(r-\frac{1}{2}\sigma_R^2)T)/\sigma\sqrt{T}$ and separate the integral into two terms. You will then need to complete the square in one of the terms. Also note that $\frac{1}{\sqrt{2\pi}}\int_x^{\infty} ds\ e^{-\frac{1}{2}s^2} = 1-\phi(x) = \phi(-x)$.]

Black and Scholes arrived at this formula through a much more difficult path to arbitrage free pricing using stochastic differential equations and solving the resulting PDE that the price of the European call option should satisfy. Given the risk neutral pricing framework, our approach is conceptually much simpler. Nevertheless the PDE approach is interesting and leads to efficient algorithms, which we will discuss later.

One important thing to notice about this pricing formula is that there is no dependence on the real world drift, only on the real world volatility. In this sense, speculative trading of options is sometimes refered to as trading volatility – if I have a better estimate of the volatility than my counterparty, then I can make money off them because I will have the more accurate price.

Estimating the Real World Parameters μ_R, σ_R^2. We briefly mention how to estimate μ_R and σ_R^2. Given a historical price time series of a stock, $S_0, S_1, S_2, \ldots, S_n$ at times $t_0, t_1, t_2, \ldots, t_n$, we consider the logarithmic differences, $\Delta \log S_i = \log \frac{S_i}{S_{i-1}}$ and the time increments $\Delta t_i = t_i - t_{i-1}$ for $i = 1, \ldots, n$.

We develop a maximum likelihood approach to estimating μ_R, σ_R^2. We have that $\Delta \log S_i$ are *i.i.d.* random samples where each $\Delta \log S_i$ has a normal distribution, $\Delta \log S_i \sim N(\mu_R \Delta t_i, \sigma_R^2 \Delta t_i)$. We compute the log-likelihood as

$$
\begin{aligned}
\log P[\{\Delta \log S_i\}_{i=1}^n | \mu_R, \sigma_R^2] &= \log \left[\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_R^2 \Delta t_i}} e^{-\frac{(\Delta \log S_i - \mu_R \Delta t_i)^2}{\sigma_R^2 \Delta t_i}} \right], \\
&= -\left(\frac{1}{\sigma_R^2} \sum_{i=1}^n \frac{(\Delta \log S_i - \mu_R \Delta t_i)^2}{\Delta t_i} + \frac{n}{2} \log \sigma_R^2 \right) + \text{constant},
\end{aligned}
$$

In order to obtain the maximum likelihood estimates of μ_R, σ_R^2, we ma maximize the log-likelihood, which may be accomplished by setting the partial derivatives with respect to μ_R and σ_R^2 to zero. The final result is given as an exercise.

Exercise 2.21

Show that the maximum likelihood estimates of μ_R and σ_R^2 are given by

$$
\begin{aligned}
\mu_R &= \frac{\log S_n - \log S_0}{t_n - t_0}, \\
\sigma_R^2 &= \frac{1}{n} \sum_{i=1}^n \frac{(\log S_i - \log S_{i-1})^2}{t_i - t_{i-1}} - \frac{(\log S_n - \log S_0)^2}{n(t_n - t_0)}.
\end{aligned}
$$

3 The Monte Carlo Approach to Pricing Derivatives

As we have seen, in some cases, it is possible to derive the price of a derivative analytically by taking the expectation of the discounted cashflows with respect to the risk neutral measure in the geometric brownian motion model. We gave the European call option and the barrier option as examples. In general, however, this is not possible, and Monte Carlo simulation is an alternative approach to the problem.

The fact that the price is the expected discounted cashflows in the risk neutral world provides a very conceptually simple algorithm for approximating the price by virtue of the fact that a sample average over independently generated samples provides a good randomized approximation to the expectation. Specifically, if we could generate sample paths independently according to the probabilities in the risk neutral world, then for a single path, the cash flows can be computed as a deterministic cash flow stream. We know how to compute the present value of a deterministic cash flow stream. So by generating a path we generate a sample (according to the risk neutral probabilities) of the discounted cash flows. Repeating this many times and taking the average yields a Monte Carlo estimate of the expected discounted cashflows. This Monte Carlo pricing algorithm is very general and can be applied to *any* derivative whose cash flow function has been specified. The high level view of the algorithm is summarized below.

Obtain μ_R, σ_R^2 for the real world geometric Brownian motion.
Compute the Risk Neutral World dynamics.
for $i = 1$ to $Npaths$ **do**
    Generate a stock price path p_i according to the Risk Neutral World dynamics.
    Compute the discounted cash flows along path p_i.
    Update average discounted cash flow.
end for
return Average discounted cash flow over the $Npaths$ paths.

There are two ways to generate a stock price path. The first is to obtain the risk neutral binomial model for some discrete time step δt. The second is to use the risk neutral geometric Brownian motion and use some time resolution Δt to generate stock prices

$$S(0), S(\Delta t), S(2\Delta t), \ldots,$$

where for $i \geq 1$,

$$S(i\Delta t) = S((i-1)\Delta t)e^{\eta_i},$$

with $\eta_i \sim N((r - \frac{1}{2}\sigma_R^2)\Delta t, \sigma_R^2 \Delta t)$.

Exercise 3.1

Consider the following options on a stock S (in all cases S_0 is the initial stock price):

1. $C(S_0, K, T)$: the *European Call* option with strike K and exercise time T.
2. $B(S_0, K, B, T)$: the *Barrier option* with payoff K, barrier B and horizon T. If and when the price hits the barrier B the holder may buy at B and sell at K.
3. $A(S_0, T)$: the *average strike Asian call option* with expiry T – a European call option with strike at the average value of the stock price over $[0, T]$.
4. $M(S_0, T)$: the *minimum strike call option* – a European call option with strike at the minimum price over $[0, T]$.

Use Monte Carlo simulation to price the 4 options. Assume that $S_0 = 1,\ \mu = 0.07,\ r = 0.03,\ \sigma = 0.2,\ T = 2$. For each case, use each of the three modes above, and compute the price using each of the three time discretizations, $\Delta t = 0.1, 0.01, 0.0001$.

In all cases, make some intelligent choice for the number of Monte Carlo samples that you need to take to get an accurate price. [Hint: first take a few samples to get an estimate of the variance of the Monte Carlo sample values.]

(a) Compute $C(1, 1, 2)$ as efficiently as you can and compare with the analytic formula.

(b) Compute $B(1, 1, 0.95, 2)$

(c) Compute $A(1, 2)$, for three possible definitions of "average": the harmonic, arithmetic, and geometric means. Explain the relative ordering of these prices.

(d) Compute $M(1, 2)$.

4 Stochastic Differential Equations and Alternate Representations of the Geometric Brownian Motion

Exercise 4.1

Assume the initial stock price is S_0 and it follows real and risk neutral dynamics given by

$$\Delta S = \mu S \Delta t + \sigma S \Delta W \qquad\qquad \Delta\tilde{S} = r\tilde{S}\Delta t + \sigma\tilde{S}\Delta\tilde{W}.$$

Write a program that takes as input $\mu,\ r,\ \sigma,\ S_0,\ T, \Delta T$ and simulates the stock price from time 0 to T in time steps of Δt for the risk neutral world, using each of the following modes:

(a) Binomial mode I: compute $\lambda_\pm$ from μ, σ assuming that $p = \frac{1}{2}$, and then computing $\tilde{p}$.

(b) Binomial mode II: compute $\lambda_\pm$ from μ, σ assuming that $p = \frac{2}{3}$, and then computing $\tilde{p}$.

(c) Continuous mode: using the continuous risk neutral dynamics r, σ generate at time step Δt as if the discrete model were taken to the limit $dt \to 0$.

For each of the three methods, give plots of representative price paths for $S_0 = 1,\ \mu = 0.07,\ r = 0.03,\ \sigma = 0.2,\ T = 2$ using $\Delta t = 0.1, 0.01, 0.0001$.

5 The PDE Approach to Arbitrage Free Pricing