1 From Binomial Model to Random Walk

The essence and the simplicity of the binomial model are the following: from time $t_n$ to the next time level $t_{n+1}$, the stock price can only have two possible moves - an up move when the stock price is multiplied by $u > 1 + r\Delta t$, or a down move when the stock price is multiplied by $d = 1/u$. In order for the model to have a local variance that matches the stock volatility $\sigma$, we require $u = \exp(\sigma\sqrt{\Delta t})$. Under the risk-neutral probability measure, the above is expressed as

$$S_{n+1} = S_n e^{\sigma\sqrt{\Delta t}\tilde{X}_{n+1}},$$

where the independent, identically distributed (i.i.d.) random variables

$$\tilde{X}_{n+1} = \begin{cases} 1 & \text{if } \omega_{n+1} = H, \\ -1 & \text{if } \omega_{n+1} = T, \end{cases}$$

with the head and tail probabilities $\tilde{p}$ and $\tilde{q} = 1 - \tilde{p}$, respectively. The immediate advantage of this notation is that we can represent

$$\log \left( \frac{S_{n+1}}{S_n} \right) = \sigma \Delta t \tilde{X}_{n+1},$$

and recursively this leads to

$$S_n = S_{n-1} e^{\sigma\sqrt{\Delta t}\tilde{X}_n} = (S_{n-2} e^{\sigma\sqrt{\Delta t}\tilde{X}_{n-1}}) e^{\sigma\sqrt{\Delta t}\tilde{X}_n} = \ldots = S_0 e^{\sigma\sqrt{\Delta t} \sum_{j=1}^{n} \tilde{X}_j}.$$  

If we introduce the random walk $\tilde{M}_n = \sum_{j=1}^{n} \tilde{X}_j$, we can represent our binomial model in terms of a random walk:

$$S_n = S_0 e^{\sigma\sqrt{\Delta t}\tilde{M}_n}.$$  

Unfortunately, $\tilde{M}_n$ is not a symmetric random walk since the risk-neutral probabilities

$$P\{\omega_n = H\} = \tilde{p} = \frac{1}{2} \left( 1 + \left( \frac{T}{\sigma} - \frac{\sigma}{2} \right) \sqrt{\Delta t} \right) \neq \frac{1}{2}, \quad P\{\omega_n = T\} = \tilde{q} = 1 - \tilde{p} \neq \frac{1}{2}.$$
To come up with a symmetric random walk, it is necessary to introduce a change of measure, and use the corresponding Radon-Nikodým derivative to convert the trend effect from the asymmetric random walk to a nonzero drift combined with a symmetric random walk. The purpose is to separate a martingale from the process so we can use these powerful tools associated with martingales.

2 Change of Measure

Consider the following intuitive reasoning: if the random moves of the stock price are more likely to be up than down, on average the stock is going up. Then why can’t we split the move into two steps: a deterministic upward component, and a random component that is equal likely to be up or down. Under this new measure, we expect head and tail states to have the same probability \( p^* = q^* = 1/2 \). In terms of Radon-Nikodým, we can introduce a process \( Z_n(\omega_1, \omega_2, \ldots, \omega_n) \) in which

\[
Z_{n+1} = \begin{cases} \frac{\tilde{p}}{p^*} = 1 + \left( \frac{r}{\sigma} - \frac{\sigma}{2} \right) \sqrt{\Delta t} & \text{if } \omega_{n+1} = H, \\
\frac{\tilde{q}}{q^*} = 1 - \left( \frac{r}{\sigma} - \frac{\sigma}{2} \right) \sqrt{\Delta t} & \text{if } \omega_{n+1} = T.
\end{cases}
\]

(7)

Equivalently, we can write

\[
\frac{Z_{n+1}}{Z_n} = 1 + \left( \frac{r}{\sigma} - \frac{\sigma}{2} \right) \sqrt{\Delta t} X_{n+1},
\]

(8)

and the Radon-Nikodým derivative process

\[
Z_n(\omega_1, \omega_2, \ldots, \omega_n) = \prod_{j=1}^{n} \left( 1 + \left( \frac{r}{\sigma} - \frac{\sigma}{2} \right) \sqrt{\Delta t} X_j \right).
\]

(9)

According to the change of measure relation (3.2.5) in the textbook,

\[
\mathbb{E}_n[Y] = \frac{1}{Z_n} \mathbb{E}^*_n [YZ_{n+1}],
\]

(10)

we should multiply \( Z_{n+1}/Z_n \) to \( Y = \log(S_{n+1}/S_n) \) if we express the expectation of \( Y \) under \( Q \) as an expectation under the measure \( Q^* \). Notice the modified

\[
Y^* = Y \frac{Z_{n+1}}{Z_n} = \begin{cases} (r - \frac{\sigma^2}{2}) \Delta t + \sigma \sqrt{\Delta t} & \text{if } \omega_{n+1} = H, \\
(r - \frac{\sigma^2}{2}) \Delta t - \sigma \sqrt{\Delta t} & \text{if } \omega_{n+1} = T,
\end{cases}
\]

(11)

under the new measure where head and tail states have the same probability 1/2. If we introduce \( X_n \) similar to Eq.(??), but with \( p^* = q^* = 1/2 \) so the resulting random walk is symmetric, then

\[
Y^* = \log \left( S_{n+1}^* \right) = (r - \frac{\sigma^2}{2}) \Delta t + \sigma \sqrt{\Delta t} X_{n+1}.
\]

(12)

Notice the difference between Eq.(??) where there is a deterministic part but the random walk is symmetric, and the previous definition Eq.(??) where there is no deterministic term but the random walk is asymmetric.
3 Stock Price as Described by Random Walk

From now on, we deal with the symmetric random walk based on $X_n$, and refer our risk-neutral measure to this new measure, with the obvious advantage that we will be dealing with a martingale

$$M_n = \sum_{j=1}^{n} X_j,$$

and the stock price can be represented as

$$S_n = S_{n-1}e^{(r - \frac{\sigma^2}{2})\Delta t + \sigma \sqrt{\Delta t} X_n} = S_0 \exp \left[ (r - \frac{\sigma^2}{2})n\Delta t + \sigma \sqrt{\Delta t} M_n \right].$$

(14)

We are particularly interested in the distribution of $S_N$ when $N$ is large, $\Delta t$ is small, such that $N\Delta t = T$, where $T$ is a fixed time horizon. Let us consider the distribution of $\log(S_N/S_0)$ first:

$$\log \left( \frac{S_N}{S_0} \right) = (r - \frac{\sigma^2}{2})T + \sigma \sqrt{T} \left( \frac{1}{\sqrt{N}} M_N \right).$$

(15)

Here we introduced the scaling $\sqrt{N}$ for a reason soon to be revealed: that is to set up the stage for an application of the central limit theorem. Since $M_N$ is a sum of $i.i.d.$ (independent identically distributed) random variables, the central limit theorem says that $\frac{1}{\sqrt{N}} M_N$ will converge to a standard normal random variable (standard means that the mean is 0 and the variance is 1), in distribution, as $N \to \infty$. It follows that for large $N$, the distribution of $\log(S_N/S_0)$ is approximately normal, with mean $(r - \frac{1}{2}\sigma^2)T$ and variance $\sigma^2 T$.

We have established the fundamental notion that the binomial model leads to a distribution for the stock price that is approximately normal, thus the name of lognormal distribution. On the other hand, we would like to view the model as a dynamics model for the stock return over an infinitesimal period of time $\Delta t$, as it will become the discretized version of the Black-Scholes model, to be studied next semester. To be more specific, we need to derive $(S_{n+1} - S_n)/S_n$ in terms of $\log(S_{n+1}/S_n)$. From calculus, we know that these two are very close if they are both small. Let us see if it is still the case here. From Taylor’s series expansion, for small $|x| < 1$,

$$\log(1 + x) = x - \frac{1}{2} x^2 + \frac{1}{3} x^3 \ldots$$

(16)

We now use this to approximate

$$\log \left( \frac{S_{n+1}}{S_n} \right) = \log \left( 1 + \frac{S_{n+1} - S_n}{S_n} \right) = \log \left( 1 + \frac{\Delta S_n}{S_n} \right),$$

(17)

here $\Delta S_n = S_{n+1} - S_n$. Let $x = \Delta S_n/S_n$,

$$\log \left( \frac{S_{n+1}}{S_n} \right) = \frac{\Delta S_n}{S_n} - \frac{1}{2} \left( \frac{\Delta S_n}{S_n} \right)^2 + \frac{1}{3} \left( \frac{\Delta S_n}{S_n} \right)^3 + \ldots$$

(18)
How many terms should we keep? We are assuming small $\Delta t$ so it is reasonable to keep terms up to $O(\Delta t)$. Does this mean we can drop terms of order $(\Delta S_n S_n^*)^2$ and higher?

Let us take a look at what should be in $\Delta S_n / S_n$, which is expected to be similar to $\log(S_{n+1} / S_n)$. For the sake of illustration, let us assume

$$\frac{\Delta S_n}{S_n} = \alpha \Delta t + \beta \sqrt{\Delta t} X_{n+1}$$

(19)

which is similar to Eq.(??). This assumption can be justified but we would rather skip at this point. Here we are interested in obtaining $\alpha$ and $\beta$. Just for curiosity, we look at the square term and try to collect the lowest order terms, and find

$$\left( \frac{\Delta S_n}{S_n} \right)^2 = \beta^2 \Delta t + o(\Delta t),$$

(20)

Here the leading term is $O(\Delta t)$ because of the $\sqrt{\Delta t}$ term in Eq.(??) and the fact that $X_{n+1}^2 = 1$. Comparing coefficients with Eq.(??), we conclude

$$\alpha = r, \quad \beta = \sigma$$

(21)

which gives

$$\frac{\Delta S_n}{S_n} = r \Delta t + \sigma \sqrt{\Delta t} X_{n+1}.$$ 

(22)

This is actually the starting point of the Black-Scholes-Merton model. Notice that there is a difference $\sigma^2 \Delta t / 2$ between $\Delta S_n / S_n$ and $\log(S_{n+1} / S_n) = \log(1 + \Delta S_n / S_n)$, which can be explained by the result of our exercise (to calculate $\mathbb{E}[e^{X}]$ for a normally distributed $X$).

4 First Passage Time Distribution

At this point, we have established our binomial model expressed by a random walk:

$$S_n = S_0 e^{\mu n \Delta t + \sigma \sqrt{\Delta t} M_n / \sqrt{n}}$$

(23)

where $\mu = (r - \frac{1}{2} \sigma^2)$. As we pointed out, the scaling in $M_n / \sqrt{n}$ is introduced so that $M_n / \sqrt{n}$ is expected to converge to a standard normal random variable, in distribution, as $n \to \infty$, thanks to the central limit theorem.

As mentioned early, one of the advantages of the symmetric random walk is the dealing of a martingale, which has very nice properties, especially when it comes to crossings. The event that $S_n \geq S^*$ for a given $S^* > 0$ is the same as the event $\mu n \Delta t + \sigma \sqrt{\Delta t} M_n \geq \log(S^* / S_0)$, and the first time this happens is denoted by

$$\tau_{S^*} = \min_{n \geq 0} \{ n : \sigma \sqrt{\Delta t} M_n \geq \log(S^* / S_0) - \mu n \Delta t \}.$$ 

(24)
The goal is to find the distribution of \( \tau_{S^*} \): \( \mathbb{P}\{ \tau_{S^*} \leq t \} \). For the barrier options, we need joint distribution functions such as

\[
F_\mu(b, c) = \lim_{N \to \infty} \mathbb{P}\{\sigma\sqrt{T} \frac{M_N}{\sqrt{N}} + \mu T < b, \tau_c > T\}, \tag{25}
\]

where the limit is taken under the condition that \( N\Delta t = T \).

It is helpful to look at the following simplest problem: find the distribution of \( \tau_m = \min\{n : M_n \geq m\} \). \tag{26}

Here we use a slightly different notation compared to Eq.(5.2.1) on page 120 of the Shreve textbook. The reason for this is that we will need to extend the barrier from integers \( m \) to real numbers.

One way to bypass the explicit dealing with \( \tau_m \) is to introduce the maximum process

\[
\tilde{M}_n = \max_{0 \leq m \leq n} M_m, \tag{27}
\]

the maximum of \( M_m \) after \( n \) steps, or the farthest to the right the random walk has gone after \( n \) steps. It is obvious that \( \tau_m \leq N \) is equivalent to \( \tilde{M}_N \geq m \).

It is proved in the Shreve textbook that (i): \( \mathbb{P}\{\tau_m < \infty\} = 1 \), and (ii): \( \mathbb{E}\[\tau_m\] = \infty \).

To obtain the pricing formulas for the barrier options, we will need more detailed results than that. It turns out that the reflection principle is useful to compute the distribution of \( \tau_m \). When \( \Delta t \to 0 \), \( n \to \infty \) such that \( n\Delta t = t \) is fixed, the distribution can be even expressed in terms of the cumulative normal distribution.

The main claim of the reflection principle is the following: assuming \( m > l > 0 \),

\[
\mathbb{P}\{M_N < l, \tau_m \leq N\} = \mathbb{P}\{M_N < l, \tilde{M}_N \geq m\} = \mathbb{P}\{M_N > 2m - l\}. \tag{28}
\]

The last probability is much easier to compute than the probability on the left, which involves a joint distribution. In particular, when \( N \to \infty \) and \( \Delta t \to 0 \) such that \( N\Delta t = T \) is fixed, \( \sigma\sqrt{\Delta t} M_N \) converges to a normal random variable with mean zero and variance \( \sigma^2 T \), in distribution. Then

\[
\mathbb{P}\{\sigma\sqrt{\Delta t} M_N < b, \sigma\sqrt{\Delta t} \tilde{M}_N \geq c\} = \mathbb{P}\{\sigma\sqrt{\Delta t} M_N > 2c - b\} \approx N\left(-\frac{2c-b}{\sigma\sqrt{T}}\right), \tag{29}
\]

here \( N(x) \) is the cumulative normal distribution function. Consequently we have the joint distribution for \( \sigma\sqrt{\Delta t} M_N \) and \( \sigma\sqrt{\Delta t} \tilde{M}_N \)

\[
F_\theta(b, c) = \lim_{N \to \infty, \Delta t \to 0} \mathbb{P}\{\sigma\sqrt{\Delta t} M_N < b, \sigma\sqrt{\Delta t} \tilde{M}_N < c\}
= \lim_{N \to \infty, \Delta t \to 0} \left[\mathbb{P}\{\sigma\sqrt{\Delta t} M_N < b\} - \mathbb{P}\{\sigma\sqrt{\Delta t} M_N < b, \sigma\sqrt{\Delta t} \tilde{M}_N \geq c\}\right]
= N\left(\frac{b}{\sigma\sqrt{T}}\right) - N\left(\frac{b-2c}{\sigma\sqrt{T}}\right), \quad \text{for } c \geq b.
\]
For $c < b$, we should have $F_0(b,c) = F_0(c,c)$. The joint density function is obtained through differentiation,

$$f_0(b,c) = \begin{cases} \frac{2(2c-b)}{\sigma^3\sqrt{T}} N' \left( \frac{b-2c}{\sigma\sqrt{T}} \right), & c \geq b, \\ 0, & c < b. \end{cases} \quad (30)$$

This would be the result we need for $\tau_{S^*}$ only if $\mu = 0$. For $\mu \neq 0$, there is an $n$ factor in the drift part, which leads to a non-horizontal barrier. This turns out to be a nontrivial problem and our approach is to turn to the change of measure trick for help again. The idea behind this is the celebrated Girsanov’s theorem which we will only motivate, but not formerly state here.

Here is how it works: a symmetric random walk plus a known drift is what we have under the current measure, but it can be viewed as an asymmetric random walk without a drift term under another measure, in the sense that the mean and variance are reproduced. That asymmetric random walk can be modified by changing the probabilities to become symmetric random walk under another measure. Of course there is a price to be paid, that is we need to introduce a Radon-Nikodym derivative again.

To be specific, we write

$$\sigma \sqrt{\Delta t} M_n + \mu n \Delta t = \sigma \sqrt{\Delta t} B_n. \quad (31)$$

The left-hand-side is the symmetric random walk combined with the known drift, which is viewed under the original measure $P$. We would like to think of it as just a random walk, drift free, under a new measure $Q'$. In order for the distribution of the increment to be preserved, we require

$$E_n^P \left[ \sigma \sqrt{\Delta t} X_{n+1} + \mu \Delta t \right] = E_n^{Q'} \left[ \sigma \sqrt{\Delta t} (B_{n+1} - B_n) \right]. \quad (32)$$

This gives us some guidance as how to adjust the probabilities under $Q'$. We notice

$$B_{n+1} - B_n = X_{n+1} + \frac{\mu \sqrt{\Delta t}}{\sigma} = Y_{n+1}. \quad (33)$$

We would like to have

$$Y_{n+1} = \begin{cases} 1 & \text{if } \omega_{n+1} = H \text{ under } Q' \\ -1 & \text{if } \omega_{n+1} = T \end{cases} \quad (34)$$

For Eq. (??) to be satisfied, we need the probabilities

$$P^{Q'} \{ \omega_{n+1} = H \} = \frac{1}{2} \left( 1 + \frac{\mu \sqrt{\Delta t}}{\sigma} \right), \quad P^{Q'} \{ \omega_{n+1} = T \} = \frac{1}{2} \left( 1 - \frac{\mu \sqrt{\Delta t}}{\sigma} \right). \quad (35)$$
The final change of measure take us to a world where \( Y_n \) takes head and tail with equal probability \( 1/2 \) and the Radon-Nikodým derivative (from \( Q' \) to \( Q \), or equivalently from \( P \) to \( Q \)) in this case is given by

\[
\frac{\tilde{Z}_{n+1}}{\tilde{Z}_n} = \frac{1}{1 + \frac{\mu \Delta t}{\sigma} X_{n+1}} \approx 1 - \frac{\mu}{\sigma} \sqrt{\Delta t} X_{n+1} \approx \exp \left( -\frac{\mu}{\sigma} \sqrt{\Delta t} X_{n+1} - \frac{1}{2} \left( \frac{\mu}{\sigma} \right)^2 \Delta t \right),
\]

for \( \Delta t \) small enough. Again, the second term in the exponential is introduced so that the error of the last approximation is \( o(\Delta t) \), rather than \( O(\Delta t) \), when we compare the expectations. The advantage of the use of the exponential is that when we accumulate the changes, we have

\[
\tilde{Z}_n = \prod_{j=1}^{n} \frac{\tilde{Z}_j}{\tilde{Z}_{j-1}} = \exp \left( -\frac{\mu}{\sigma} \sqrt{\Delta t} M_n - \frac{1}{2} \left( \frac{\mu}{\sigma} \right)^2 n \Delta t \right)
\]

and when viewed through the variable \( B_n \),

\[
\tilde{Z}_n^{-1} = \exp \left( \frac{\mu}{\sigma} \sqrt{\Delta t} B_n - \frac{1}{2} \left( \frac{\mu}{\sigma} \right)^2 n \Delta t \right).
\]

This is the Radon-Nikodým derivative we need.

Now we can write the joint distribution for the case \( \mu \neq 0 \), under the original measure, as

\[
F_{\mu}(b, c) = \lim_{N \to \infty, \Delta t \to 0} \mathbb{P}\{\sigma \sqrt{\Delta t} M_N + \mu T < b, \sigma \sqrt{\Delta t} \tilde{M}_N + \mu T < c\}
\]

\[
= \lim_{N \to \infty, \Delta t \to 0} \mathbb{E}_P \left[ I_{\{\sigma \sqrt{\Delta t} B_N < b, \sigma \sqrt{\Delta t} \tilde{B}_N < c\}} Z_N^{-1} \right]
\]

\[
= \int_0^c \int_{-\infty}^b \exp \left( \frac{\mu}{\sigma^2} x - \frac{1}{2} \left( \frac{\mu}{\sigma} \right)^2 T \right) f_0(x, y) dy dx
\]

\[
= N \left( \frac{b - \mu T}{\sigma \sqrt{T}} \right) - e^{2\mu c} N \left( \frac{b - 2c - \mu T}{\sigma \sqrt{T}} \right).
\]

The integration in \( y \) is over \((0, c)\) because \( \tilde{B}_n \geq 0 \).

Now we can compute the joint density that is required in many of the barrier option calculations:

\[
f_{\mu}(b, c) = \begin{cases} 
\frac{2(2c-b)}{\sigma \sqrt{T} \sigma^2} N' \left( \frac{2c-b}{\sigma \sqrt{T}} \right) \exp \left( \frac{\mu b - \frac{1}{2} \mu^2 T}{\sigma^2} \right), & c \geq b, \\
0, & c < b.
\end{cases}
\]

\[39\]
5 Valuation of the Barrier Options

Here we give one example of the barrier options, utilizing the density function given by (??). First we notice that the distribution of $\tau_{S^*}$ can be expressed as

$$P\{\tau_{S^*} > T\} = P\left\{\sigma \sqrt{\Delta t} B_N < \log \left( \frac{S^*}{S_0} \right), \sigma \sqrt{\Delta t} \tilde{B}_N < \log \left( \frac{S^*}{S_0} \right) \right\} = F_\mu \left( \log \left( \frac{S^*}{S_0} \right), \log \left( \frac{S^*}{S_0} \right) \right).$$

To work on the lookback call option, it is necessary to rewrite

$$\min_{0 \leq t \leq T} S_t = S_0 e^{-\max(-\sigma \sqrt{\Delta t} B_n)}.$$  \hspace{1cm} (40)

The difference between $B_n$ and $-B_n$ is just in the drift ($-\mu$ vs. $\mu$). Let us introduce $G_n = -B_n$, $\tilde{G}_n = \max G_n = \max(-B_n) = -\min B_n$, the option price can be computed as

$$C_{\text{lookback}} = e^{-rT} \mathbb{E}_P \left[ S_0 e^{-\sigma \sqrt{\Delta t} G_N} - S_0 e^{-\max_{0 \leq n \leq N}\left(\sigma \sqrt{\Delta t} G_n\right)} \right]$$

$$= S_0 - e^{-rT} S_0 \mathbb{E}_P \left[ e^{-\max_{0 \leq n \leq N}\left(\sigma \sqrt{\Delta t} G_n\right)} \right]$$

$$= S_0 - e^{-rT} S_0 \int_0^\infty \int_{-\infty}^y e^{-y} f_{-\mu}(x, y) dx dy.$$

The integration in $x$ is from $-\infty$ to $y$ since $G_N \leq \max_{0 \leq n \leq N} G_n = \tilde{G}_N$. 

\[8\]