Verifying the Martingale Property for Subsequent Periods

In Section 5.2.2.2, we verified that the stock price relative to the bond (the bond as the numeraire) satisfies the martingale property from time 0 to 1. Here, we will do the same for the stock price from time 1 to 2 and then in subsequent periods. We will also demonstrate that the required measure that makes the discounted stock price into a martingale is the risk-neutral probability $\mathbb{Q}$. This appendix is essentially a continuation of Section 5.2.2.2.

Suppose, we are given the stock price $S_{0,1;u}$ at time 1 after an upjump at time 1. Denote $S_{0,2;u,u}$ and $S_{0,2;u,d}$ as the stock prices at time 2 given the outcomes $\omega_{u,u}$ (consecutive upjumps at time 1 and 2) and $\omega_{u,d}$, respectively. Denote $q_{1,2;u,u}$ and $q_{1,2;u,d}$ as the risk-neutral probabilities that at time 1, the stock will have an upjump or downjump at time 2 given that it had an upjump at time 1. Let $B_{1,2}$ (discount function) be the time 1 value of a bond that pays 1 at time 2. Analogous to equation (5.15), the forward prices from time 1 to 2 given the stock price $S_{0,1;u}$ at time 1, must satisfy the system of equations:

(1)
$$\begin{bmatrix} S_{0,2;u,u} & S_{0,2;u,d} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \psi_{1,2;u,u} \\ \psi_{1,2;u,d} \end{bmatrix} = \begin{bmatrix} S_{0,1;u} \\ B_{1,2} \end{bmatrix}.$$

If we divide the matrix equation by $B_{1,2}$, and define the risk-neutral probabilities $q_{1,2;u,u} = \psi_{1,2;u,u}/B_{1,2}$ and $q_{1,2;u,d} = \psi_{1,2;u,d}/B_{1,2}$ as required by equation (5.12) we get:

(2)
$$\begin{bmatrix} S_{0,2;u,u} & S_{0,2;u,d} \\ 1 & 1 \end{bmatrix} \begin{bmatrix} q_{1,2;u,u} \\ q_{1,2;u,d} \end{bmatrix} = \begin{bmatrix} S_{0,1;u}/B_{1,2} \\ 1 \end{bmatrix}.$$

The first row of matrix equation (2) is the equation:

(3)
$$S_{0,2;u,u} q_{1,2;u,u} + S_{0,2;u,d} q_{1,2;u,d} = \frac{S_{0,1;u}}{B_{1,2}}.$$

Equation (3) can be expressed in the form:

(4)
$$E_{\mathbb{Q}}\left[S_2 \middle| \frac{S_{0,1;u}}{B_{1,2}} \right] = \frac{S_{0,1;u}}{B_{1,2}}.$$

We can repeat the calculation given the stock price $S_{0,1;d}$. Analogous to equations (3) and (4), we would obtain:

(5)
$$S_{0,2;d,u} q_{1,2;d,u} + S_{0,2;d,d} q_{1,2;d,d} = \frac{S_{0,1;d}}{B_{1,2}}$$

and

(6)
$$E_{\mathbb{Q}}\left[S_2 \middle| \frac{S_{0,1;d}}{B_{1,2}} \right] = \frac{S_{0,1;d}}{B_{1,2}}.$$

Equations (4) and (6) together verify the martingale property going from time 1 to 2 that the expected value of the stock at time 2 given its value at time 1 (in bond numeraire) is equal to the value of the stock at time 1 (in bond numeraire).

Next, we study the stock model going from time 0 to 2. Let, $q_{0,2;u,d}$, $q_{0,2;d,u}$, and $q_{0,2;d,d}$ denote the risk-neutral probabilities at time 0 that the stock price will experience the outcomes $\omega_{u,u}$, $\omega_{u,d}$, $\omega_{d,u}$, and $\omega_{d,d}$, respectively, in going from time 0 to time 2. Let $B_{0,2}$ be the time 0 value of a bond that pays 1 at time 2. Note that this bond value must satisfy the equation $B_{0,2}=B_{0,1}B_{1,2}$. Recall that $\psi_{0,2;u,u}$ is the price of the pure security at time 0 that pays 1 at time 2 after the upjump in the price of the stock from time 0 to time 1 and also from time 1 to time 2. The prices $\psi_{0,2;u,d}$, $\psi_{0,2;d,u}$, and $\psi_{0,2;d,d}$ are defined analogously. Similar to previous analysis, we have:

(7)
$$\begin{bmatrix} S_{0,2;u,u} & S_{0,2;u,d} & S_{0,2;d,u} & S_{0,2;d,d} \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} \psi_{0,2;u,u} \\ \psi_{0,2;u,d} \\ \psi_{0,2;d,u} \\ \psi_{0,2;d,d} \end{bmatrix} = \begin{bmatrix} S_0 \\ B_{0,2} \end{bmatrix}.$$

If we divide both sides of equation (7) by $B_{0,2}$ and use equation (5.12) to define the risk-neutral probabilities, we obtain:

(8)
$$\begin{bmatrix} S_{0,2;u,u} & S_{0,2;u,d} & S_{0,2;d,u} & S_{0,2;d,d} \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} q_{0,2;u,u} \\ q_{0,2;u,d} \\ q_{0,2;d,u} \\ q_{0,2;d,d} \end{bmatrix} = \begin{bmatrix} S_0/B_{0,2} \\ 1 \end{bmatrix}.$$

Notice that the first row of the matrix equation (8) is the equation:

(9)
$$S_{0,2;u,u}q_{0,2;u,u}+S_{0,2;u,d}q_{0,2;u,d} + S_{0,2;d,u}q_{0,2;d,u}+S_{0,2;d,d}q_{0,2;d,d} = \frac{S_0}{B_{0,2}}.$$

The left side of equation (9) is the expected value with respect to the probability measure $\mathbb{Q}$ of the stock price at time 2 given its value at time 0 in bond numeraire, Thus equation (9) states that this expected value equals the value of the stock at time 0 in bond numeraire:

(10)
$$E_{\mathbb{Q}}[S_2|S_0/B_{0,2}] = \frac{S_0}{B_{0,2}}.$$

This is the required martingale property for the stock going from time 0 to 2. Now, we need to check that this result is consistent with the results going from time 0 to 1 followed by going from time 1 to 2. Since $q_{0,2;u,u}$ is the probability that the stock had upjumps at times 1 and 2, $q_{1,2;u,u}$ is the conditional probability that the stock had an upjump at time 2 given that it had an upjump at time 1, and $q_{0,1;u}$ is the probability that the stock had an upjump at time 1, then by the definition of conditional probability, we require $q_{1,2;u,u}= q_{0,2;u,u}/ q_{0,1;u}$, or equivalently: $q_{0,2;u,u}= q_{0,1;u} \, q_{1,2;u,u}$. The relationships for the other risk-neutral probabilities are computed in a similar way giving us:

(11)
$$q_{0,2;u,u} = q_{0,1;u}q_{1,2;u,u}, \; q_{0,2;u,d} = q_{0,1;u}q_{1,2;u,d}, \; q_{0,2;d,u} = q_{0,1;d}q_{1,2;d,u},$$

$$q_{0,2;d,d} = q_{0,1;d}q_{1,2;d,d}$$

Note that equations (11) define the risk-neutral probabilities from time 0 to time 2, and we need to check that they are consistent with equation (9). Since the left side of equation (9) is $E_{\mathbb{Q}}[S_2|S_0/B_{0,2}]$ and substituting the results of equations (11) into the left side of equation (9) gives:

$$E_{\mathbb{Q}}[S_2|S_0/B_{0,2}] = S_{0,2;u,u}q_{0,1;u}q_{1,2;u,u} + S_{0,2;u,d}q_{0,1;u}q_{1,2;u,d}$$

$$+S_{0,2;d,u}q_{0,1;d}q_{1,2;d,u} + S_{0,2;d,d}q_{0,1;d}q_{1,2;d,d}$$

$$= (S_{0,2;u,u}q_{1,2;u,u} + S_{0,2;u,d}q_{1,2;u,d})q_{0,1;u} + (S_{0,2;d,u}q_{1,2;d,u} + S_{0,2;d,d}q_{1,2;d,d})q_{0,1;d}$$

Referring to equations (3) and (5), we see that:

$$E_{\mathbb{Q}}[S_2|S_0/B_{0,2}] = \frac{S_{0,1;u}q_{0,1;u} + S_{0,1;d}q_{0,1;d}}{B_{1,2}}.$$

Referring to equation (5.17) and using the fact that $B_{0,2}=B_{0,1}B_{1,2}$, we get:

(12)
$$E_{\mathbb{Q}}[S_2|S_0/B_{0,2}] = \frac{S_0}{B_{0,2}}.$$

Equation (10) is identical to equation (12), so the results are consistent. Equation (12) states that the expected value with respect to the probability measure $\mathbb{Q}$ of the stock price at time 2 given its price at time 0 (in bond numeraire) is equal to the stock price at time 0 (in bond numeraire). Thus, we have verified all of the martingale properties up to time 2. The proof to show the martingale properties for all subsequent times essentially duplicates the work we have done using an induction argument. We omit the details. The induction argument would give the result:

$$E_{\mathbb{Q}}\left[S_n \middle| \frac{S_0}{B_{0,n}}, \frac{S_1}{B_{1,n}}, \dots, \frac{S_k}{B_{k,n}}\right] = \frac{S_k}{B_{k,n}}$$

for all $k=0,1,2,\dots,n-1$.

Since $E_{\mathbb{Q}}[S_n|S_0,S_1,\dots,S_k] = S_k \; \forall \; k < n$ if the stock prices are expressed in the units of the bond as the numeraire, it is clear that S_t is a martingale, despite its having been a submartingale when priced relative to currency.

Brownian Motion as a Limit of a Discrete Random Walk

Think of the discrete random walk and Brownian motion taking place on the z-axis, where we start at the origin ($z=0$) at time 0. We will then study these two processes as a function of time. We can also graph each of these two processes if we wish with position z as the vertical axis versus time t as the horizontal axis, analogous to the graph of a stock price (S) versus time (t). The graph of the discrete random walk versus time for a specific history will consist of a large number of discrete points in the t-z plane, while the graph of Brownian motion versus time for a specific history will consist a continuous zig-zag graph in the t-z plane. To create Brownian motion from the discrete random walk with $p = 1/2$, first divide each unit of time on the continuous time axis into n equal subintervals, so that the width of each subinterval is $\Delta t = 1/n$. This results in the discrete times 0, $1/n$, $2/n$,...,k/n,... on the continuous time axis. Define the following discrete random walk as a function of time k (the kth discrete time on the continuous time axis at time k/n) by:

$$X_k^{(n)} = \sum_{i=1}^{k} \frac{Z_i}{\sqrt{n}}.$$

Assume that each Z_i is a random variable that equals 1 with probability ½ and equals -1 with probability ½ at time i, and that the Z_i's are pairwise independent of one another. This discrete random walk, starting at 0, takes independent small steps of length $\pm\frac{1}{\sqrt{n}}$ at each moment of time from $(i-1)/n$ to i/n as i goes from 1 to k until it arrives at the position $X_k^{(n)}$ at time k/n. Since the random walk $X_k^{(n)}$ takes ever-smaller steps approaching length 0 as $n \to \infty$, it is approaching a continuous path as $n \to \infty$ for each specific history. This justification for the existence and continuity of the path isn't rigorous. A rigorous argument is beyond the level of this text.

We next claim that as $n \to \infty$, the random walk will converge to a process that has a normal distribution (which is one the requirements to be Brownian motion).To prove this, pick a particular time t on the continuous non-negative real line. Pick the integer k, $k \leq nt$, so that it best approximates $k/n \approx t$. Note that $|k/n$-$t|<1/n$ so that the approximation improves as n approaches ∞. Also note that $k \approx nt$, so that $k \to \infty$ as $n \to \infty$. Define the sample mean random variable

$$\bar{Z} = \frac{\sum_{i=1}^{k} Z_i}{k}.$$

Clearly, $X_k^{(n)} = \frac{k}{\sqrt{n}}\bar{Z}$. Since the random variables Z_i are independent and identically distributed random variables with mean $\mu = 0$ and variance $\sigma^2 = 1$, by the Central Limit Theorem from section 2.7, we conclude that $\sqrt{k}\bar{Z} = \frac{\bar{Z}}{1/\sqrt{k}}$ approaches the standard normal random distribution $Z\sim N(0,1)$ as $k \to \infty$. This implies that

$$X_k^{(n)} = \sqrt{k}\bar{Z}\sqrt{\frac{k}{n}} \to Z\sqrt{t}$$

as $k \to \infty$. Since $Z\sqrt{t} \sim N(0, t)$, this shows that $X_k^{(n)}$ approaches a normal distribution, say X_t, with mean 0 and variance t. We have completed the main step to show that our random walk converges to a process that has a normal distribution with mean 0 and variance t.

The rest of the properties are easier to justify. If we choose the integer j so that it j/n best approximates a value s, with $s < t$, on the time axis, then by a proof similar to that we just completed, we have $X_k^{(n)} - X_j^{(n)}$ approaching a normal distribution $X_t - X_s$ with mean 0 and variance $t - s$. Consider the values $s < t < u < v$ on the time axis. Choose the integers $j < k < l < m$ so that they best approximate $j/n \approx s$, $k/n \approx t$, $l/n \approx u$, and $m/n \approx v$. Since

$$X_k^{(n)} - X_j^{(n)} = \sum_{i=j+1}^{k} \frac{Z_i}{\sqrt{n}}, \quad X_m^{(n)} - X_l^{(n)} = \sum_{i=l+1}^{m} \frac{Z_i}{\sqrt{n}},$$

and $j +1 \leq k \leq l + 1 \leq m$, then the random variables $X_k^{(n)} - X_j^{(n)}$ and $X_m^{(n)} - X_l^{(n)}$ are independent. This is easily seen by observing that the Z_i's in the first sum are independent of the Z_i's in the second sum in this case. But $X_k^{(n)} - X_j^{(n)}$ approaches $X_t - X_s$ and $X_l^{(n)} - X_m^{(n)}$ approaches $X_u - X_v$ as $n \to \infty$. Thus, $X_t - X_s$ and $X_u - X_v$ should be independent of one another. Each of the requirements in order that $X_k^{(n)}$ will converge to the standard Brownian motion process as $n \to \infty$ have been addressed, and we conclude that the convergent process X_t is in fact standard Brownian motion.

Geometric Brownian Motion as a Limit of a Discrete Stochastic Process

In the same way that we derived Brownian motion as a limit of a discrete stochastic process, we divide each unit of time into n equal subintervals, so that the width of each subinterval is $\Delta t = 1/n$. This results in the discrete times $0, 1/n, 2/n, ..., k/n, ...$ on the continuous time axis. Define the price of a stock $S_k^{(n)}$ as a discrete stochastic process as a function of time k (the kth discrete time on the continuous time axis at time k/n) so that:

$$\frac{S_k^{(n)} - S_{k-1}^{(n)}}{S_{k-1}^{(n)}} = \frac{\mu}{n} + \frac{\sigma Z_k}{\sqrt{n}}.$$

Assume that each Z_k is a random variable that equals 1 with probability $\frac{1}{2}$ and equals -1 with probability $\frac{1}{2}$ at time i, and that the Z_i's are pairwise independent of one another. Observe that μ is the return per unit time per unit value of the stock and σ is the degree of volatility of the stock. Solving for $S_k^{(n)}$ gives:

$$S_k^{(n)} = \left(1 + \frac{\mu}{n} + \frac{\sigma Z_k}{\sqrt{n}}\right) S_{k-1}^{(n)}.$$

Substituting k with k -1 in this equation and replacing the result for $S_{k-1}^{(n)}$ in terms of $S_{k-2}^{(n)}$ on the right side of the equation above, and continuing this process all the way back to time 0 results in:

$$S_k^{(n)} = S_0 \left(1 + \frac{\mu}{n} + \frac{\sigma Z_1}{\sqrt{n}}\right)\left(1 + \frac{\mu}{n} + \frac{\sigma Z_2}{\sqrt{n}}\right)\cdots\left(1 + \frac{\mu}{n} + \frac{\sigma Z_k}{\sqrt{n}}\right).$$

Taking logarithms, we have:

$$lnS_k^{(n)} = lnS_0 + \sum_{i=1}^{k} ln\left(1 + \frac{\mu}{n} + \frac{\sigma Z_i}{\sqrt{n}}\right).$$

From calculus, we know that the Taylor series for $ln(1+x)$ centered at zero is

$$\ln(1 + x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots$$

For a particular history ω of the stock, each $Z_i(\omega)$ will take on the definite value of either 1 or -1. Applying the Taylor series above to

$$x = \frac{\mu}{n} + \frac{\sigma Z_i(\omega)}{\sqrt{n}}$$

for each i results in:

$$lnS_k^{(n)}(\omega) = lnS_0 + \sum_{i=1}^{k}\left[\frac{\mu}{n} + \frac{\sigma Z_i(\omega)}{\sqrt{n}} - \frac{1}{2}\left(\frac{\mu}{n} + \frac{\sigma Z_i(\omega)}{\sqrt{n}}\right)^2 + \frac{1}{3}\left(\frac{\mu}{n} + \frac{\sigma Z_i(\omega)}{\sqrt{n}}\right)^3 - \cdots\right]$$

$$= lnS_0 + \mu\frac{k}{n} + \sigma\sum_{i=1}^{k}\frac{Z_i(\omega)}{\sqrt{n}} - \frac{\sigma^2}{2n}\sum_{i=1}^{k}(Z_i(\omega))^2 + error\ terms.$$

Since $Z_i(\omega)$ is always 1 or -1, then $(Z_i(\omega))^2 = 1$. So, we have:

$$lnS_k^{(n)}(\omega) = lnS_0 + (\mu - \frac{\sigma^2}{2})\frac{k}{n} + \sigma\sum_{i=1}^{k}\frac{Z_i(\omega)}{\sqrt{n}} + error\ terms.$$

Similar to the derivation of Brownian motion from the discrete model, pick a particular time t on the continuous time axis. Choose the integer k so that $k/n \leq t$ and k/n best approximates t. As before, $|k/n - t| \leq 1/n$, and so $k/n \to t$ as $n \to \infty$. It is also not hard to show that the error terms are bounded by $C(t, \mu, \sigma)/\sqrt{n}$ where $C(t,\mu,\sigma)$ is a value independent of n. Thus, the error

terms approach 0 as $n \to \infty$. We showed in the derivation of Brownian motion from the discrete model that $\sum_{i=1}^{k} \frac{Z_i(\omega)}{\sqrt{n}}$ converges to $Z_t(\omega)$ where Z_t is standard Brownian motion. These observations show $lnS_k^{(n)}(\omega)$ converges to

$$lnS_0 + \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma Z_t(\omega)$$

as $n \to \infty$. Exponentiating, we have $S_k^{(n)}(\omega)$ converges to

$$S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma Z_t(\omega)}.$$

Let's label this resulting continuous stochastic process as $S_t(\omega)$. Since this result holds for all possible histories ω in the sample space, then

$$S_t = S_0 e^{\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma Z_t}.$$

Verification of the Tower Property

First, we state the Tower Property: *Suppose X_t is a stochastic process with an adapted filtration $\{\mathcal{F}_t\}$. If $s \leq t$, then for any $T \geq t$, we have*

$$E[E[X_T|\mathcal{F}_t]| \mathcal{F}_s] = E[X_T|\mathcal{F}_s].$$

To verify this property for the case that the filtrations are finite, consider the following more general theorem. Suppose $\mathcal{G}$ and $\mathcal{F}$ are σ–algebras with $\mathcal{G} \subseteq \mathcal{F}$, and X is a random variable with respect to the σ – algebra with $\mathcal{F}$. Then, the following expectation result holds:

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

Observe that the tower property follows from this result by choosing $\mathcal{G} = \mathcal{F}_s$ and $\mathcal{F} = \mathcal{F}_t$. We are going to prove this more general statement for the case that the σ-algebras are finite.

To interpret the expectation equation immediately above, notice that the σ – algebra $\mathcal{F}$ is a "finer" σ – algebra than $\mathcal{G}$, since it contains all of the measurable sets that are in $\mathcal{G}$. Although a proof is provided below, this result is understandable intuitively. Roughly speaking, the equation states that if one averages a random variable conditioned on the finer subdivision $\mathcal{F}$ and then averages these averages conditioned on the coarser subdivision $\mathcal{G}$, one gets the same average as conditioning on the coarser subdivision $\mathcal{G}$ alone.

Verification of the theorem: Since we are assuming the σ–algebra $\mathcal{G}$ is finite, then $\mathcal{G}$ consists of a finite number of disjoint sets G_1, G_2,…, G_n and all of the possible unions of these

sets. Since $\mathcal{G} \subseteq \mathcal{F}$ and $\mathcal{F}$ is finite, then we can further decompose each set G_i into disjoint sets F_{ij} for $j = 1, \dots, n_i$ so that

$$G_i = \bigcup_{j=1}^{n_i} F_{ij}$$

and $\mathcal{F}$ consists of the sets F_{ij} and all of the possible unions of these sets. To find $E[X|\mathcal{G}]$ means to calculate $E[X|G_i]$ for each i. Suppose $p(x)$ denotes the probability that the random variable $X = x$. Then, by definition,

$$E[X|G_i] = \frac{\sum_{x \in G_i} x p(x)}{\sum_{x \in G_i} p(x)}.$$

To find $E[X|\mathcal{F}]$ means to calculate $E[X|F_{ij}]$ for each i and j. By definition,

$$E[X|F_{ij}] = \frac{\sum_{x \in F_{ij}} x p(x)}{\sum_{x \in F_{ij}} p(x)}.$$

To find $E[E[X|\mathcal{F}]|\,\mathcal{G}]$ means to calculate $E[E[X|\mathcal{F}]|\,G_i]$ for each i. $E[X|\mathcal{F}]$ is a random variable that takes on the values $E[X|F_{ij}]$ with probability $\sum_{x \in F_{ij}} p(x)$. Thus:

$$E[E[X|\mathcal{F}]|G_i] = \frac{\sum_{j=1}^{n_i} E[X|F_{ij}](\sum_{x \in F_{ij}} p(x))}{\sum_{x \in G_i} p(x)}$$

$$= \frac{\sum_{j=1}^{n_i} \left[\frac{\sum_{x \in F_{ij}} x p(x)}{\sum_{x \in F_{ij}} p(x)}\right](\sum_{x \in F_{ij}} p(x))}{\sum_{x \in G_i} p(x)} = \frac{\sum_{j=1}^{n_i} \sum_{x \in F_{ij}} x p(x)}{\sum_{x \in G_i} p(x)}$$

Since the sets F_{ij} are disjoint from one another and $\bigcup_{j=1}^{n_i} F_{ij} = G_i$, then

$$\sum_{j=1}^{n_i} \sum_{x \in F_{ij}} x p(x) = \sum_{x \in G_i} x p(x).$$

We conclude that

$$E[E[X|\mathcal{F}]|\,\mathcal{G}] = \frac{\sum_{x \in G_i} x p(x)}{\sum_{x \in G_i} p(x)} = E[X|G_i]$$

as we wished to prove.