3. Multiperiod Security Market

3.1 Model Specifications, Filtrations, and Stochastic Processes

Model of the market with the following submodels

(i) $T + 1$ trading dates, $t = 0, 1, \ldots, T$.

(ii) Finite sample space: 
$$\Omega = \{\omega_1, \ldots, \omega_k\}.$$ 

(iii) Probability measure: 
$$P: \Omega \rightarrow [0, 1], P(\omega) > 0, \forall \omega \in \Omega.$$ 

(iv) Filtration: 
$$F = \{F_t : t = 0, 1, \ldots, T\},$$ 

a submodel describing how information about the security prices are revealed to the investors.

(v) Bank account process: 
$$B = \{B_t : t = 0, 1, \ldots, T\},$$ 

where $B_t$ is a stochastic process with $B_0 = 1$ and $B_t(\omega) > 0, \forall t$ and $\forall \omega \in \Omega$.

At $0 < t < T$ information cumulate enabling investors ruling out certain states as impossible and concentrate to those among which the final state will be. Thus we can model information accumulation as partitioning $\Omega$ to finer and finer subsets. Let $P_t$ denote time $t$ partition* of $\Omega$.

**Example.** Coin tossing. Let $S_t$ be such that $S_t = S_{t-1} + \$1$, if head (H) and $S_t = S_{t-1} - \$1$, if tail (T) with $S_0$ the initial amount of money. Three tosses. Then the sample space $\Omega$ is

\[
\begin{align*}
\{\text{HHH}\} &= \omega_1, \\
\{\text{HHT}\} &= \omega_2, \\
\{\text{HTH}\} &= \omega_3, \\
\{\text{THH}\} &= \omega_4, \\
\{\text{HTT}\} &= \omega_5, \\
\{\text{THT}\} &= \omega_6, \\
\{\text{TTT}\} &= \omega_7.
\end{align*}
\]

$S_t$ is a stochastic process. Thus for example if $S_2 = 0$ then $S_2(\omega_2) = [0, 1, 2, 1]$ is the sample path corresponding to state $\omega_2$.

*{$F_1, \ldots, F_m$} is a partition of $\Omega$, if $\Omega = \bigcup_m F_i$ and $F_i \cap F_j = \emptyset$ for $i \neq j$. 

Usually $B_t$ is nondecreasing, so that time interval $(t-1, t)$ return is nonnegative, i.e.,

$$r_t = \frac{B_t - B_{t-1}}{B_{t-1}} \geq 0, t = 1, \ldots, T.$$ 

(vi) $N$ risky security processes 
$$S_t = \{S_t(\omega) : t = 0, 1, \ldots, T\},$$ 

where $S_t$ is a nonnegative stochastic process for all $t = 1, \ldots, N$.

$S_t(\omega) = S_t(\omega, \omega_t)$ is the return of the risky security (at state $\omega \in \Omega$).

Two new features:

(a) The information submodel $\mathcal{E}$.

(b) The stochastic process submodel $S_H$.

3.1.1 Information Structures

How to model information revealed to the investor?

At $t = 0$ every state $\omega \in \Omega$ is possible. Nothing ruled out.

At $t = T$ the investors learn the true state.

The information accumulation process as time goes is for example in the case of $\omega_5$ as follows (if we are only interested in the final state)

<table>
<thead>
<tr>
<th>$t$</th>
<th>Outcome</th>
<th>Subset of the possible outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$H$</td>
<td>${\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7}$</td>
</tr>
<tr>
<td>2</td>
<td>$T$</td>
<td>${\omega_5}$</td>
</tr>
<tr>
<td>3</td>
<td>$T$</td>
<td>${\omega_5}$</td>
</tr>
</tbody>
</table>

At each step we observe either $H$ or $T$, so the incremental partitioning becomes

<table>
<thead>
<tr>
<th>$t$</th>
<th>Outcome</th>
<th>Partition</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$H$ or $T$</td>
<td>$P_0 = {\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7}$</td>
</tr>
<tr>
<td>1</td>
<td>$H$ or $T$</td>
<td>$P_1 = {\omega_1, \omega_2, \omega_3} \cup {\omega_4, \omega_5, \omega_6, \omega_7}$</td>
</tr>
<tr>
<td>2</td>
<td>$H$ or $T$</td>
<td>$P_2 = {\omega_1, \omega_2} \cup {\omega_3} \cup {\omega_4, \omega_5} \cup {\omega_6, \omega_7}$</td>
</tr>
<tr>
<td>3</td>
<td>$H$ or $T$</td>
<td>$P_3 = {\omega_1} \cup {\omega_2} \cup {\omega_3} \cup {\omega_4} \cup {\omega_5} \cup {\omega_6} \cup {\omega_7}$</td>
</tr>
</tbody>
</table>

Note that e.g. $P_2 = \{\omega_3, \omega_6, \omega_7\}$ is not possible.

In the above example we observe that the next step partition $P_{t+1}$ is obtained from $P_t$ by partitioning its sets. Schematically this corresponds the path or network diagram below.
Particularly we have

\[ \mathcal{F}_0 = \{ \Omega, \emptyset \} \]

and

\[ \mathcal{F}_T = \{ A : A \subseteq \Omega \} = \mathcal{P}(\Omega) \]

the power set of \( \Omega \), i.e., the set of all possible subsets of \( \Omega \).

**Example.** Coin tossing example continued.

<table>
<thead>
<tr>
<th>( t )</th>
<th>( \mathcal{F}_0 )</th>
<th>( \mathcal{F}_1 )</th>
<th>( \mathcal{F}_2 )</th>
<th>( \mathcal{F}_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>( \emptyset )</td>
<td>( \Omega, \emptyset )</td>
<td>( {{1}, {2}, {3}, {\emptyset, \Omega}} )</td>
<td>( {{1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}, \emptyset, \Omega})</td>
</tr>
<tr>
<td>1</td>
<td>( {{2}, {3}} )</td>
<td>( {\emptyset, \Omega, {1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}} )</td>
<td>( {\emptyset, \Omega, {1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}, {1}, {2}, {3}, {\emptyset, \Omega}} )</td>
<td>( {\emptyset, \Omega, {1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}, {1,2}, {1,3}, {2,3}, {1,2,3}, \emptyset, \Omega})</td>
</tr>
<tr>
<td>2</td>
<td>( {{1}, {3}} )</td>
<td>( {\emptyset, \Omega, {1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}, {1}, {3}} )</td>
<td>( {\emptyset, \Omega, {1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}, {1}, {3}, {\emptyset, \Omega}} )</td>
<td>( {\emptyset, \Omega, {1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}, {1}, {3}, {\emptyset, \Omega}} )</td>
</tr>
<tr>
<td>3</td>
<td>( {{1}} )</td>
<td>( {\emptyset, \Omega, {1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}, {1}} )</td>
<td>( {\emptyset, \Omega, {1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}, {1}} )</td>
<td>( {\emptyset, \Omega, {1}, {2}, {3}, {1,2}, {1,3}, {2,3}, {1,2,3}, {1}} )</td>
</tr>
</tbody>
</table>

Thus \( \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_3 \). \( \mathcal{F} = \{ \mathcal{F}_t : t = 0, 1, 2, 3 \} \) is a filtration.

The partitions \( \mathcal{P}_t \) consist of the smallest nonempty sets of the corresponding algebra \( \mathcal{F}_t \), i.e., \( \mathcal{P}_t = \{ A \in \mathcal{F}_t : A \neq \emptyset, A \not\subseteq B = A \text{ or } \emptyset \subseteq B \in \mathcal{F}_t \} \) (when the sample space is finite).

**Note.** (i) \( \emptyset \in \mathcal{F} \) where \( \emptyset \) is the empty set, (ii) if \( F, G \in \mathcal{F} \) then \( F \cap G \in \mathcal{F} \).

Note that in probability theory (sub)sets are called events.

Algebra is a family of subsets which is closed with under set operations. That is if a set \( D \) is obtained from sets in \( \mathcal{F} \) by some set operations then \( D \in \mathcal{F} \).

**Note.** If \( \{A_1, \ldots, A_m\} \) is a partition of \( \Omega \) then there is a unique algebra \( \mathcal{F}_A \) corresponding to this partition. Conversely for each algebra \( \mathcal{F} \) corresponds a unique partition of \( \Omega \).

Consequently, because algebras are more convenient in probability theory, the information structure will be modeled in terms of sequences \( \mathcal{F}_t \) algebras rather than partitions.

For the purpose we denote the sequence as

\[ \mathcal{F} = \{ \mathcal{F}_t : t = 0, 1, \ldots, T \} \]

where \( \mathcal{F}_t \subseteq \mathcal{F}_{t+1} \), that is if \( F \in \mathcal{F}_t \) than \( F \in \mathcal{F}_{t+1} \).

### 3.1.2 Stochastic Process Models of Security Prices

A stochastic process \( S_n \) is a real valued function \( S_n(t, \omega) \) of both \( t \) and \( \omega \). That is

\[ S_n : \{0, 1, \ldots, T\} \times \Omega \to \mathbb{R} \]

For each fixed \( \omega \in \Omega \) the function \( t \to S(t, \omega) \) is called the sample path.

**Example.** In our previous example with \( \omega = \omega_3 = HTH \) the sample path is \( S(t, \omega_3) = \{1, 0, 1\} \).

For fixed \( t \), the function \( \omega \to S_n(t, \omega) \) is a random variable.

**Example.** For \( t = 2 \) the possible values for the random variable \( S(2, \omega) \) are 2, 0 or -2.

To make the information structure consistent with the random variable \( S_n \) of security prices we need link the values of \( S_n \) to our information structure. This is done via the measurability concept.

Generally we say that a random variable \( X : \Omega \to \mathbb{R} \) is measurable with respect to algebra \( \mathcal{F} \), if

\[ \{ \omega \in \Omega : X(\omega) = x \} \in \mathcal{F} \text{ for all } x \in \mathbb{R} \]
Filtrations with which the security prices are adapted may be interpreted as learning processes of the investors about security prices. Usually, however, it is possible to construct several filtrations such that the price processes are adapted, but some of which may be unacceptable (against the financial theory or practical intuition).

Example. (Continued) If we define
\[ P_1 = \{\{w_1, w_2, w_3\}, \{w_4, w_5\}\}, \]
and
\[ P_2 = \{\{w_1, w_2\}, \{w_3, w_4, w_5\}\}. \]

Let \( F_0 \) and \( F_1 \) be the corresponding algebras (construct them as an exercise) then \( F = \{ F_0, F_1, F_2, F_3 \} \), where \( F_0 = (\emptyset, \Omega) \) and \( F_3 = \{ A : A \subseteq \Omega \) is the power set\} is a filtration and \( S(t) \) is adapted to \( F \). Then of course \( S(2, \omega) = S(2, \omega) \) is measurable, but it is also \( F_1 \) measurable! What are the consequences? Suppose that at time \( t = 1 \) we observe some \( \omega \) in \( \{w_3, w_5\} \). Then we know \( S(1, \omega) = 1 \), but because \( S(2, \omega) = S(2, \omega) = 0 \) we know exactly what the price will be at the next period \( t = 2 \) that is the price is fully predictable on the basis of the information at \( t = 1 \). The same is true in all other cases whenever we know the subset where \( \omega \) is at time \( t = 1 \).
Nevertheless there always exists one filtration that corresponds to learning about the prices as time goes on, but learning nothing more. This corresponds at each time point to the coarsest possible (i.e., the generating partitions consistent with the stock price process.

An acceptable filtration is constructed from the partitions generated by the stochastic process. The result is a filtration that is the coarsest possible (i.e., the generating partitions that are fixed) consistent with the stock price process. Nevertheless there always exists one filtration that is the coarsest possible (i.e., the generating partitions consistent with the stock price process.

Example. (Continued) Suppose instead of the coin tosses the investor observes the price process

\[
\begin{array}{c|c|c|c|c}
\omega & S_0 & S_1 & S_2 & S_3 \\
\hline
\omega & 0 & 1 & 2 & 3 \\
\omega_1 & 0 & 1 & 2 & 3 \\
\omega_2 & 0 & 1 & 2 & 3 \\
\omega_3 & 0 & 1 & 2 & 3 \\
\omega_4 & 0 & 1 & 2 & 3 \\
\end{array}
\]

This gives exactly the same partitions and algebras as earlier. Here the implied filtration is acceptable.

3.1.3 Trading Strategies

A trading strategy

\[ H_n = (H_0, H_1, \ldots, H_N) \]

is a vector of stochastic processes

\[ H_n = \{ H_n(t); t = 1, 2, \ldots, T \}, n = 0, 1, \ldots, N. \]

It will be assumed that \( H_n(t) \) is predictable.

We say that a stochastic process \( H_n(t) \) is predictable with respect to filtration \( \mathcal{F} = \{ \mathcal{F}_t; t = 0, 1, \ldots, T \} \) if \( H_n(t+1) \) is \( \mathcal{F}_t \) measurable. Note that predictable processes are always adapted.

Note. (i) \( H_0(0) \) is unspecified, because we interpret that, for \( n \geq 1 \), \( H_n(t) \) equals the number of units (e.g., shares) invested from time \( t-1 \) to \( t \). \( H_0(t)H_{n-1} \) equals the amount of money invested in the bank account at time \( t-1 \).

(ii) \( H_n(t) \) can be positive (saving, long position) or negative (borrowing, short position).

Example 3.3 Suppose \( N = 1, K = 4, \) and \( T = 2 \) and the markets are going to evolve as follows

\[
\begin{array}{c|c|c|c|c}
\omega & S_0 & S_1 & S_2 & S_3 \\
\hline
\omega & 5 & 8 & 9 & \\
\omega_1 & 5 & 8 & 6 & \\
\omega_2 & 5 & 4 & 6 & \\
\omega_3 & 5 & 4 & 3 & \\
\end{array}
\]

At time \( t = 0 \) investors observe \( S_0 = 5 \) for all \( \omega \in \Omega \), so \( \mathcal{F}_0 = \{ \Omega \} \). At time \( t = 1 \) investors observe either \( S_1 = 8 \) or \( S_1 = 4 \) giving rise to \( \mathcal{F}_1 = \{ \{ \omega_1 \}, \{ \omega_2 \}, \{ \omega_3 \} \} \), so that \( \mathcal{F}_2 = \{ \{ \omega_1 \}, \{ \omega_2 \}, \{ \omega_3 \} \} \) and \( \mathcal{F}_2 = \{ \{ \omega \}, \{ \omega_1 \}, \{ \omega_2 \}, \{ \omega_3 \} \} \). At time \( t = 2 \) investors observe \( S_2 \) and thereby deduce \( \omega \), and hence the partition is \( \{ \{ \omega \}, \{ \omega_1 \}, \{ \omega_2 \}, \{ \omega_3 \} \} \).

\( \mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \) is a filtration, and \( S_1 \) is adapted to it. As an information tree the structure is

\[
\begin{align*}
\mathcal{F}_0 & = \{ \Omega \} \\
\mathcal{F}_1 & = \{ \{ \omega_1 \}, \{ \omega_2 \}, \{ \omega_3 \} \} \\
\mathcal{F}_2 & = \{ \{ \omega \}, \{ \omega_1 \}, \{ \omega_2 \}, \{ \omega_3 \} \}
\end{align*}
\]

Figure 3.3 Information structure and risky securities for example 3.3.

(iii) A rationale for modeling \( H_n \) as a stochastic is that at each time point \( t \) investors make decision based on the available information, \( \mathcal{F}_t \) and nothing more, i.e., "adapting" to the filtration \( \mathcal{F} \). This (technically) makes it also predictable, because positions are changed on the basis of information revealed at time \( t \), so that position at time \( t+1 \) when information is next time revealed, are known already at time \( t \).

Example 3.1 (Continued) Suppose the investor takes initial position \( H_0(1) = h_0, H_1(1) = h_1, \ldots, H_N(1) = h_N \), i.e., the positions carried forward from time \( t = 0 \) to time \( t = 1 \). Then obviously \( H_0(1, \omega) = h_0 \) for all \( \omega \in \Omega \in \mathcal{F}_0 \). Thus \( H_1(1) \in \mathcal{F}_0 \). At time \( t = 1 \) the trader can adjust the position based on the information which becomes available, that is, on the observation whether \( \omega \in \{ \omega_1, \omega_2, \omega_3, \omega_4 \} \). The investor can choose some value for \( H_0(2, \omega) \), and some other value for \( H_2(2, \omega) \) if \( \omega \in \{ \omega_3, \omega_4 \} \). Thus \( H_2(2, \omega) = H_2(2, \omega_3) = H_2(2, \omega_4) \) and \( H_2(2, \omega_1) = H_2(2, \omega_2) = H_2(2, \omega_5) \), so that \( H_2(2) \in \mathcal{F}_2 \). In a similar fashion at \( t = 2 \) we deduce that the investor can change the positions such that \( H_3(3, \omega_1) = H_3(3, \omega_2), H_3(3, \omega_3) = H_3(3, \omega_4), \ldots, H_3(3, \omega_7) = H_3(3, \omega_8) \). That is \( H_n(t) \in \mathcal{F}_t \).
3.1.4 Value Processes and Gains Processes

The value process
\[ V = \{ V_t; t = 0, 1, \ldots, T \} \]
is a stochastic process defined by
\[ V_t = \begin{cases} H_0(1)B_0 + \sum_{n=1}^{N} H_n(1)S_n(0), & \text{for } t = 0 \\ H_0(t)B_t + \sum_{n=1}^{N} H_n(t)S_n(t), & \text{for } t \geq 1. \end{cases} \]

Note: (i) \( V_0 \) is the initial value of the portfolio.
(ii) \( V_t \), \( t \geq 1 \) is the time \( t \) value of the portfolio before any transactions.
(iii) \( V_t \) is an adapted process, because if the subset of the relevant partition \( \mathcal{P} \) is known, then \( H_n(t), B_t, \) and \( S_n(t) \) are known, and \( V_t \) can be calculated.

The one period change in security price \( S_n(t) \) is
\[ \Delta S_n(t) = S_n(t) - S_n(t - 1). \]

If the investor has at the end of period \( t \) \( H_n(t) \) shares, the gain (or loss) from the security is
\[ H_n(t)\Delta S_n(t), \quad t \geq 1. \]

3.1.5 Self-Financing Strategies

A trading strategy is said to be self-financing if
\( V_t = V_0 + G_t, \quad t = 1, 2, \ldots, T. \)

That is, the time \( t \) value of the portfolio just before and just after any time \( t \) transactions are equal (no money is added or withdrawn).

A trading strategy \( H \) is self-financing if and only if
\( V_t = V_0 + G_t, \quad t = 1, 2, \ldots, T. \)

Proof. Exercise.

Example 3.3 (continued) For \( H \) to be self-financing, we must have
\[ V_t = \begin{cases} (1 + r)H_0(1)B_0 + 8H_1(1), & \text{for } t = 0 \\ (1 + r)H_0(t)B_t + 28H_1(t), & \text{for } t \geq 1. \end{cases} \]

3.1.6 Discounted Prices

The discounted price process \( S_n^*(t) = \{ S_n^*(t); t = 0, 1, \ldots, T \} \) is
\[ S_n^*(t) = S_n(t)/B_t, \quad t = 0, 1, \ldots, T, \quad n = 1, \ldots, N. \]

Example 3.3 (continued) Let \( B_t = (1+r)^t \), with \( r \geq 0 \) a constant, \( t = 0, 1, \ldots, T \). The value process is
\[ V_0 = H_0(1) + 5H_1(1), \quad \forall \omega \in \Omega \]
\[ V_1 = \begin{cases} (1 + r)H_0(1) + 8H_1(1), & \omega = \omega_1, \omega_2 \\ (1 + r)H_0(1) + 4H_1(1), & \omega = \omega_3, \omega_4 \end{cases} \]
\[ V_2 = \begin{cases} (1 + r)^2H_0(2) + 9H_1(2), & \omega = \omega_1 \\ (1 + r)^2H_0(2) + 6H_1(2), & \omega = \omega_2, \omega_3 \end{cases} \]
\( H_n(1) \) is known, then
\[ H_n(1) = (1 + r)H_n(2) + 3H_1(2), \quad \omega = \omega_4 \]

The gains process is
\[ G_1 = \begin{cases} rH_0(1) + 3H_1(1), & \omega = \omega_1, \omega_2 \\ rH_0(1) + H_1(1), & \omega = \omega_3, \omega_4 \end{cases} \]
\[ G_2 = \begin{cases} rH_0(1) + 3H_1(1) + r(1 + r)H_0(2) + H_1(2), & \omega = \omega_1 \\ rH_0(1) + 3H_1(1) + r(1 + r)H_0(2) - 2H_1(2), & \omega = \omega_2 \end{cases} \]
\[ rH_0(1) - H_1(1) + r(1 + r)H_0(2) + 2H_1(2), & \omega = \omega_3 \]
\[ rH_0(1) - H_1(1) + r(1 + r)H_0(2) - H_1(2), & \omega = \omega_4 \]

Again, because \( G = \{ G_t; t = 1, \ldots, T \} \) is a simple sum of adapted processes, it is itself an adapted stochastic process.

Discounted value process \( V^* = \{ V^*_t; t = 0, 1, \ldots, T \} \)
\[ V^*_t = \begin{cases} H_0(1) + \sum_{n=1}^{N} H_n(0)S_n^*(0), & t = 0 \\ H_0(t) + \sum_{n=1}^{N} H_n(t)S_n^*(t), & t = 1, \ldots, T \end{cases} \]

Discounted gains process \( G^* = \{ G^*_t; t = 1, \ldots, T \} \)
\[ G^*_t = \sum_{n=1}^{N} \sum_{u=1}^{t} H_n(u)\Delta S_n^*(u), \quad t = 1, \ldots, T, \]
where \( \Delta S_n^*(t) = S_n^*(t) - S_n^*(t-1) \).

Again, all these processes are adapted. Furthermore, \( V^*_t = V^*_0 + G^*_t \).

Verify these as exercises.
3.2 Return and Dividend Processes

Return process \( R_n = \{ R_n(t); t = 0, 1, \ldots, T \} \) with \( R_n(0) = 0 \) is defined for \( t = 1, \ldots, T \) as

\[
\Delta R_n(t) = \begin{cases} 
\Delta S_n(t)/S_n(t-1), & S_n(t-1) > 0 \\
0, & S_n(t-1) = 0.
\end{cases}
\]

(5)

Straightforwardly

\[
\Delta S_n(t) = S_n(t-1) \Delta R_n(t), t = 1, \ldots, T
\]

(6)

\[
S_n(t) = S_n(0) + \sum_{u=1}^{t} S_n(u-1) \Delta R_n(u), t = 1, \ldots, T
\]

(7)

\[
S_n(t) = S_n(0) \prod_{u=1}^{t} (1 + \Delta R_n(u)), t = 1, \ldots, T
\]

(8)

3.2.1 Returns for Discounted Price Processes

Using discounted prices in (5) returns, \( \Delta R_n^*(t) \), for the discounted prices. We can write these in terms of the undiscounted returns. Noting that \( S_n^*(t) = S_n(t)/B_t \), using (6) and \( B_t = (1 + r) B_{t-1} \), so

\[
S_n^*(t) = S_n(t)/B_t
\]

(9)

\[
= \frac{(1 + \Delta R_n(t)) S_n(t-1)/B_{t-1}}{1 + \Delta R_n(t)}
\]

(10)

\[
= 1 + \Delta R_n(t)
\]

(11)

\[
= \frac{S_n^*(t-1)}{S_n^*(t)} - 1
\]

(12)

\[
= \frac{\Delta R_n^*(t)}{1 + \Delta R_n^*(t)}
\]

(13)

Hence, finally

\[
\Delta R_n^*(t) = \frac{S_n^*(t)}{S_n^*(t-1)} - 1
\]

(14)

\[
= \frac{\Delta R_n(t)}{1 + \Delta R_n(t)}
\]

(15)

Furthermore

\[
S_n^*(t) = S_n^*(0) \prod_{u=1}^{t} [1 + \Delta R_n^*(u)]
\]

(16)

\[
= S_n(0) \prod_{u=1}^{t} \left[ 1 + \Delta R_n(u)/1 + \Delta R_n^*(u) \right] = S_n(t)/B_t
\]

(17)

3.2.2 Returns for the Value and Gains Process

Using

\[
H_n(t) \Delta S_n(t) = H_n(t) S_n(t-1) \Delta R_n(t)
\]

in the gains process, we can write

\[
G_i = \sum_{u=1}^{i} M_u(u) \Delta R_n(u) + \sum_{u=1}^{N} \sum_{i=1}^{N} M_u(u) \Delta R_n(u),
\]

where

\[
M_n(t) = \begin{cases} 
H_0(t) B_{t-1}, & n = 0 \\
H_0(t) S_n(t-1), & n = 1, \ldots, N
\end{cases}
\]

is the money invested in security \( n \) beginning at time \( t-1 \). Thus

\[
M = \{ M_n(t); t = 1, \ldots, T, n = 0, 1, \ldots, N \}
\]

gives the trading strategy in terms of invested money.

Let

\[
F_n(t) = \frac{M_n(t)}{V_{t-1}}, \quad n = 0, 1, \ldots, N
\]

denote fraction of money invested in security \( S_n \) at time \( t-1 \) then (especially) for the self-financing portfolio the return of the value process is

\[
\Delta R(t) = \frac{V_t - V_{t-1}}{V_{t-1}} = \sum_{n=0}^{N} F_n(u) \Delta R_n(t).
\]

Note that \( F_0(t) = 1 - F_1(t) \ldots - F_n(t) \) refers to the fraction invested in the bank account, so

\[
F = \{ F_n(t); t = 1, \ldots, T, n = 1, \ldots, N \}
\]
gives yet another way to describe trading strategies. Again \( F \) is adapted. The value process of the (self-financing) portfolio can be then written

\[
V_t = V_0 \prod_{u=1}^{t} [1 + \Delta h(u)] = V_0 \prod_{u=1}^{t} \left[ 1 + \sum_{u=0}^{N} F_n(u) \Delta R_n(u) \right].
\]

In summary, the trading strategy can be expressed in three different ways by \( H, M \) or \( F \).
For the value process we can calculate also the discounted versions. Particularly as with the price processes, we have

\[ \Delta R^*(t) = \frac{\Delta R(t) - \Delta R_0(t)}{1 + \Delta R_0(t)} \]

and

\[ V_t^* = V_0^* \prod_{u=1}^{t} \left[ 1 + \Delta R^*(u) \right] = V_0^* \prod_{u=1}^{t} \left[ 1 + \Delta R(u) \right] \].

Verify these as exercises.

3.2.3 Dividend Processes

Let

\[ D_n = \{ D_n(t); t = 0, 1, \ldots, T \} \]

denote the (cumulative) dividend process for security \( S_n \), \( n = 1, \ldots, N \), where \( D_n(0) = 0 \)
and let \( \Delta D_n(t) = D_n(t) - D_n(t-1) \) denote the dividend paid per security unit (dividend per share) at time point \( t \). Thus \( D_n(t) \) represents the cumulative dividend process up to time \( t \) and \( \Delta D_n(t) \) the payment at time \( t \).

The return series including dividends is then

\[ R_n(t) = \frac{\Delta S_n(t) + \Delta D_n(t)}{S_n(t-1)} \]

Using discounted price series gives

\[ R^*_n(t) = \frac{\Delta S^*_n(t) + \Delta D_n(t)/B_t}{S^*_n(t-1)} \].

3.3 Conditional Expectations and Martingales

In the multiperiod case the risk neutral probability measure, needed in the no-arbitrage pricing, is defined in terms of martingales (mathematical model for fair game). These in turn are defined in terms of conditional expectations.

Before continuing, let us define probability measure.

Let \( F \) be an algebra on \( \Omega \). We say that function \( P : F \to \mathbb{R} \) is probability if

\[
\begin{align*}
(P1) & \quad P(A) \geq 0, \quad \forall A \in F \\
(P2) & \quad P(\Omega) = 1 \\
(P3) & \quad A, B \in F \text{ and } A \cap B = \emptyset, \text{ then } P(A \cup B) = P(A) + P(B)
\end{align*}
\]

So given that we have \( P(\{\omega\}) > 0 \), for all \( \omega \in \Omega \) We can write

\[ P(Y = y) = P[Y(\omega) = y] = \sum_{\omega \in A_y} P(\{\omega\}) \]

where \( A_y = \{ \omega : Y(\omega) = y \} \). So we can write

\[ E[Y] = \sum_y yP(Y = y) = \sum_{\omega \in \Omega} Y(\omega)P(\{\omega\}). \]

In elementary statistics the conditional probability is defined in terms of the probabilities as

\[ P(A|B) = \frac{P(A \cap B)}{P(B)} \]

and the conditional expectation \( E(Y|A) \) as

\[ E(Y|A) = \sum_y yP(Y = y|A) = \frac{1}{P(A)} \sum_y yP(Y = y \cap A). \]
Example 3.3 (continued) Suppose $P(\omega) = 1/4$. Let $A = \{\omega: S_1(\omega) = 8\} = \{\omega_1, \omega_2\}$. Then $P(S_2 = 9 | A) = (1/4)/(1/4 + 1/4) = 1/2 = P(S_2 = 6 | A)$ and zero otherwise, so

$$E[S_2 | S_1 = 8] = E[S_2 | A] = 1/2 \times 9 + 1/2 \times 6 = 7.5$$

Similarly $E[S_2 | S_1 = 4] = 4.5$.

This is the method we can always use in practical calculations.

For general modeling purposes of stock prices as stochastic processes on filtered probability spaces, it is, however, more convenient to extend the definition and consider conditional expectations with respect to appropriate algebras. Before getting into the notation, let us make brief introduction.

A short introduction to conditional expectation

Consider a filtration $\mathcal{F} = \{\mathcal{F}_t: t = 0, 1, \ldots, T\}$ which is for us a model how information accumulates. Let $S_t$ denote stock price series, where $S_t$ is measurable on $\mathcal{F}_t$ (i.e., stochastic process $S_t$ is adapted). We can say that $S_t$ is observable on $\mathcal{F}_t$.

Example (coin tossing) $S(2)$ is observable on $\mathcal{F}_2$, because we can determine all the possible values of $S(2)$ on the information based on the events in $\mathcal{F}_2$. On the other hand $S(3)$ is not measurable on $\mathcal{F}_3$, and hence it is not observable on $\mathcal{F}_2$ (we cannot deduce it possible values on the basis of the events in $\mathcal{F}_3$). Furthermore, there are many other random variables $X$ that are also measurable on $\mathcal{F}_2$. For example $X = c$ (a constant), $S(1, \omega) = \omega \in \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and $S(1, \omega) = -1, \omega \in \{\omega_4, \omega_5, \omega_6, \omega_7\}$, and so on.

Consider next the usual (unconditional) expectation of $S_t$, $E[S_t]$, which is a real number. It can be considered as "the best possible" single real number to estimate $S_t$ in the case when one has no additional background information on $S_t$, except its (unconditional) distribution. That is, it can be interpreted "the best number" among all the real numbers to estimate $X$.

Suppose next that we have information $\mathcal{F}_{t-1}$, and consider (in principle) all random variables that are measurable on $\mathcal{F}_{t-1}$. Then each of these are observable on $\mathcal{F}_{t-1}$, and assume values depending on what set $A \in \mathcal{F}_{t-1}$ we select. Now to predict $S_t$, which is not observable on $\mathcal{F}_{t-1}$, our task is to select "the best possible" estimate of $S_t$ based on all the information in $\mathcal{F}_{t-1}$, which means that we select among all the random variables the one that best approximates $S_t$. That particular random variable will be the conditional expectation of $S_t$, given all the information $\mathcal{F}_{t-1}$, and will be denoted as $E[S_t | \mathcal{F}_{t-1}]$.

Thus in summary as the usual unconditional expectation, $E[S_t]$, selected the best estimate among the real numbers, the conditional expectation, $E[S_t | \mathcal{F}_{t-1}]$, can be interpreted to select a $\mathcal{F}_{t-1}$-measurable random variable that best approximates $S_t$, given all the information in $\mathcal{F}_{t-1}$.

### Example (continued)

Suppose that the coin is biased such that $P(\text{Heads}) = 1/4$ and $P(\text{Tails}) = 3/4$. Then what is $E[S(3) | \mathcal{F}_2]$? Denote for short $S = S(3)$. Then

| $\omega$ | $X_1$ | $P(X_1)$ | $P(S = 3 | X_1)$ | $E[S | X_1]$ |
|---|---|---|---|---|
| $\omega_1$ | 1 | $P(S = 3 | \omega_1) = 1$ | $E[S | \omega_1]$ |
| $\omega_2$ | 0 | $P(S = 3 | \omega_2) = 0$ | $E[S | \omega_2]$ |
| $\omega_3$ | 0 | $P(S = 3 | \omega_3) = 0$ | $E[S | \omega_3]$ |
| $\omega_4$ | 0 | $P(S = 3 | \omega_4) = 0$ | $E[S | \omega_4]$ |
| $\omega_5$ | 0 | $P(S = 3 | \omega_5) = 0$ | $E[S | \omega_5]$ |
| $\omega_6$ | 0 | $P(S = 3 | \omega_6) = 0$ | $E[S | \omega_6]$ |
| $\omega_7$ | 0 | $P(S = 3 | \omega_7) = 0$ | $E[S | \omega_7]$ |
| $\omega_8$ | 0 | $P(S = 3 | \omega_8) = 0$ | $E[S | \omega_8]$ |

So the general notation of the conditional expectation is $E[Y | \mathcal{G}]$, where $Y$ is measurable on an algebra $\mathcal{G}$ and $\mathcal{G} \subset \mathcal{F}$ is a sub-algebra in $\mathcal{F}$.

Another interpretation of $E[Y | \mathcal{F}]$ is that the unconditional expectation $E[Y]$ is the summary of $Y$ over all its possible values, the conditional expectation $E[Y | \mathcal{F}]$ is a summary of all its conditional expectations. The result is a random variable $X$ whose values will be constant on sets in the partition $\mathcal{P}$ related to $\mathcal{F}$. That is $X = E[Y | \mathcal{G}]$ such that for each $A \in \mathcal{P}$, $X_A = E[Y | \mathcal{G}]|_A$

$$= E[Y | A] = \sum_y P(Y = y | A),$$

where $1_A$ is an indicator function of $A$, such that

$$1_A = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{otherwise} \end{cases}$$
Thus we observe that the conditional expectation is a new random variable with associated probability distribution. For example in the case of $t = 2$:

\[ E[S(t)] \begin{pmatrix} \frac{-1}{16} \\ \frac{-1}{16} \\ \frac{3}{16} \\ \frac{2}{16} \end{pmatrix} \]

The conditional expectation obeys several properties. One of the most important for us is the law of iterated expectations

(10) If $\mathcal{F}_1 \subset \mathcal{F}_2$, then $E[E[Y|\mathcal{F}_2]|\mathcal{F}_1] = E[Y|\mathcal{F}_1]$. As an exercise verify this with the above coin toss example.

A special case is obtained by selecting in (10) $\mathcal{F}_1 = \mathcal{F}_0 = (\Omega, \emptyset)$. Then obviously $E[Y|\mathcal{F}_0] = E[Y]$, and because for any algebra $\mathcal{F}$, $\mathcal{F}_0 \subset \mathcal{F}$, we have

\[ E[E[Y|\mathcal{F}]] = E[E[Y|\mathcal{F}_0]] = E[Y|\mathcal{F}_0] = E[Y]. \]

Note: The order in (10) does not matter. I.e.,


Example (continues)

\[ E[E[S|F_2]] = \frac{-1}{2} \cdot \frac{3}{16} + \frac{1}{2} \cdot \frac{2}{16} = \frac{3}{16} \]

The conditional expectation has also the following properties:

Let $X_1$, $X_2$, $Y_1$, and $Y_2$ be random variable such that $X_1, X_2 \in \mathcal{F}$ then

(12) $E[X_1Y_1 + X_2Y_2|\mathcal{F}] = X_1E[Y_1|\mathcal{F}] + X_2E[Y_2|\mathcal{F}]$.

(13) If $X \in \mathcal{F}$ then $E[X|\mathcal{F}] = X$.

Furthermore, if $X$ is independent of $\mathcal{F}$ then

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

Note: $X$ is independent of $\mathcal{F}$ if and only is $P((X = x)|\mathcal{F}) = P(X = x)$ for all $x \in R$ and $A \in \mathcal{F}$.

A general mathematical definition of the conditional expectation is that given an arbitrary random variable $Y$, the conditional expectation $E[Y|\mathcal{F}]$ is the unique random variable such that

(15) \[
\begin{align*}
(a) & \quad E[Y|\mathcal{F}] \in \mathcal{F} \\
(b) & \quad E[E[Y|\mathcal{F}]A] = E[Y_1A], \quad \forall A \in \mathcal{F}
\end{align*}
\]

Martingales

Let $S = \{S(t); t = 0, 1, \ldots, T\}$ be an adapted stochastic process with respect to filtration $\mathcal{F} = \{F_t; t = 0, 1, \ldots, T\}$ The process $S$ is a martingale if

\[ E[S(t + 1)|\mathcal{F}_t] = S(t) \]

$S$ is submartingale if

\[ E[S(t + 1)|\mathcal{F}_t] \leq S(t) \]

and supermartingale if

\[ E[S(t + 1)|\mathcal{F}_t] \geq S(t). \]

Note: Using the iterated expectations formula given by (10), if $S_t$ is a martingale then

\[ E[S(t + s)|\mathcal{F}_t] = S(t), \; \forall s \geq 0. \]

The same is true for the sub- and supermartingales.

Example: If in the coin toss example $P(H) = P(T) = 1/2$ the $S(t)$ process is a martingale.

Note further that if $S$ is martingale as above, then we easily see that

\[ E[S(t)] = E[E[S(t)|\mathcal{F}_{t-1}]] = E[S(t-1)] = \cdots = E[S(0)], \]

i.e., a martingale is 'constant on average'. If $S(t)$ is supermartingale then

\[ E[S(t)] = E[E[S(t)|\mathcal{F}_{t-1}]] \leq E[S(t-1)] \leq \cdots \leq E[S(0)], \]

so supermartingale is 'decreasing on average', and deducing in similar fashion, submartingale is 'increasing on average'.

If $S_t$ is martingale then we, again, see easily that

\[ E[\Delta S(t + 1)|\mathcal{F}_t] = 0, \]

where $\Delta S(t + 1) = S(t + 1) - S(t)$, called martingale difference.
3.4 Economic Considerations

Trading strategy $H$ is an arbitrage opportunity if

(a) $V_0 = 0$
(b) $V_T > 0$
(c) $EV_T > 0$
(d) $H$ is self-financing.

In terms of the discounted value process $V_t^*$ defined by (3) we have:

The self-financing trading strategy is an arbitrage opportunity if and only if

(a) $V_t^* = 0$
(b) $V_T^* > 0$

or equivalently (because of (4))

(a) $G_T^* > 0$
(b) $EG_T^* > 0$.

3.4 Economic Considerations (continued) $B_t = (1 + r)^t$, $r \geq 0$ constant. If $Q$ is martingale measure, then

$t=0, s=1$: $5(1 + r) = 8Q(\omega_1) + 8Q(\omega_2) + 4Q(\omega_3) + 4Q(\omega_4)$
$t=0, s=2$: $5(1 + r) = 9Q(\omega_1) + 6Q(\omega_2) + 6Q(\omega_3) + 3Q(\omega_4)$
$t=1, s=1$: $8(1 + r) = [6Q(\omega_1) + 3Q(\omega_2)]/[Q(\omega_1) + Q(\omega_2)]$
$1 = Q(\omega_1) + Q(\omega_2) + Q(\omega_3) + Q(\omega_4)$

Solving the equations (a system of three of the first equations and the last one) yields

$Q(\omega_1) = \frac{1}{2}(1 + 5r)(1 + 4r)$
$Q(\omega_2) = \frac{1}{2}(1 + 5r)(1 - 2r)$
$Q(\omega_3) = \frac{1}{2}(1 - 5r)(1 + 4r)$
$Q(\omega_4) = \frac{1}{2}(1 - 5r)(1 - 2r)$

From these we observe that $Q(\omega) > 0$ if $0 \leq r < 1/8$, and the martingale measure exists. Otherwise it does not exist.

The main result of this section is:

(3.19) There are no arbitrage opportunities if and only if there exists a martingale measure.

Example 3.3 (continued) If $B_t = 1$, $r = 0.125$, no arbitrage opportunities exist. Suppose $B_0 = (1 + r)^0$ with $r \geq 12.5\%$. Let $V_0 = 0$, and if at $t = 1$ $S_1 = 8$ sell one unit of it short and save the amount to the bank account $B_t$. (Thus the strategy at time $t = 2$ with $S_1(1) = 8$ is $H_1(2) = 1$ and $H_1(2) = -1$). Then

$V_2 = \begin{cases} (1 + r)^2 H_0(2) + 9 H_1(2) = 8(1 + r) - 9 > 0, \omega = \omega_1 \\ (1 + r)^2 H_0(2) + 6 H_1(2) = 8(1 + r) - 6 > 0, \omega = \omega_2, \end{cases}$

i.e., arbitrage opportunity.

Risk neutral probabilities

In single period market the risk neutral probabilities were defined in terms of ordinary expectations. In multiperiod markets they are defined in terms of conditional expectations, or more precisely in terms of martingales. A risk neutral probability measure (or martingale measure) is a probability measure $Q$ such that

(1) $Q(\omega) > 0$ for all $\omega \in \Omega$
(2) $E_Q[S_n(t + s) | F_t] = S_0^*(t)$, $t, s \geq 0$

for every $n = 1, \ldots, N$.

That is

(18) $E_Q[B_tS_n(t + s) | B_t] = S_0^*(t)$, $t, s \geq 0$.

The proof of this result is similar as in the single period models. The easy part is to show the sufficiency direction. For the purpose we derive the following useful results

If $Z$ is a martingale and $H$ is a predictable process, then

(20) $G_t = \sum_{u=1}^{t} H_u \Delta Z_u$

is also a martingale.

To see this we apply the law of iterated expectations given in (10). First we observe

$E[H_{t+1} \Delta Z_{t+1} | F_t] = H_{t+1} E[\Delta Z_{t+1} | F_t]$  

because $H$ is predictable. Furthermore, because

$\Delta Z_{t+1} = Z_{t+1} - Z_t$

and $Z$ is martingale

$E[\Delta Z_{t+1} | F_t] = E[Z_{t+1} | F_t] - E[Z_t | F_t] = Z_t - Z_t = 0$

so

$E[G_{t+1} | F_t] = E[G_t + H_{t+1} \Delta Z_{t+1} | F_t] = G_t + H_{t+1} E[\Delta Z_{t+1} | F_t] = G_t$

which proves the assertion.
The other important result is

\[ (3.21) \] If \( Q \) is a martingale measure and \( H \) is a self-financing strategy, then \( V^* \), the discounted value process corresponding to \( H \), is a martingale under \( Q \).

This follows immediately from (20) when we write \( V^* \) in the form given in (4), i.e.,

\[ V^n_t = V^n_0 + G^n_t \]

with

\[ G^n_t = \sum_{i=1}^{n} H^n_i \Delta S^n_i \]

Now we are finally ready to demonstrate the sufficiency part of (3.19).

Let \( Q \) be a martingale measure, \( H \) a self-financing strategy, and suppose there exist an arbitrage opportunity \( V^* \), a portfolio based on \( H \) such that \( V^n_0 = 0, V^n_1 > 0 \) and \( E_Q[V^n_2] > 0 \). However, since \( V^* \) is martingale under \( Q \), we have \( 0 = V^n_0 = E_Q[V^n_1 | \mathcal{F}_0] = E_Q[V^n_2] > 0 \), a contradiction. So there cannot be arbitrage opportunity.

The necessary part is more tedious and technical. Although it is instructive and demonstrates how the martingale measure can be in principle constructed, we skip it here, and give only an example.

Before that consider the following yet another important result.

\[ (3.22) \] If the multiperiod model does not have any arbitrage opportunities, then none of the underlying single period model has any arbitrage opportunities in the single period sense.

Intuitively this is quite obvious. Because, (by a bit oversimplifying) if there were a period with an arbitrage opportunity when a particular event occurs, we could just simply sit an wait unless the time comes for the period, make a single period arbitrage position, close it after the period, have the arbitrage profit and again wait until the end of the total period \( T \) with the profit in pocket.

Following the above (simplified) logic, the risk neutral probability measure can be constructed considering each time point as a single market model an calculating risk neutral probabilities. These prove to be conditional probabilities of the desired martingale probability measure. That is, given the partition \( \mathcal{P}_t = \{ A_{i1}, \ldots, A_{in} \} \) of \( \mathcal{F}_t \), with \( n \) the number of set in \( \mathcal{P} \). Then using the risk neutral probabilities for moving from \( t \) to \( t+1 \), the martingale measure can be calculated (in principle) recursively as follows. Let \( Q(A_{i1} | \mathcal{A}_t) \) denote the risk neutral probability between time points, given the observed event is \( A_{i1} \). At each step, because of the hierarchy \( A_{i1+1} \cap A_{i2} \) is either \( A_{i1+1} \) or the \( 0 \). Consequently

\[ Q(A_{i1+1} | \mathcal{A}_t) = Q(A_{i1})Q(A_{i1+1} | A_{i1}) \]

Thus we have an algorithm to generate the risk neutral probability measure as

\[ \text{(*)} \]

\[ Q(A_{i1+1}) = Q(A_{i1})Q(A_{i1+1} | A_{i1}) \]

Example 3.3 (continued) \( B_t = (1+r)^t, r \geq 0 \) constant. Starting \( \mathcal{F}_0 = \emptyset, \mathcal{A}_0 = \Omega, \mathcal{A}_1 = \{ \omega_1, \omega_2 \}, \mathcal{A}_2 = \{ \omega_1, \omega_2, \omega_3 \} \). Using the risk neutral principle

\[ S_0 = E_Q [S_1/(1+r)] = \frac{S_1(\omega_1, \omega_2)}{1+r} \]

i.e.,

\[ \delta = \frac{S_1(\omega_1, \omega_2)}{1+r} \]

Solving for \( q \) gives

\[ q = \frac{1+5r}{4} \]

provided that \( 0 \leq r < 4/5 \). Thus

\[ Q(\omega_1, \omega_2) = Q(\omega_1, \omega_2) \Omega = q = \frac{1+5r}{4} \]

and

\[ Q(\omega_2, \omega_3) = 1 - q = \frac{3-5r}{4} \]

In a similar fashion for \( t = 1 \) one obtains

\[ Q[(\omega_1)](\omega_1, \omega_2) = \frac{2}{3}(1+4r) \]

\[ Q[(\omega_2)](\omega_1, \omega_2) = \frac{4}{3}(1-8r) \]

\[ Q[(\omega_2)](\omega_2, \omega_3) = \frac{1}{3}(1+4r) \]

\[ Q[(\omega_2)](\omega_2, \omega_3) = \frac{1}{3}(1-4r) \]

which are admissible nonzero probabilities if \( 0 \leq r < 1/8 \). Using (*) gives, for example

\[ Q[(\omega_1)] = Q[(\omega_1, \omega_2)]Q[(\omega_1)](\omega_1, \omega_2) \]

\[ = \frac{2}{3}(1+5r) \times \frac{2}{3}(1+4r) \]

\[ = \frac{4}{3}(1+5r) \times \frac{1}{3}(1+4r) \times \frac{1}{3}(1-4r). \]

That is, exactly the same we obtained already earlier.

The martingale measure can be defined also in terms of the return process. If \( Q \) is a martingale measure then

\[ E_Q[S^n_{t+1} | \mathcal{F}_t] = S^n_t \]

we have

\[ (21) \]

\[ E_Q[\Delta S^n_t | \mathcal{F}_t] = 0. \]

Furthermore, since

\[ \Delta_R^n(t+1) = \Delta S^n_t(t+1) / S^n_t(t) \]

so that \( \Delta S^n_t(t+1) = \Delta_R^n(t+1) S^n_t(t) \), and hence, using (21)

\[ 0 = E_Q[\Delta S^n_t(t+1) | \mathcal{F}_t] = S^n_t(t) E_Q[\Delta R^n(t+1)| \mathcal{F}_t] \]
Because $S_n(t) > 0,$
\begin{equation}
E_Q[\Delta R_n^*(t+1)|\mathcal{F}_t] = 0^\dagger
\end{equation}
implies that $\Delta R_n^*(t)$ is a martingale difference wrt $Q,$ and hence, $R_n^*(t)$ is martingale wrt $Q.$ In summary we have
\begin{equation}
(3.25) \text{ The strictly positive probability measure } Q \text{ is a martingale measure if and only if } R_n^*(t) \text{ is a martingale with respect to } Q, \ n = 1, \ldots, N.
\end{equation}

Note, as seen earlier, we can also write
\begin{equation}
\Delta R_n^*(t+1) = \frac{\Delta R(t+1) - \Delta R_0(t+1)}{1 + \Delta R_0(t+1)}.
\end{equation}

$^\dagger$Note that $E[R_n^*(t)] = 0,$ because $E[R_n^*(t+1)] = E[E[R_n^*(t+1)|\mathcal{F}_t]]$ by the properties of conditional expectations.

3.5 The Binomial Model

Each period there are two possibilities; up by a factor $u > 1$ or down by a factor $d, 0 < d < 1.$

\begin{figure}
\begin{center}
\begin{tikzpicture}
\node (S_t) at (0,0) {$S_t$};
\node (t) at (-1,1) {$t$};
\node (t+1) at (1,1) {$t + 1$};
\node (uS_t) at (2,1) {$uS_t$};
\node (dS_t) at (2,-1) {$dS_t$};
\draw (t) edge (S_t);
\draw (S_t) edge (uS_t);
\draw (S_t) edge (dS_t);
\end{tikzpicture}
\end{center}
\caption{A node in the binomial model.}
\end{figure}

If we consider the changes at each step independent and
\begin{equation}
S(t+1) = \begin{cases} uS(t), \text{ with probability } p \\ dS(t), \text{ with probability } 1 - p \end{cases}
\end{equation}
it leads to a Bernoulli process.

\begin{example}
\text{Dividend paying stocks}
\end{example}

With dividends
\begin{equation}
\Delta R_n(t+1) = \frac{\Delta S_n(t+1) + \Delta D_n(t+1)}{S_n(t)}
\end{equation}
and using discounted series
\begin{equation}
\Delta R_n^*(t+1) = \frac{\Delta S_n^*(t+1) + \Delta D_n(t+1)/B_{t+1}}{S_n^*(t)}
\end{equation}
Now, in order that the (discounted) return process $R_n^*(t)$ to be martingale wrt $Q,$ it must hold
\begin{equation}
(3.27) \text{ if } Q \text{ is a risk neutral probability measure, then for any } t \geq 0, s \geq 1
\end{equation}
\begin{equation}
E_Q[S_n^*(t+s) + \Delta D_n(t+s)/B_{t+s+1}] = S_n^*(t), \quad t < T.
\end{equation}

Using this recursively with the law of iterated expectations, we can generalize (3.25) to include dividend paying stocks, i.e.,
\begin{equation}
(3.28) \text{ if } Q \text{ is a risk neutral probability measure, then for any } t \geq 0, s \geq 1
\end{equation}
\begin{equation}
E_Q[S_n^*(t+s) + \Delta D_n(t+s)/B_{t+s+1}] = S_n^*(t), \quad t < T.
\end{equation}

Let $X = \{X_t; t = 1, 2, \ldots, T\},$ with $X_t$ either 0 or 1, is a Bernoulli process, if
\begin{equation}
P(X_t = 1) = 1 - P(X_t = 0) = p
\end{equation}
and $X_t$'s are independent (coin tossing process with probability of head, say, equal to $p$).

In our application we restrict the "number of flips" to $T,$ so that we observe sequences of $T$ tosses, like
\begin{equation}
\omega = (0, 1, 0, 1, 1, 0, \ldots, 1).
\end{equation}

Thus the sample space $\omega \in \Omega$ consists of $2^T$ vectors like this. If $n$ denotes the number of 1's then
\begin{equation}
P(\omega) = p^n (1 - p)^{T-n}.
\end{equation}

Let
\begin{equation}
N_t(\omega) = X_1(\omega) + \cdots + X_t(\omega),
\end{equation}
then we have a process $\{N_t; t = 1, \ldots, T\}$ where $N_t$ indicates the number of heads in $t$ first $t$ coin flips.
The distribution of $N_t$ is the binomial distribution such that

$$P(N_t = n) = \binom{t}{n} p^n (1-p)^{t-n}, \quad n = 0, 1, \ldots, t.$$  

with

$$\binom{t}{n} = \frac{t!}{n!(t-n)!}$$

where $t!(t-1)(t-2)\ldots1$ (note that 0! = 1).

The binomial security price model becomes

$$S_t = S_0 u^N t d^{-N_t}, \quad t = 1, \ldots, T,$$

where $N_t$ refers to the number of up events.

The restriction above is that $0 < q < 1$, so there exists a risk neutral measure if and only if $u > 1 + r > d$.

Probability $q$ gives the conditional probability given information $\mathcal{F}_t$, so the unconditional probability, i.e. the martingale probability measure, becomes according to the binomial law

$$(32) \quad Q_t = Q(S_t = S_0 u^N t d^{-N_t}) = \binom{t}{n} q^n (1-q)^{t-n}.$$  

with $n = 0, 1, \ldots, t$.

Example. Consider a European call option with expiry $T$ and strike or exercise price $K$ written on (one share of) stock $S$. Let $B_t = 1 + r, \ r \geq 0$ a constant. Then the arbitrage free price process for the call option $C(t)$ is under the binomial model as follows: At $t = 0$

$$C(0) = \frac{1}{(1+r)^T} E_q [(S_T - e)^+ | \mathcal{F}_0] = \frac{1}{(1+r)^T} E_q [(S_T - e)^+],$$

where $(S_T - K)^+ = \max(S_T - e, 0)$. Now $S_T = S_0 u^N t d^{-N_t}$. So using the binomial model

$$C(0) = \frac{1}{(1+r)^T} \sum_{n=0}^{T} \binom{T}{n} q^n (1-q)^{T-n} (S_0 u^N t d^{-N_t} - e)^+$$

At time $t = 1$ we have the above case, where the expiry is at $T = 1$ and observed market price $S_1$. For general $t$ we have the above case with expiry $T = 1$ and observed market price $S_t$. So the arbitrage free price process as $C(t)$ is

$$C(t) = \frac{1}{(1+r)^T} \sum_{n=0}^{T-t} \binom{T-t}{n} q^n (1-q)^{T-t-n} (S_0 u^N t d^{-N_t} - e)^+$$

Using the binomial model

$$(30) \quad P(S_t = S_0 u^N t d^{-N_t}) = \binom{t}{n} q^n (1-p)^{t-n}, \quad n = 0, 1, \ldots, t$$

Equivalently we could work with the returns such that

$$(31) \quad \Delta R(t) = u X_t q^{1-X_t} - 1, \quad t = 1, \ldots, T,$$

so that, either $\Delta R(t) = u$ or $\Delta R(t) = d$.

Now assuming $B_t = 1 + r, \ r \geq 0$ constant, the arbitrage free condition implies (with [conditional] risk neutral probabilities $q$, given information $\mathcal{F}_t$) in period $t$ to $t+1$

$$S_{t+1} = E_q [\frac{S_t}{1+r} = q S_t u + (1-q) S_t d \frac{1}{1+r}]$$

so that

$$1 = q \frac{u}{1+r} + (1-q) \frac{d}{1+r}$$

or

$$0 = q \frac{u - r}{1+r} + (1-q) \frac{d - r}{1+r}$$

From which

$$q = \frac{1 + r - d}{u - d}.$$  

Implementing the Binomial Model

First, note that if $X$ is binomially distributed, with the probability parameter $p$ and the number of experiments $n$, so that we can denote $X \sim \text{Bin}(m, p)$ then

$$E[X] = mp$$

and

$$\text{var}[X] = mp(1-p)$$

Accepting the binomial model as an approximation to the stock returns we can implement the model as follows:

Suppose the time period for the stock observed stock prices is $[0, T]$. We can always scale such that $T = 1$ (for example day, month or year). Divide the interval to $n$ equal subintervals $0 = t_0 < t_1 < \cdots < t_n = 1$ of equal length such that

$$\Delta t = t_i - t_{i-1} = \frac{1}{n}$$

The binomial model implies (30).
Then obviously \( u \) and \( d \) must depend on the density of the division. So denote \( u = u(\Delta t) \) and \( d = d(\Delta t) \). In the same fashion the probability \( p \) in (28) can be assumed to depend on the length of the interval. For convenience write

\[
u(\Delta t) = e^{\sigma \sqrt{\Delta t}}, \quad d(\Delta t) = e^{-\sigma \sqrt{\Delta t}}
\]

where \( \sigma \) is the standard deviation of the returns, such that

\[\sigma^2 = \text{var}(\log S_t / S_0)\]

with \( \sigma^2 \) denoting the "one period" (e.g. day, month, year) variance. Write further

\[p = \frac{1}{2}(1 + a \Delta t),\]

where \( a \) is an appropriate constant.

In these notations, using (30) with \( X = \sum_{i=1}^{m} X_i \), we find

\[\log(1 + \Delta R) = \log S_0 = Xu + (m - X)d\]

where \( X \sim \text{Bin}(m,p) \). So

\[
\text{var}[X\log(u) + (m - X)\log(d)] = np(1 - p)[\log(u/d)]^2
\]

\[= np(1 - p)[\log(e^{\sigma \sqrt{\Delta t}})]^2 = np(1 - p)\sigma^2 \Delta t = p(1 - p)\sigma^2\]

\[\rightarrow \sigma^2 = \text{var}[\log S_t / S_0],\]

because \( p = \frac{1}{2}(1 + \Delta t) \rightarrow \frac{1}{2} \) as \( m \rightarrow \infty \).

3.6 Markov Models

Makrov property: 'Future is independent of the past, given the present values of the process'. That is given filtration \( \mathcal{F} = \{F_t; t = 0, 1, \ldots, T\} \) generated by the process \( X = \{X_t; t = 0, 1, \ldots, T\} \)

\[E[X_{t+1}|\mathcal{F}_t] = E[X_{t+1}|X_t].\]

More precisely, assume the process \( X \) can take values in some finite set \( E \), called state space. If \( X = \{j\} \) is in state \( j \). We assume that there is some sample space \( \Omega \) and a probability measure \( P \) on it (i.e., for any \( \omega \in \Omega \), \( P(\{\omega\}) > 0 \)) and \( \mathcal{F}_t \) is generated by the present and past values of \( X \).

Then, generally we say that the stochastic process is a Markov chain if

\[P(X_{t+1} = j|\mathcal{F}_t) = P(X_{t+1} = j|X_t).\]

This generalizes to

(33) \[P(X_{t+s} = j|\mathcal{F}_t) = P(X_{t+s} = j|X_t)\]

Given the risk free return \( r \), the risk neutral probability is then

\[q = \frac{1 + r \Delta t - e^{-\sigma \sqrt{\Delta t}}}{e^{\sigma \sqrt{\Delta t} - 1}}\]

Now it is easy to implement the binomial model. The input needed is: current stock price \( S_0 \), exercise price \( e \), annual interest rate \( r \), stock price volatility \( \sigma \) (annual), time to maturity \( T \), and number of periods \( n \).


Note. The conditional probability \( P[Y = y|\mathcal{F}] \) is defined by

\[P[Y = y|\mathcal{F}] = E[1(Y=y)|\mathcal{F}]\]

where \( \{Y = y\} = \omega \in \Omega : Y(\omega) = y \).

Using the law of iterated expectations and the definition of the conditional probability we find by the Markov property \( P[X_{t+s} = j|\mathcal{F}_t] = P[X_{t+s} = j|X_t] \) that

\[E[1(X_{t+s} = j)|\mathcal{F}_t] = E[1(X_{t+s} = j)|X_t].\]

Using this and the law of iterated expectations we can prove the claim iteratively: For \( s = 1 \) the claim follows by definition, and is presented above. For \( s > 1 \)

\[P(X_{t+s} = j|\mathcal{F}_t) = E[1(X_{t+s} = j)|\mathcal{F}_t] = E[E[1(X_{t+s} = j)|X_t]|\mathcal{F}_t] = E[E[1(X_{t+s} = j)|X_t]|X_t].\]

Now \( E[1(X_{t+s} = j)|X_{t+s}] = P[X_{t+s} = j|X_{t+s}] \) is a function of \( X_{t+s} \). For convenience, write \( f(X_{t+s}) = E[1(X_{t+s} = j)|X_{t+s}] \). So

\[P(X_{t+s} = j|\mathcal{F}_t) = E[f(X_{t+s})|X_t] = E[E[f(X_{t+s})|X_{t+s}]|X_t] = E[f(X_{t+s})|X_t].\]

(by definition of expectation) \( = \sum_x f(x)P(X_{t+s} = x|\mathcal{F}_t) \)

(Markov property) \( = \sum_x f(x)P(X_{t+s} = x|X_t) \)

\( = E[f(X_{t+s})|X_t] \)

(Iterated exp) \( = E[1(X_{t+s} = j)|X_t] \)

\( = P(X_{t+s} = j) \)

Thus (30) holds for \( s = 2 \). Just following the same procedure we find that (30) true for arbitrary \( s \geq 0 \).
The Markov chain is said to be **stationary** or time-homogeneous if the conditional probabilities \( P\{X_{t+1} \mid F_t\} \) do not depend on time \( t \). In this case we can define the **transition probabilities**

\[
P(i, j) = P\{X_{t+1} = j \mid X_t = i\},
\]

the probability that the system moves to state \( j \) given that it was in state \( i \). The **transition matrix** is then

\[
P = [P(i, j)].
\]

We observe that the each row sums to one, i.e.,

\[
\sum_j P(i, j) = 1.
\]

In the case of non-stationary Markov chain

\[
P_t(i, j) = P\{X_{t+1} = j \mid X_t = i\}
\]

depends on time \( t \), implying a distinct transition matrix for each time point

\[
P_t = [P_t(i, j)].
\]

Example 3.5 Let \( N = \{N_t; t = 1, 2, \ldots, T\} \) with \( N_t \) denoting the number of head in \( t \) independent flips of a coin. If \( P(H) = p \) then (\( 1 < t < T \))

\[
P(N_{t+1} = j \mid F_t) = \begin{cases} p, & \text{if } j = N_t + 1 \\ 1 - p, & \text{if } j = N_t \\ 0, & \text{otherwise} \end{cases}
\]

So obviously the conditional probability depends only on the previous value of the process. That is

\[
P(N_{t+1} = j \mid F_t) = P(N_{t+1} = j \mid N_t)
\]

The following is an important property of Markov chains

(3.35) **If** \( Y = f(X_t, X_{t+1}, \ldots, X_T) **for some function** \( f **, then**

\[
E[Y \mid F_t] = E[Y \mid X_t]
\]

To see this we observe

\[
E[Y \mid F_t] = \sum_{x_t, x_{t+1}, \ldots, x_T} f(x_t, x_{t+1}, \ldots, x_T)P(X_t = x_t, X_{t+1} = x_{t+1}, \ldots, X_T = x_T \mid F_t)
\]

Exercise. Verify (3.35)

The main result for us is (without explicit proof):

(3.36) **Suppose**

(i) there are no arbitrage opportunities,

(ii) the discounted price process \( S^* \), is a Markov chain under \( P \),

(iii) the filtration \( \mathcal{F}^* \) is generated by \( S^* \),

then there exist a martingale measure \( Q \) under which \( S^* \) is a Markov chain.