Chapter 20

MSMYGF Mathematics Of Finance

(20.1) Statistical Considerations

Modelling of financial processes is done using differential equations which also take into account randomness. The randomness is expressed using statistical (or rather probabilistic) techniques, so it is first necessary to introduce these.

(20.1.1) σ Algebra

As was the case in Chapter 6 probability is thought of as a function P acting on a set of events $\mathcal{F}$ which consists of subsets of the sample space Ω. It is necessary for the elements of $\mathcal{F}$ to have certain properties. It is these properties which define $\mathcal{F}$ as a σ algebra.

Definition 1 *Let Ω be a set and let $\mathcal{F}$ be a set of subsets of Ω. $\mathcal{F}$ is a σ algebra on Ω if*

1. $\emptyset \in \mathcal{F}$
2. $F \in \mathcal{F} \Rightarrow F^c \in \mathcal{F}$
3. $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

From DeMorgan's laws it is readily seen that

$$\bigcap_{i=1}^{\infty} A_i = \left(\bigcup_{i=1}^{n} A_i^c \right)^c \in \mathcal{F}$$

The concept of a set of subsets of some other set is rather involved, and indeed somewhat abstract from the elements of Ω, the elementary events. This is where σ algebrae are useful: they can be used to build intricate structures for holding information. In the case of probability this relates to which elementary events constitute a 'real' event.

Clearly the presence of a set A in a σ algebra dictates that $\Omega, \emptyset$, and A^c must also be in the algebra. The set $\{\emptyset, \Omega, A, A^c\} = \mathcal{F}(A)$ is called the minimal algebra of A. The smallest algebra is $\{\emptyset, \Omega\}$ which is called the trivial algebra or the degenerate algebra, and often denoted by $\mathcal{O}$.

Definition 2 *The pair $(\Omega, \mathcal{F})$ is called a measurable space.*

Readers familiar with measure and integration will feel quite at home with this. Other readers should just accept this, as its importance is low. On a measurable space it is possible to define a measure (unsurprisingly), and defining the probability measure P will yield the familiar probability space $(\Omega, \mathcal{F}, P)$.

(20.1.2) The Probability Space

Definition 3 *A probability measure P on a measurable space $(\Omega, \mathcal{F})$ is a function $P\colon \mathcal{F} \to [0,1]$ such that*

1. $P(\emptyset) = 0$
2. $P(\Omega) = 1$
3. *If $A_1, A_2, \dots \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$ for all $i \neq j$ then* $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right)$

It is possible to define measures other than the probability measure. These differ by the omission of the condition $P(\emptyset) = 0$. Measures are in fact quite common. Distance in $\mathbb{R}$, area in $\mathbb{R}^2$, volume in $\mathbb{R}^3$, and the cardinality of A for $A \in \mathcal{F}$ are all examples of measures.

Enough information has now been presented to make meaningful the probability space $(\Omega, \mathcal{F}, P)$. Note the events—the subsets of Ω that are in $\mathcal{F}$—are called $\mathcal{F}$ measurable. The probability space has the following features

- Anything that happens in 'real life' corresponds to a single element $\omega \in \Omega$, an elementary event in the sample space.
- Meaningful events are represented by collections of elementary events, the sets in $\mathcal{F}$. These are called events.

Note it is not the case that $\sum_{A_i \in \mathcal{F}} P(A_i) = 1$. This is because in general $\bigcup_{A_i \in \mathcal{F}} A_i$ is not a disjoint union. Note the word *if* in Definition 3.

Consider now some function $Y\colon \Omega \to \mathbb{R}^n$. By taking elements from Ω it is implicit that some event in $\mathcal{F}$ is chosen. The set $\{\omega \in \Omega \mid Y(\omega) = \mathbf{x}\}$ for some fixed $\mathbf{x} \in \mathbb{R}^n$ defines an event which may or may not be in $\mathcal{F}$. Y is said to be $\mathcal{F}$ measurable if these are always events. Formally,

Definition 4 *Let U by any open subset of $\mathbb{R}^n$. The function $Y\colon \Omega \to \mathbb{R}^n$ is called $\mathcal{F}$ measurable if*

$$Y^{-1}(U) = \{\omega \in \Omega \mid Y(\omega) \in U\} \in \mathcal{F}$$

In $\mathbb{R}$ this means that the set

$$\{\omega \in \Omega \mid a < Y(\omega) < b\} \in \mathcal{F} \quad \forall a, b \in \mathbb{R},\ a < b$$

In fact it is more generally the case that if Y is measurable on the σ algebra $\mathcal{G}$ then $\{\omega \mid X(\omega) < X\} \in \mathcal{G}$ $\forall x \in \mathbb{R}$.

Definition 5 *A random variable X on a probability space $(\Omega, \mathcal{F}, P)$ is an $\mathcal{F}$ measurable function $X\colon \Omega \to \mathbb{R}^n$.*

The condition of $\mathcal{F}$ measurability prevents X giving elementary events from the same event (in $\mathcal{F}$) different values.

Now, P acts on $\mathcal{F}$, so implicitly acts on some $\omega \in A \subset \mathcal{F}$. As X acts on Ω, there must be some function $f\colon \mathbb{R}^n \to \mathbb{R}$ which somehow represents the probability of X assuming its various values.

The measure $\mu_X\colon \mathbb{R}^n \to \mathbb{R}$ which corresponds to P is defined by $\mu_X(B) = P(X^{-1}(U))$. Because X is $\mathcal{F}$ measurable $X^{-1}(U) \in \mathcal{F}$ and so $P(X^{-1}(U))$ is defined.

In the one dimensional case μ_X is the cumulative distribution function, $F_X(x)$. It represents the probability that any of the ωs in $X^{-1}(B)$ happen, which translates to $P(X \leqslant x)$. The probability density function is defined to be the non-negative function f for which $\int_{\mathbb{R}} f(x)\,\mathrm{d}x = 1$ and $F_X(x) = \int_{-\infty}^{x} f(y)\,\mathrm{d}y$.

On the probability space on which they are defined all random variables are $\mathcal{F}$ measurable. However, it may be the case that some of the sets in $\mathcal{F}$ are not required for measurability. This begs the question which sets are required in $\mathcal{F}$. The smallest σ algebra with respect to which a random variable X is measurable is denoted $\mathcal{F}(X)$.

Example 6 *Consider throwing two coins, so that $\Omega = \{HH, HT, TH, TT\}$. Now let X be the random variable defined by*

$$X(\omega) = \begin{cases} 1 & \text{if the first coin is } H \\ 0 & \text{if the first coin is } T \end{cases}$$

The smallest σ algebra defined by X is the set of subsets of Ω—not the set $\{H, T\}$—such that

$$\{\omega \in \Omega \mid X(\omega) = 1\} \in \mathcal{F} \quad \text{and} \quad \{\omega \in \Omega \mid X(\omega) = 0\} \in \mathcal{F}$$

In English, $\mathcal{F}(X)$ contains the subsets of Ω which start with a H, and the subsets which start with a T. $\mathcal{F}(X) = \{\emptyset, \Omega, \{HH, HT\}, \{TT, TH\}\}$.

Consider now a random variable Y defined in the same way but for the second throw of the coin. Now let $Z = X + Y$. Hence

$$Z(\omega) = \begin{cases} 2 & \text{if } \omega = HH \\ 1 & \text{if } \omega = HT \text{ or } \omega = TH \\ 0 & \text{if } \omega = TT \end{cases}$$

$\mathcal{F}(Z)$ must contain $\{HH\}$, $\{HT, TH\}$, and $\{TT\}$, representing the 3 different cases. It must contain all possible unions of these, as well as $\emptyset$, and Ω. This gives a total of 8 elements.

Most generally $\mathcal{F} = \mathcal{P}(\Omega)$, which has 16 elements in this case.

Definition 7 *A stochastic process may be defined in terms of a set of random variables, $\{X_t\}_{t \in T}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ which take values on $\mathbb{R}^n$.*

Observe that for fixed t, $\omega \mapsto X_t(\omega)$. However, fixing ω gives $t \mapsto X_t(\omega)$. This is called a path of X_t. Working in this way allows the behaviour of a particular ω in an experiment (say) to be modelled over time. In this way ω defines a function $\omega \colon T \to \mathbb{R}^n$ given by $t \mapsto X_t(\omega)$.

In this way, $\mathcal{F}$ now contains sets of functions which define subsets of $\mathbb{R}^n$.

(20.1.3) Expectation

Definition 8 *The expected value of a random variable X or any function g of X is defined as*

$$\begin{aligned} \mathbb{E}(X) &= \int_\Omega X(\omega)\, \mathrm{d}P(\omega) = \int_{\mathbb{R}^n} X\, \mathrm{d}\mu_X(X) \\ \mathbb{E}(g(X)) &= \int_\Omega g(X(\omega))\, \mathrm{d}P(\omega) = \int_{\mathbb{R}^n} g(X)\, \mathrm{d}\mu_X(X) \end{aligned}$$

Note it is possible to calculate expectation over a subset of Ω i.e. what is the expected value if certain elementary events are excluded? For $A \subset \Omega$ this gives $\mathbb{E}(X; A) = \int_A X\, \mathrm{d}P(X)$.

The linearity of the expectation operator follows from its definition in terms of integrals—the integration operator is linear. Clearly the definition of expectation is not in the form familiar to statistics. The variable of integration must be changed from $\mu_X(X)$ to the real variable X (working in $\mathbb{R}$).

$$\begin{aligned}
F_X(X) &= \int_{-\infty}^{x} f(x)\,\mathrm{d}x \\
\text{so } \frac{\mathrm{d}F_X}{\mathrm{d}x} &= f(x) \\
\text{by definition, } \mathbb{E}\,(X) &= \int_{\mathbb{R}^n} X\,\mathrm{d}\mu_X(X) \\
\text{now substituting, } &= \int_{-\infty}^{\infty} Xf(x)\,\mathrm{d}x \text{ but } X = X(\omega) = x \\
&= \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x
\end{aligned}$$

This calculation is a bit confusing as x and X are used for very different things. The expectation has a few important results.

Theorem 9 (Chebychev's Inequality) *With notation as used above, and any non-negative non-decreasing function f,*

$$P(X \geqslant a) \leqslant \frac{\mathbb{E}\,f(X)}{a}$$

Theorem 10 (Jensen's Inequality) *Let $c\colon G \to \mathbb{R}$ be a convex function on an open interval $G \subset \mathbb{R}$. Let X be a random variable of finite expectation and $P(X \in G) = 1$. Then*

$$\mathbb{E}\,c(X) \geqslant c(\mathbb{E}\,X)$$

A convex function has the property $c\left(\frac{x+y}{2}\right) \leqslant \frac{1}{2}f(x) + \frac{1}{2}f(y)$. The tangent line at any point lies wholly below the function. For example $f(x) = x^2$.

In considering stochastic processes sequences of random variables have been introduced. It is of interest, therefore, when and how these sequences are convergent. Let $\{X_n\}$ be a sequence of random variables

- $X_n \to X$ almost surely as $n \to \infty$ if $P(X_n \to X) = 1$.
- $\{X_n\}$ is mean square convergent to X if $\lim_{n\to\infty} |X_n - X| = 0$. From this it can be shown that $\mathbb{E}\,X_n \to \mathbb{E}\,X$ also.

Conditional Expectation

Mentioned after the definition of expectation was the possibility of finding expectation over only a subset of Ω. This is used in the definition of conditional expectation.

Definition 11 *Let $(\Omega, \mathcal{F}_0, P)$ be a probability space, let $\mathcal{F} \subset \mathcal{F}_0$ be a σ algebra, and let X be a random variable that is $\mathcal{F}_0$ measurable and of finite expectation. Define the random variable $Y = \mathbb{E}\,(X \mid \mathcal{F})$ such that*

1. *Y is $\mathcal{F}$ measurable*
2. *$\int_A X\,\mathrm{d}P = \int_A Y\,\mathrm{d}P\ \forall A \in \mathcal{F}$.*

By excluding some of $\mathcal{F}_0$ the information available for finding $\mathbb{E}\,X$ is reduced. Consider the example of throwing 2 coins and let X be the random variable defined by $X(H) = 1$ and $X(T) = 0$ for the first coin. Any events relating to the second coin can be happily thrown away, but notice that doing so does not reduce the portion of Ω being worked with—all 4 elementary events are still under consideration.

If now Y is defined in the same way as X but for the second coin. Using the same reduced even space (σ algebra) it is clear there is insufficient information to calculate $\mathbb{E}\,Y$. It is, however, possible to calculate $\mathbb{E}\,(Y \mid \mathcal{F})$ where $\mathcal{F}$ is the reduced event space.

Conditional expectation has all the same properties as normal expectation with the addition of the following.

1. $\mathbb{E}\,(\mathbb{E}\,(X \mid \mathcal{F})) = \mathbb{E}\,X$
2. If $\mathcal{F}_1 \subset \mathcal{F}$ then
   (a) $\mathbb{E}\,(\mathbb{E}\,(X \mid \mathcal{F}_1) \mid \mathcal{F}) = \mathbb{E}\,(X \mid \mathcal{F}_1)$
   (b) $\mathbb{E}\,(\mathbb{E}\,(X \mid \mathcal{F}) \mid \mathcal{F}_1) = \mathbb{E}\,(X \mid \mathcal{F}_1)$

   So the smaller σ algebra dictates the expectation.
3. If $X \in \mathcal{G} \subset \mathcal{F}$, X is $\mathcal{G}$ measurable, Y is a random variable of finite expectation, and $\mathbb{E}\,(XY)$ is finite then $\mathbb{E}\,(XY \mid \mathcal{G}) = X\,\mathbb{E}\,(Y \mid \mathcal{G})$.

This last property holds because X is $\mathcal{G}$ measurable. Y on the other hand may not be.

Conditional Probability

Attention is drawn to conditional probability, Bayes Law, etc. In the context of σ algebrae, for $B \in \mathcal{F}$ define

$$\mathcal{F}_B = B \cap \mathcal{F} = \{A \cap B \mid A \in \mathcal{F}\}$$

Observe that $\mathcal{F}_B$ is a set of sets of elementary elements. DeMorgan's laws readily show that $\mathcal{F}_B$ is itself a σ algebra. Hence define the measure

$$P^*(A \cap B) = \frac{P(A \cap B)}{P(B)}$$

So $(B, \mathcal{F}_B, P^*)$ is a probability space. Ω has been reduced to B because of the way $\mathcal{F}_B$ is defined. Alternatively simply define the conditional probability measure $P_B(A) = P(A \mid B)$. This yields the probability space $(\Omega, \mathcal{F}, P_B)$.

Consider a partition of Ω, $B_1, B_2, \ldots$. Now use the measure P_{B_i} to calculate the expectation of some random variable X, $*$

$$\mathbb{E}_{B_i}(X) = \int X \, \mathrm{d}P_{B_i}$$

(20.1.4) Martingales

Definition 12 *A filtration is a sequence of σ algebrae such that*

$$n \leqslant m \quad \Leftrightarrow \quad \mathcal{F}_n \subset \mathcal{F}_m$$

Considering the subscripts as time, more and more events are added as time increases. This relates to more information being available in the sense that the event "the second coin is a head" can be happily ignored until the second coin is thrown.

A natural filtration is one generated by a stochastic process X_t, so the σ algebrae will be $\mathcal{F}(X_t)$. Clearly for a natural filtration, X_t is always $\mathcal{F}_t$ measurable, but for other filtrations this is not necessarily the case.

Definition 13 *A family of random variables, X_t, is adapted to the filtration $\mathcal{F}_i$ if for each t, X_t is $\mathcal{F}_t$ measurable.*

Observe that if a random variable is adapted, then any function of that random variable which is itself a random variable is also adapted to the same filtration. This is because taking a function effects the value of the random variable, but not what elementary events produce which value. Measurability is therefore maintained.

Definition 14 *Let $(\Omega, \mathcal{F}, P)$ be a probability space, let $\mathcal{F}_i$ be a filtration in $\mathcal{F}$, and let X_t be a collection of random variables that are adapted to $\mathcal{F}_i$.*

1. *X_t is a martingale if $\mathbb{E}(X_t \mid \mathcal{F}_s) = X_s$ whenever $s \leqslant t$.*
2. *X_t is a submartingale if $\mathbb{E}(X_t \mid \mathcal{F}_s) \geqslant X_s$ whenever $s \leqslant t$.*
3. *X_t is a supermartingale if $\mathbb{E}(X_t \mid \mathcal{F}_s) \leqslant X_s$ whenever $s \leqslant t$.*

A martingale is an adapted collection of random variables with the property that the σ algebra $\mathcal{F}_t$ only contains enough information to calculate the expected value of X_t and earlier random variables. Essentially, this means that it isn't possible to tell the future in advance of it happening.

To show that something is a martingale three properties must be verified: adaption, that $\mathbb{E}|X_n|$ are all finite, and that one of the three martingale properties holds.

The practical situation is as follows. An experiment is performed and is done so sequentially. Each time a bit more of the experiment is performed more events are 'available'. However, $\mathcal{F}_t \subsetneq \mathcal{F}_{t+1}$ as the sample spaces are completely different. For example when throwing two coins, after the first throw the elementary events have only one letter—'H' or 'T'. After the second coin is thrown the elementary events have two letters.

Stopping Times

Definition 15 *Let $\mathcal{F}_t$ for $t \in I$ be a filtration in a set Ω. Define a function $T\colon \Omega \to I$ such that for $T(\omega) \leqslant t$, T is $\mathcal{F}_t$ measurable. Then T is called a stopping time.*

A stopping time is really a random variable: Depending on which $\omega \in \Omega$ occurs, it will have different values. If X_n is a martingale, then when t is the value of T obtained at a particular step in the sequence of events

$$X_{n \wedge T} = \begin{cases} X_n & \text{if } n < t \\ X_t & \text{if } n > t \end{cases}$$

Suppose a martingale represents the value of shares. A stopping time could be used to decide when to sell the shares and from this expected profit calculated. The question of optional stopping is therefore very important.

Theorem 16 *A stopped $\binom{super}{sub}$martingale is a $\binom{super}{sub}$martingale.*

Theorem 17 (Optional Stopping) *Let X_i be a supermartingale and T be a stopping time. T is integrable and $\mathbb{E}\, X_T \leqslant \mathbb{E}\, X_0$ if any one of the following hold*

1. *T is bounded.*
2. *X is bounded for every t and T is almost surely finite i.e. $Pr\{T_n \to T \text{ as } n \to \infty\} = 1$.*
3. *$\mathbb{E}\, T$ is finite and there exists K such that*

$$|X_s(\omega) - X_t(\omega)| \leqslant K \quad \textit{for all } s, t, \omega$$

If in fact X_i is a martingale, then the equality holds.

Example 18 *Let X_n be a martingale with respect to a filtration $\mathcal{F}_n$. The martingale is stopped when its value reaches $-a$ or b. Hence*

$$T = \min\{n \mid X_n = -a \text{ or } X_n = b\}$$

Now, say $p = Pr\{X_T = -a\}$.
Since X_n is a martingale, $\mathbb{E}\, X_t = 0$. Hence by the optional stopping theorem $\mathbb{E}\, X_T = 0$. However, by the definition of the random variable T this gives

$$0 = -ap + b(1-p) \quad \Rightarrow \quad p = \frac{b}{a+b}$$

(20.1.5) Applications In Economics

Of the plethora of economic theories, only some are compatible with the maths covered thus far. It is clear, for example, that the use of martingales will be relevant to theories proposing the current information about securities* is reflected in its present market price.

Alternative 'fundamentalist' theories suggest that the value of a security is equal to "the value of the discounted cashflow that security generates". In this model, money is made by trading securities whose price is different from its fundamental price, which an analyst can calculate.

Stock Pricing

At some time t let p_t be the price of stock, d_t be the dividends paid, $\mathcal{F}_t$ be "the information available", and r the the interest rate available on a safe investment†.

Say an investor wishes to invest an amount X_0 at time zero. At time t the investment would be worth $(1+r)^t X_0$ if left safe in the bank. This should be taken account for in the calculations—this is the concept of a discounted cashflow. In other situations it is common to discount for inflation.

Assume at present the time is $t = 0$. The value of an investment can be modelled as a random variable, and its present worth can be found by discounting against the safe investment.

At a time t (even though at present the time is $t = 0$) the investment will be worth $p_t + d_t$. The value at t should be the expected value of the investment at the next time, but discounted against the possibility of making a safe investment instead. Hence

$$p_t = (1+r)^{-1}\, \mathbb{E}\,(p_{t+1} + d_{t+1} \mid \mathcal{F}_t)$$

Now let h_t be the number of shares held, v_t be the value of the fund at time t, and assume that dividends are re-invested. Hence

$$v_t = (1+r)^t p_t h_t$$

Imagine now what happens at the next time, $t+1$. At the start of the period h_t shares are held. They are worth p_{t+1} and pay a dividend of d_{t+1}. Once the dividend is paid more shares are bought so that h_{t+1} shares are held. Hence

$$p_{t+1}h_{t+1} = (p_{t+1} + d_{t+1})h_t$$

At present ($t = 0$) interest lies in the value of the fund, which is really just a question of finding $\mathbb{E}\, v_t$.

*The umbrella term 'securities' covers shares, options, bonds, etc. Each of these are defined in due course.

†r represents the return achievable from a safe investment such as a bank account. If the risky investment in shares cannot do as well as a safe investment, then it is not a good investment.

However,

$$\begin{aligned}
\mathbb{E}\,(v_{t+1} \mid \mathcal{F}_t) &= \mathbb{E}\,\left((1+r)^{-t-1}h_{t+1}p_{t+1} \mid \mathcal{F}_t\right)\\
&= \mathbb{E}\,\left((1+r)^{-t-1}(p_{t+1}+d_{t+1})h_t \mid \mathcal{F}_t\right)\\
&= (1+r)^{-t}h_t\left((1+r)^{-1}\,\mathbb{E}\,\left(p_{t+1}+d_{t+1} \mid \mathcal{F}_t\right)\right)\\
&= (1+r)^{-t}h_t p_t\\
&= v_t
\end{aligned}$$

Hence v_t is a martingale. Note that the actual stock prices are not martingales, this will be shown in the next section. From this point it is possible to find an expression for p_t but this is a fundamentalist concept.

Valuation Of A Portfolio

A portfolio is simply a collection of shares and bonds. Bonds are safe investments. Since the superscript is used to denote time, the subscript is now used to index, so let X_t^1 and X_t^2 be random variables (stochastic processes) representing the price of a bond and of a stock, respectively. Let h_t^i represent the number of stocks and bonds held, so h_t^i is $\mathcal{F}_{t-1}$ measurable i.e. it is predictable. The value of the portfolio is simply

$$v_t = h_t^1 X_t^1 + h_t^2 X_t^2$$

Define now the gains process, which measures how much is made between consecutive times (considering time as a discrete quantity).

$$G_t = v_{t+1} - v_t = h_t^1(X_{t+1}^1 - X_t^1) + h_t^2(X_{t+1}^1 - X_t^2) \quad \text{so} \quad v_t = v_0 + \sum_{j=0}^{t} G_j$$

Note the term G_t^i may be used to denote the gains for the investment i, and $X_{t+1}^i - X_t^i$ may be written ΔX_{t+1}^i.
*

This is all very well, but how does the value of the portfolio compare to the value of a safe investment? This consideration is incorporated into the model in the usual way—discounting. Say

$$\overline{X}_t^i = \frac{X_t^i}{N_t}$$

The normalising quantity N_t often represents the value of a safe investment, $N_t = X_t^0 = (1+r)^t X_0^0$ say, where X^0 is the value of the safe investment. This now gives

$$\overline{v}_t = v_0 + \sum_{i=1}^{n} \overline{G}_i$$

where $\overline{G}_i = \frac{G_i}{N_i}$.

Definition 19 *Let a be a self financing portfolio and $V_t(a)$ represent the value of a at time t. If $V_0(a) = 0$ and $V_t(a) \geqslant 0$ and $\mathbb{E}\,V_T(a) > 0$ then a is called an arbitrage.*

An arbitrage is a portfolio that makes something from nothing. It is required that a portfolio is not an arbitrage, and furthermore that its value has a lower bound—there is a limit to indebtedness. It an be shown the non existence of arbitrage is equivalent to the existence of an equivalent martingale measure. This is a measure under which the normalised prices become martingales.

Model For A Market

In modelling a market, consider the value of a contingent claim[‡]. Modelling this, a contingent claim is a non-negative random variable F which is $\mathcal{F}$ measurable and represents the payment of £$F(\omega)$ at time T if ω occurs. The claim can be sold at any time $t \leqslant T$ and the question is its worth at any such time.

Consider for example F as a call option—the option to buy shares at the pre-determined price k. Let X_t be the price of the shares at time t.

- If $X_T > k$ then the option has value $k - X_T$ and would be exercised.
- If $X_T < k$ then the option has value 0, and would not be exercised.

Hence $F = \max(0, X_T - k)$, a random function of ω. The claim has two values—one to the buyer and one to the seller.

For the seller let the value at $t = 0$ be Y. The equation

$$Y + \sum_{i=0}^{T} \theta_i \Delta X_i \geqslant F(\omega)$$

must be satisfied. While the seller would like Y to be as large as possible, the minimum price acceptable is the minimum value of Y for which this equation holds. The equation says the money received for the claim plus the profit made by the seller's portfolio must exceed the payment due in honouring the option when exercised.

For the buyer the equation

$$-Y + \sum_{i=0}^{T} \theta_i \Delta X_i + F(\omega) \geqslant 0$$

must be satisfied. Although the buyer clearly wants Y to be as small as possible, interest lies in the largest value of Y that satisfies this equation. The equation says that whatever is paid for the option plus whatever the buyer makes from his portfolio (a different one to the seller's) plus the value of the option, must be positive.

These two situations must be solved simultaneously if the claim can be sold. Assume no arbitrage and assume completeness so that for any given $F(\omega)$ there exists θ such that

$$F(\omega) = Y + \sum_{i=0}^{T} \theta_i \Delta X_i$$

The seller wishes to solve this for Y. Now, there exists a measure Q which makes $\overline{X}_i$ a martingale so for a self financing process

$$\begin{aligned} Y + \sum_{i=0}^{T} \theta_i \Delta \overline{X}_i &= \frac{F(\omega)}{X_T^0} \\ \mathop{\mathbb{E}}_{Q}\left(Y + \sum_{i=0}^{T} \theta_i \Delta \overline{X}_i\right) &= \mathop{\mathbb{E}}_{Q}\left(\frac{F(\omega)}{X_T^0}\right) \\ Y &= \mathop{\mathbb{E}}_{Q}\left(\frac{F(\omega)}{X_T^0}\right) \qquad \text{since } \overline{X}_i \text{ is a martingale} \end{aligned}$$

[‡]A contingent claim is an asset with random price that depends on the value of some other asset—the 'underlying'. These include options.

So a value for Y can be found. To do this the measure Q has to be determined. Let X_t^0 the the value of a bond and let the stock price move with increments

$$R_t = \frac{X_{t+1}}{X_t} = \begin{cases} a & \text{with probability } p \\ b & \text{with probability } 1-p \end{cases}$$

so the Rs are independently and identically distributed random variables. This is a binomial model for the market. The value of p determines the measure Q and it is found to make the discounted price process a martingale.

$$\overline{X}_t = \frac{X_t}{X_t^0} = \frac{X_t}{(1+r)^t X_0^0}$$

The value of p is found so that the martingale property holds, so

$$\begin{aligned}
\overline{X}_{t-1} &= \mathop{\mathbb{E}}_{Q}\left(\overline{X}_t \mid \mathcal{F}_{t-1}\right) \\
\frac{X_{t-1}}{(1+r)^{t-1}X_0^0} &= \mathop{\mathbb{E}}_{Q}\left(\frac{X_{t-1}R_{t-1}}{(1+r)^t X_0^0} \mid \mathcal{F}_{t-1}\right) \\
1+r &= \mathop{\mathbb{E}}_{Q}\left(R_{t-1} \mid \mathcal{F}_{t-1}\right) \\
&= \mathop{\mathbb{E}}_{Q} R_{t-1} \quad \text{since } R_{t-1} \text{ is not } \mathcal{F}_{t-1} \text{ measurable} \\
&= ap + b(1-p) \\
p &= \frac{1+r-b}{a-b}
\end{aligned}$$

This value of p defines Q, and hence an appropriate measure is found.

(20.1.6) Brownian Motion

Brownian motion is a familiar concept to physicists as it can be used to model diffusion. Brownian motion is useful in mathematical finance as a way to add randomness to an otherwise smooth curve—the underlying trend.

Definition 20 *An m dimensional Brownian motion is a stochastic process $\mathbf{B}_t = (B_t^1, B_t^2, \ldots, B_t^m)$ taking values in $\mathbb{R}^m$ such that*

1. *If $t_0 < t_1 < \cdots < t_n$ then the random variables $B_{t_0}, B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent.*
2. *Where $s \geqslant t$ the increments are Gaussianly distributed, so that where $A \subset \mathbb{R}^m$,*

$$Pr\{B_{s+t} - B_t \in A\} = \int_A \frac{1}{\sqrt{2\pi t}} e^{\frac{-|x|^2}{2t}} \, \mathrm{d}\mathbf{x}$$

3. *The paths are continuous with probability 1. i.e. the function mapping t to $\mathbf{B}_t$ is continuous.*

The probability density function in this definition is very important. Note that in the definition t represents time difference and x represents change in value. Therefore in the one dimensional case the following variations occur.

$$\begin{aligned}
&\text{When } B_0 = 0 \qquad \mathbb{E}\left(f(B_t)\right) = \int_{-\infty}^{\infty} f(z)\frac{1}{\sqrt{2\pi t}}e^{\frac{-|z|^2}{2t}}\ \mathrm{d}z \\
&\text{When } B_0 = x \qquad \mathop{\mathbb{E}}_{x}\left(f(B_t)\right) = \int_{-\infty}^{\infty} f(z)\frac{1}{\sqrt{2\pi t}}e^{\frac{-|x-z|^2}{2t}}\ \mathrm{d}z \\
&\text{When } B_0 = 0 \quad \mathbb{E}\left(f(B_t - B_s)\right) = \int_{-\infty}^{\infty} f(z)\frac{1}{\sqrt{2\pi(t-s)}}e^{\frac{-|z|^2}{2(t-s)}}\ \mathrm{d}z
\end{aligned}$$

In particular it is worth noting that $(B_t - B_s) + B_s$ has the same distribution as $B_{t-s} + B_s$, as this will be of much use.

Stock Pricing

A common model for the price of a stock is

$$\frac{\Delta s}{s} = rt + \sigma \Delta B \qquad \text{giving} \qquad s_t = s_0 \exp\left(at + \sigma B_t\right) \quad \text{where } a = r - \frac{\sigma^2}{2} \tag{21}$$

This can be shown to have expected value $s_0 e^{rt}$ and to be a martingale under an appropriate measure. In fact Brownian Motion itself is a martingale.

$$\begin{aligned}
\mathbb{E}\left(B_t \mid \mathcal{F}_s\right) &= \mathbb{E}\left(B_t - B_s + B_s \mid \mathcal{F}_s\right) \quad t > s \\
&= B_s + \mathbb{E}\left(B_t - B_s \mid \mathcal{F}_s\right) \\
&= B_s + \int_{-\infty}^{\infty} \frac{y}{\sqrt{2\pi(t-s)}} \exp\left(\frac{-y^2}{2(t-s)}\right)\ \mathrm{d}y \\
&= B_s
\end{aligned}$$

Note that time is continuous, not discrete. The conditional expectations of Brownian Motions have a special property. $\mathcal{F}_s$ provides no more information about B_t than B_s does. Hence $\mathbb{E}\left(B_t \mid \mathcal{F}_s\right) = \mathbb{E}\left(B_t \mid B_s\right)$. Furthermore, because B_{t-s} is independent,

$$\mathop{\mathbb{E}}_{0}\left(B_t \mid \mathcal{F}_s\right) = \mathop{\mathbb{E}}_{0}\left(B_t - B_s + B_s \mid \mathcal{F}_s\right) = \mathop{\mathbb{E}}_{B_s} B_{t-s}$$

Also

$$\mathop{\mathbb{E}}_{0}\left(f(B_t) \mid \mathcal{F}_s\right) = \mathop{\mathbb{E}}_{B_s}\left(f(B_{t-s})\right) = \int_{-\infty}^{\infty} \frac{f(y)}{\sqrt{2\pi(t-s)}} \exp\left(\frac{-(y-B_s)^2}{2(t-s)}\right)\ \mathrm{d}y$$

where B_s in the integrand is just a number.

Theorem 22 (Markov Property For Brownian Motion) *Let Y be a bounded measurable function and let θ_s be the shift operator. Then $\mathbb{E}\left(Y \circ \theta_s \mid \mathcal{F}_s\right) = \mathop{\mathbb{E}}_{B_s} Y$.*

This means at any particular time s the Brownian Motion after s is dependent only on the value B_s. Brownian Motions "forget their past". This theorem has uses in finding things such as $\mathbb{E}\left(s_{s+t} \mid \mathcal{F}_s\right)$.

Further to the Markov Property, Brownian Motion also has the Strong Markov Property. This simply means the above holds when s is a random time i.e. determined by a random variable—as stopping time.

(20.2) Itô Calculus

How far has a particle undergoing Brownian Motion travelled? This is an obvious question, and its solution is clearly an integral of some form.

(20.2.1) Itô Integration

Consider a partition of the time interval $[a, b]$, say $a = t_0 < t_1 < \cdots < t_n = b$. Now if $f(t, \omega)$ is a function it would be usual to approximate the 'area under it' with a Riemann sum with summands $f(t_i, \omega)(t_{i+1} - t_i)$ but this does not work as there are many values ω could take. Instead take

$$\int_0^t f(t, \omega)\ \mathrm{d}B_t \stackrel{\text{def}}{=} \lim_{n\to\infty} \sum_{i=0}^{n-1} f(t_i, \omega)(B_{t_{i+1}} - B_{t_i})$$

This is not convergent for any fixed ω, and the sum is only convergent when

$$\mathbb{E}\left(\left|f(t_i, \omega)(B_{t_{i+1}} - B_{t_i}) - I\right|^2\right) \to 0 \quad \text{as} \quad n \to \infty$$

Since in f, ω represents dependence on Brownian Motion, the concept of the composite map $t \to B_t \to f$ justifies the presence of the term $(B_{t_{i+1}} - B_{t_i})$.

The Itô integral has the following important properties

1. linearity
2. $\mathbb{E}\left(\int_a^b f\ \mathrm{d}B_t\right) = 0$
3. $\mathbb{E}\left(\left(\int_a^b f\ \mathrm{d}B_t\right)^2\right) = \mathbb{E}\left(\int_a^n f^2\ \mathrm{d}B_t\right)$ this is called the Itô isometry.
4. $M_t = \int_0^a f\ \mathrm{d}B_s$ is a martingale whenever f is bounded.

Itô Processes

Definition 23 *An Itô process is a stochastic process X_t of the form*

$$X_t = X_0 + \int_0^t u(s, \omega)\ \mathrm{d}s + \int_0^t v(s, \omega)\ \mathrm{d}B_s$$

where u and v have the properties $\int_0^t u\ \mathrm{d}s$ and $\int_0^t v^2\ \mathrm{d}s$ are finite. The process may be written in differential form,

$$\mathrm{d}X_t = u\ \mathrm{d}t + v\ \mathrm{d}B_t$$

The function u is the drift of the stochastic process, and v is the diffusion. These give an underlying trend with randomness added. It can be shown that an Itô process is a martingale only if its drift is zero.

Recall the value of a stock in a portfolio may be expressed in discrete time as

$$v_t = \sum_{i=1}^{t} \theta_t(X_{i+1} - X_i)$$

As increments of time between 0 and t decreases in size, this approaches the continuous case for which it is evident that

$$v_t = \sum_{i=1}^{t} \theta_t(X_{i+1} - X_i) \quad \to \quad \int_0^t \theta_t\ \mathrm{d}X_t$$

The Itô Formula

Calculating an Itô integral is quite different from calculating a normal integral. Take for example the case $f(t,\omega) = B_t$. Then

$$\begin{aligned}\int_0^t B_t \, \mathrm{d}B_t &= \lim_{n\to\infty} \sum_{i=1}^{n-1} B_{t_i} \left(B_{t_{i+1}} - B_{t_i}\right) \\ &= \lim_{n\to\infty} \sum_{i=1}^{n-1} \frac{1}{2}\left(B^2_{t_{i+1}} - B^2_{t_i}\right) - \frac{1}{2}\left(B_{t_{i+1}} - B_{t_i}\right)^2 \\ &= \lim_{n\to\infty} \left(\frac{1}{2}\left(B^2_{t_{n-1+1}} - B^2_{t_0}\right) - \sum_{i=1}^{n-1} \left(B_{t_{i+1}} - B_{t_i}\right)^2 \right)\end{aligned}$$

Now, the remaining sum is convergent to t. Also, $B_0 = 0$ and $t_n = t$ hence

$$\begin{aligned} &= \frac{1}{2} B_t^2 - \frac{1}{2} t \\ \text{So } B_t^2 &= t + 2\int_0^t B_t \, \mathrm{d}B_t \end{aligned}$$

This re-arrangement shows that B_t^2 is an Itô process with drift 1 and diffusion $2B_t$. However, B_t is itself an Itô process (it is a Brownian motion and so is a martingale) which has drift 0 and diffusion 1. Hence seek a way to find Itô processes from non-linear transformations of Itô diffusions—in the above case $B_t \mapsto B_t^2$.

Theorem 24 (The Itô Formula) *Let $f(t,x)$ be continuously differentiable in t and continuously twice differentiable in x i.e. $f(t,x) \in C^{(1,2)}$ and let X be an Itô process. Define the new stochastic process $Y_t = f(t, X_t)$, then Y_t is the Itô process*

$$Y_t = Y_0 + \int_0^t \frac{\partial f}{\partial t} + \frac{1}{2} v^2 \frac{\partial^2 f}{\partial x^2} + u \frac{\partial f}{\partial x} \, \mathrm{d}t + \int_0^t \frac{\partial f}{\partial x} \, \mathrm{d}B_t$$

or more generally

$$\mathrm{d}y = \frac{\partial y}{\partial t} \, \mathrm{d}t + \sum_{i=1}^{n} \frac{\partial y}{\partial x_i} \, \mathrm{d}x_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 y}{\partial x_i \partial x_j} \, \mathrm{d}x_i \, \mathrm{d}x_j$$

It is often preferable to work with the differential form where the identities

$$\mathrm{d}t \, \mathrm{d}t = \mathrm{d}t \, \mathrm{d}B_t = 0 \qquad \mathrm{d}B_t \, \mathrm{d}B_t = \mathrm{d}t \qquad \mathrm{d}x_i \, \mathrm{d}x_j = 0 \text{ when } i \neq j$$

may be used.

Example 25 *Define the Itô process $M_t = e^{at+bB_t}$. Find conditions on a and b such that M_t is a martingale.*

Proof. Solution Consider the function $f(t,x) = e^{at+bx}$ which is in $C^{(1,2)}$. Set $x = B_t$ which is an Itô process with drift 0 and diffusion 1. Hence in the Itô formula,

$$M_t = \left(a + \frac{b}{2}\right) \int_0^t e^{at+bB_t} \, \mathrm{d}t + b \int_0^t B_t \, \mathrm{d}t$$

Now, an Itô process is a martingale when it has no drift, so for M_t to be a martingale, $a = \frac{-b^2}{2}$. □

The Itô formula can be extended into multiple dimensions. So when $Y_t = f(t, X_t^1, X_t^2, \ldots, X_t^m)$

$$\mathrm{d}Y_t = \frac{\partial f}{\partial t} \, \mathrm{d}t + \sum_{i=1}^{m} \frac{\partial f}{\partial x_i} \, \mathrm{d}X_t^i + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_i \partial y_j} \, \mathrm{d}X_t^i \, \mathrm{d}X_t^j$$

where $\mathrm{d}X_t^i \, \mathrm{d}X_t^j = 0$ whenever $i \neq j$.

(20.2.2) Stochastic Differential Equations

Definition 26 *A stochastic differential equation is an equation of the form*

$$X_t = X_0 + \int_0^t v(t,\omega)\ \mathrm{d}t + \int_0^t u(t,\omega)\ \mathrm{d}B_t$$

One such equation is a model for the price of an asset,

$$\begin{aligned} \mathrm{d}S_t &= \mu S_t\ \mathrm{d}t + \sigma S_t\ \mathrm{d}B_t \qquad (27)\\ \text{i.e. } S_t - S_0 &= \int_0^t \mu S_t\ \mathrm{d}t + \int_0^t \sigma S_t\ \mathrm{d}B_t \end{aligned}$$

Dividing through the differential form by S_t the left would integrate to $\ln S_t$ in normal calculus. Consider therefore the function $f(t,x) = \ln x$. Put $x = S_t$ and use the Itô formula

$$\begin{aligned} \mathrm{d}f &= \frac{\partial f}{\partial t}\ \mathrm{d}t + \frac{\partial f}{\partial x}\ \mathrm{d}x + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\ (\ \mathrm{d}x)^2 \\ &= \frac{1}{S_t}\ \mathrm{d}x + \frac{-1}{2S_t^2}\ (\ \mathrm{d}x)^2 \\ &= \frac{1}{S_t}\ (\mu S_t\ \mathrm{d}t + \sigma S_t\ \mathrm{d}B_t) + \frac{-1}{2S_t^2}\ (\mu S_t\ \mathrm{d}t + \sigma S_t\ \mathrm{d}B_t)^2 \\ &= \frac{1}{S_t}\ (\mu S_t\ \mathrm{d}t + \sigma S_t\ \mathrm{d}B_t) + \frac{-1}{2S_t^2}\sigma^2 S_t^2\ \mathrm{d}t \\ &= \left(\mu - \frac{\sigma^2}{2}\right)\ \mathrm{d}t + \sigma\ \mathrm{d}B_t \\ \ln S_t - \ln S_0 &= \left(\mu - \frac{\sigma^2}{2}\right)\int_0^t\ \mathrm{d}t + \sigma\int_0^t\ \mathrm{d}B_t \\ &= \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t \\ S_t &= S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right) \end{aligned}$$

This is equation (21).

Generator Of A Diffusion

Associated with each Itô process is a second order linear partial differential operator A. This is defined as

$$A\,f(x) \stackrel{\text{def}}{=} \lim_{t\to\infty} \frac{\mathbb{E}\,f(X_t) - f(x)}{t}$$

This is clearly quite useless so alternatives are sought. Let $\mathbf{X}_t$ be an Itô process in m dimensions with initial value $\mathbf{x}$.

Theorem 28 *Let* $\mathbf{u}$ *be an* n *dimensional vector,* $\mathbf{B}_t$ *be an* m *dimensional Brownian Motion. If*

$$\mathbf{X}_t = \mathbf{x} + \int_0^t \mathbf{b}(s,\omega)\ \mathrm{d}t + \int_0^t \boldsymbol{œ}(s,\omega)\ \mathrm{d}\mathbf{B}_s(\omega)$$

then

$$A = \sum_{i=1}^n b_i \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n (\boldsymbol{œ}\boldsymbol{œ}^T)_{ij}\frac{\partial^2}{\partial x_i \partial x_j}$$

The duality between Itô processes and partial differential equations allows differential equations to be solved by taking expectation over a stochastic process.

Example 29 *Find the solution to the diffusion equation as the expected value of a stochastic process, given the Cauchy boundary conditions $u(0, x) = g(x)$.*

Proof. Solution The solution to the given equation is known to be

$$u(x,t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp\left(\frac{-(x-y)^2}{2t}\right) g(y)\, \mathrm{d}y$$

Already the integrand seems familiar as the probability density function of some kind of Brownian motion. By inspection this is clearly equal to

$$\mathop{\mathbb{E}}_x \left(g(B_t)\right)$$

Let B'_t be the same Brownian motion but starting at 0. Then

$$B_t = x + B'_t \quad \text{so} \quad B'_t = B_t - B_0 = \int_0^t 0\, \mathrm{d}t + \int_0^t \mathrm{d}B'_t$$

Hence B'_t is an Itô process with drift 0 and diffusion 1. The generator for this is

$$A = \frac{1}{2}\frac{\partial^2}{\partial x^2}$$

□

The situation may be approached from a different angle. Consider a stochastic process given by the function $f(T-t, B_t)$. Applying the Itô formula,

$$f(T-t, B_t) = f(T, B_0) + \int_0^T -\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\, \mathrm{d}t + \int_0^T \frac{\partial f}{\partial x}\, \mathrm{d}B_t$$

The drift vanishes when

$$\frac{\partial f}{\partial t} = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}$$

So when f obeys the diffusion equation, replacing x with B_t gives an Itô diffusion. With Cauchy boundary conditions this gives

$$g(x) = f(T, B_0) + \int_0^T \frac{\partial f}{\partial x}\, \mathrm{d}B_t$$

Now take expectation. The integral represents a martingale so has expected value 0. Put $B_0 = x$ which not random—hence the expectation can be dropped. Since $B_0 = x$ the expectation is taken beginning at x. T is replaced by t.

$$\begin{aligned}
\mathop{\mathbb{E}}_x g(B_t) &= \mathbb{E}\, f(T, B_0) = f(t, x) \\
f(t,x) = \mathop{\mathbb{E}}_x g(B_t) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} \exp\left(\frac{-(x-y)^2}{2t}\right)\, \mathrm{d}y
\end{aligned}$$

This is precisely the same result as Example 29. This is an example of a more general result.

Theorem 30 (Feynman-Kac) *Suppose $f \in C^2(\mathbb{R}^n)$ and c is continuously differentiable. Let X_t be a Itô process with generator A. The the solution to the partial differential equation*

$$\frac{\partial u}{\partial x} = Au + cu$$

where $u(0,x) = f(x)$ *is*

$$u(t,x) = \mathop{\mathbb{E}}_{x}\left(\exp\left(\int_0^t c(X_s)\,\mathrm{d}s\right) f(X_t)\right)$$

This is an easy way to solve partial differential equations provided, of course, a suitable Itô process can be found (although one always exists). For example, the differential equation

$$\frac{\partial u}{\partial t} = b\frac{\partial u}{\partial x} + \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + cu$$

clearly has generator and associated stochastic process

$$A = b\frac{\partial}{\partial x} + \frac{1}{2}\frac{\partial^2}{\partial x^2} \quad \Leftrightarrow \quad X_t = X_0 + \int_0^t b\,\mathrm{d}t + \int_0^t \mathrm{d}B_t$$

By Feynman-Kac the solution is

$$u(t,x) = \mathop{\mathbb{E}}_{x}\left(g(X_t)\exp\left(\int_0^t c\,\mathrm{d}t\right)\right)$$

Now, assume b and c are constant, so $X_t = x + bt + B_t$ where now B_t starts at 0. Using this,

$$\begin{aligned}
&= \mathbb{E}\left(g(x+bt+B_t)e^{ct}\right) \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t}} e^{\frac{-y^2}{t}} g(x+bt+y)e^{ct}\,\mathrm{d}y \quad \text{now put } z = x+bt+y \\
&= \int_{-\infty}^{\infty} \frac{g(y)e^{ct}}{\sqrt{2\pi t}} \exp\left(\frac{-(z-x-bt)^2}{2t}\right)\,\mathrm{d}z
\end{aligned}$$

Martingale Properties

It is very convenient to work with stochastic processes that are martingales. But, for example, an Itô process is only a martingale when it has no drift. The situation can be remedied by changing probability measure.

Definition 31 *Let P and Q be probability measures defined on a σ algebra $\mathcal{F}$. P and Q are equivalent—written $P \sim Q$—if for all $A \in \mathcal{F}$ $P(A) = 0 \Leftrightarrow Q(A) = 0$.*

Consider a Brownian motion: it is expected to fluctuate about some central position and not to drift. In order to make an Itô process into a martingale the paths which drift must be given low probability and the paths which do not must be given high probability, all under an equivalent measure.

The process $\overline{B}_t = bt + B_t$ is a Brownian motion with drift. Its probability density function is deduced as follows.

$$\begin{aligned}
\Pr \overline{B} < x &= \Pr bt + B_t < x = \Pr B_t < x - bt \\
&= \int_{-\infty}^{x-bt} \frac{1}{\sqrt{2\pi t}} \exp\left(\frac{-y^2}{2t}\right)\,\mathrm{d}y \\
&= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi t}} \exp\left(\frac{-(z-bt)^2}{2t}\right)\,\mathrm{d}z
\end{aligned}$$

So the probability density function is this last integrand. Finding now the expected value for a probability measure P,

$$\mathop{\mathbb{E}}_{P} \overline{B}_t = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi t}} \exp\left(\frac{-(x-bt)^2}{2t}\right) \, \mathrm{d}x = bt$$
$$\mathop{\mathbb{E}}_{P} \overline{B}_t^2 = \int_{-\infty}^{\infty} \frac{x^2}{\sqrt{2\pi t}} \exp\left(\frac{-(x-bt)^2}{2t}\right) \, \mathrm{d}x = t + (bt)^2$$

Clearly $\overline{B}_t$ is not a martingale under P. For $\overline{B}_t$ to be a martingale a measure Q must be found such that

$$\mathop{\mathbb{E}}_{Q} \overline{B}_t = \int_{-\infty}^{\infty} \frac{x}{\sqrt{2\pi t}} \exp\left(\frac{-(x-bt)^2}{2t}\right) \, \mathrm{d}x = 0$$

With a little thought it is clear this can be achieved by using the equivalent probability measure which has

$$\frac{\mathrm{d}Q}{\mathrm{d}P} = \exp\left(\frac{-y^2}{2t}\right) \exp\left(\frac{(y-bt)^2}{2t}\right) = \exp\left(\frac{1}{2}b^2 t - by\right)$$

This factor has the effect of changing the pdf for $\overline{B}_t$ to that of B_t. The notation of differential fractions is used since

$$\mathop{\mathbb{E}}_{Q} f(\overline{B}_t) = \int_{\Omega} f(y) \, \mathrm{d}Q = \int_{\Omega} f(y) \frac{\mathrm{d}Q}{\mathrm{d}P} \, \mathrm{d}P = \mathop{\mathbb{E}}_{P} \left(f(\overline{B}_t) \frac{\mathrm{d}Q}{\mathrm{d}P} \right)$$

It should be noted that $\frac{\mathrm{d}Q}{\mathrm{d}P}$ is not a 'proper' derivative, indeed it is often denoted ξ.

Theorem 32 (Girsanov) *Let ` $= (\theta_1, \theta_2, \ldots, \theta_n)$ be a vector of square integrable stochastic processes, and B_t be a Brownian motion under probability measure P. Then*

$$\xi_t^\theta = \exp\left(-\int_0^t \theta_s \, \mathrm{d}B_s - \frac{1}{2}\int_0^t \theta_s^2 \, \mathrm{d}s\right)$$

then this is a martingale under P. Define the equivalent measure Q^θ by

$$\frac{\mathrm{d}Q^\theta}{\mathrm{d}P} = \xi_t^\theta$$

then the process

$$B_t^\theta = B_t + \int_0^t \theta_s \, \mathrm{d}s$$

is a standard Brownian motion and is a martingale under Q^θ.

Theorem 33 *Let X be the Itô process*

$$X_t = x + \int_0^t \mu_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}B_s$$

and let ν be an integrable process and θ be be a square integrable process such that $\sigma_t \theta_t = \mu_t - \nu_t$. If ξ^θ is a martingale under P then

$$X_t = x + \int_0^t \nu_s \, \mathrm{d}s + \int_0^t \sigma_s \, \mathrm{d}B_s^\theta$$

where B_t^θ is a standard Brownian motion under Q^θ and is given by

$$B_t^\theta = B_t + \int_0^t \theta_s \, \mathrm{d}s$$

This theorem allows the drift of a stochastic process to be changed, and in particular for drift to be eliminated. To eliminate drift $\sigma\theta = \mu$, so if no solution for θ can be found then the drift cannot be eliminated.

(20.3) Applications In Finance

Definition 34 *A market is an $\mathcal{F}_t^{n+1}$ adapted $n+1$ dimensional Itô process with*

$$\mathrm{d}X_0 = r(t,\omega)X_0 \,\mathrm{d}t \quad \textit{(a bond)}$$
$$\mathrm{d}X_i(t) = \mu_i(t,\omega)\,\mathrm{d}t + \sum_{j=0}^{m}\sigma_{ij}(t,\omega)\,\mathrm{d}B_j(t)$$

(20.3.1) Black-Scholes Economic Model

In the Black-Scholes economy μ, σ, and r are constant.

Definition 35 *A portfolio is a collection of assets, say*

$$`(t) = (\theta_1(t), \theta_2(t), \ldots, \theta_n(t))$$

the present value of which is

$$v_t = \sum_{i=1}^{n}\theta_i(t)X_i(t) \tag{36}$$

Separately to this, a portfolio is said to be self financing if no external finds are contributed or taken out of the portfolio. The value of the portfolio is therefore equal to the initial value plus the sum of the gains, i.e.

$$v_t = v_0 + \int_0^t \sum_{i=1}^{n}\theta_i(t)\,\mathrm{d}X_i(t)$$
$$\mathrm{d}v_t = \sum_{i=1}^{n}\theta_i(t)\,\mathrm{d}X_i(t) \tag{37}$$

For equation (36) use Itô's formula, taking both θ and X to be stochastic processes. Hence

$$\mathrm{d}v_t = \sum_{i=1}^{n}\theta_i(t)\,\mathrm{d}X_i(t) + \sum_{i=1}^{n}X_i(t)\,\mathrm{d}\theta_i(t) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}0$$

Hence using equation (37) it is evident that for a self financing portfolio

$$\sum_{i=1}^{n}X_i(t)\,\mathrm{d}\theta_i(t) = 0$$

This is quite obvious really, as it expresses the inability to change the amounts of securities held independently of eachother i.e. by external influence. The next step in the process is to discount against a safe investment with price process $X_0(t)$, observe that $\overline{X}_0(t) = 1$.

Lemma 38 *The discounted price process of a normalised portfolio is given by*

$$\overline{v}_t = \sum_{i=0}^{n}\theta_i(t)\overline{X}_i(t)$$

Proof. First of all, use Itô's formula on $\overline{X}_i(t) = \frac{X_i(t)}{X_0 t}$ to give

$$\mathrm{d}X_0 = r\,\mathrm{d}t$$
$$\mathrm{d}\overline{X_i}(t) = \frac{1}{X_0(t)}\,\mathrm{d}X_i(t) - \frac{X_i(t)}{(X_0(t))^2}$$

(This is easily seen by working through the example $\mathrm{d}X_0 = r\,\mathrm{d}t$ with $f(X_0, X_1) = \frac{X_1}{X_0}$.) In the Black-Scholes model

$$\begin{aligned}\mathrm{d}X_0 &= \mu_0\ \mathrm{d}t\\ \mathrm{d}X_i &= \mu_i\ \mathrm{d}t + \sum_j \sigma_{ij}\ \mathrm{d}B_j(t)\end{aligned}$$

which gives rise to the appropriate cancellations. Now doing the same for $\overline{v}_t = \frac{v_t}{X_0 t}$ gives

$$\begin{aligned}\mathrm{d}\overline{v}_i &= \frac{1}{x_0(t)}\ \mathrm{d}v_t - \frac{v_t}{(X_0(t))^2}\ \mathrm{d}X_0(t)\\ &= \sum_{i=0}^{n} \frac{1}{x_0(t)}\theta_i(t)\ \mathrm{d}X_i(t) - \frac{1}{(X_0(t))^2}\ \mathrm{d}X_0(t)\sum_{i=0}^{n}\theta_i(t)X_i(t)\\ &= \sum_{i=1}^{n}\theta_i(t)\ \mathrm{d}\overline{X}_i(t)\end{aligned}$$

□

An appropriate model for a portfolio has now been constructed. However, in order to progress towards option pricing a few assumptions about the market environment must be made.

Arbitrage

Roughly speaking, arbitrage is the opportunity to make money from nothing and to do so without risk. In such a case investors would try to create arbitrage portfolios, creating high demand for certain assets. This would cause disequilibrium in the market as demand would be much more than subtle.

Definition 39 *A self financing portfolio with value v_t is an arbitrage if $v_0 = 0$ and $\mathbb{E}\, v_t > 0$.*

Theorem 40 *Suppose there exists a probability measure Q that is $\mathcal{F}_t^{n+1}$ measurable and such that $P \sim Q$. If $\{\overline{v}_i(t)\}_{t\in[0,T]}$ is a martingale under Q then there is no arbitrage.*

The point to remember from this theorem is that there is no arbitrage if and only if an equivalent martingale measure exists.

In the Black-Scholes economic model

$$\begin{aligned}\mathrm{d}X_0 &= rX_0\ \mathrm{d}t\\ \mathrm{d}X_1 &= \mu X_1\ \mathrm{d}t + \sigma X_1\ \mathrm{d}B_t\end{aligned}$$

which solve to give

$$\begin{aligned}X_0(t) &= X_0(0)e^{rt}\\ X_1(t) &= X_1(0)\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma B_t\right)\end{aligned}$$

Let P be a probability measure under which B_t is a Brownian motion and assume r, μ, and σ are constants. Calculating $\mathbb{E}\, X_1(t)$ shows $X_1(t)$ to be a martingale only if $\mu = 0$. Normalising, $\overline{X}_1(t)$ is only a martingale if $\mu = r$. Now,

$$\mathrm{d}\overline{X}_1(t) = (\mu - r)\overline{X}_1(t)\ \mathrm{d}t + \sigma\overline{X}_1(t)\ \mathrm{d}B_t$$

is the normalised price process, and so a change of drift of $(\mu - r)\overline{X}_1(t)$ is required. Using Theorem 33 $\sigma_t\overline{X}_1(t)\theta_t = -(\mu - r)\overline{X}_1(t)$ so

$$\theta_t = \frac{(\mu - r)}{\sigma}$$

and hence

$$\begin{aligned}\xi_t^\theta &= \exp\left(-\int_0^t \frac{(\mu-r)}{\sigma}\,\mathrm{d}B_s - \frac{1}{2}\int_0^t \frac{(\mu-r)^2}{\sigma^2}\,\mathrm{d}s\right)\\ &= \exp\left(\frac{-(\mu-r)}{\sigma}B_t - \frac{(\mu-r)^2}{2\sigma^2 t}\right)\end{aligned}$$

Having made this change of drift the new $\overline{X}_t$ can be found as follows. First of all note that

$$\tilde{B}_t = B_t + \int_0^t \frac{\mu-r}{\sigma}\,\mathrm{d}s = B_t + \frac{\mu-r}{\sigma}t$$

and hence

$$\mathrm{d}B_t = \mathrm{d}\tilde{B}_t - \frac{\mu-r}{\sigma}\,\mathrm{d}t$$

Substituting this into the equation for $\mathrm{d}\overline{X}_1$ gives

$$\begin{aligned}\mathrm{d}\overline{X}_1 &= (\mu-r)\overline{X}_1(t)\,\mathrm{d}t + \sigma\overline{X}_1(t)\,\mathrm{d}B_t\\ &= (\mu-r)\overline{X}_1(t)\,\mathrm{d}t + \sigma\overline{X}_1(t)\left(\mathrm{d}\tilde{B}_t - \frac{\mu-r}{\sigma}\,\mathrm{d}t\right)\\ &= \sigma\overline{X}_1(t)\,\mathrm{d}\tilde{B}_t\end{aligned}$$

Solving this is similar to the process used for equation (27) and gives

$$\overline{X}_t = \overline{X}_1(0)\exp\left(\frac{\sigma^2}{2}t + \sigma\tilde{B}_t\right)$$

Now finding the expected value

$$\begin{aligned}\mathbb{E}\,\overline{X}_1(t) &= \mathbb{E}\left(\overline{X}_1(0)\exp\left(\frac{\sigma^2}{2}t + \sigma\tilde{B}_t\right)\right)\\ &= \overline{X}_1(t)\exp\left(\frac{-\sigma^2}{2}t\right)\int_{-\infty}^{\infty}\frac{\exp\frac{-y^2}{2t}}{\sqrt{2\pi t}}\exp(\sigma y)\,\mathrm{d}y\\ &= \overline{X}_1(t)\exp\left(\frac{-\sigma^2}{2}t\right)\exp\left(\frac{\sigma^2}{4\left(\frac{1}{2t}\right)}\right)\\ &= \overline{X}_1(0)\end{aligned}$$

So $\overline{X}_t$ is a martingale under Q. Hence by Theorem 40 there is no arbitrage in the Black-Scholes economy.

Using Theorem 33 for multiple assets, σ is a matrix and θ is a vector. The result means that if σ is singular (or if not square then the solutions are inconsistent) then an arbitrage can be constructed. Take for example

$$\begin{aligned}\mathrm{d}X_0 &= 0\\ \mathrm{d}X_1 &= 3\,\mathrm{d}t + \mathrm{d}B_1(t)\\ \mathrm{d}X_2 &= \mathrm{d}t + 2\,\mathrm{d}B_1(t) + 3\,\mathrm{d}B_2(t)\\ \mathrm{d}X_3 &= \mathrm{d}t + 3\,\mathrm{d}B_2(t)\end{aligned}$$

This economy gives rise to the system of equations

$$\begin{pmatrix} 1 & 0 \\ 2 & 3 \\ 0 & 3 \end{pmatrix} ` = \begin{pmatrix} 3 & 1 & 1 \end{pmatrix}$$

Clearly this has no solution and so an arbitrage can be constructed. This is done by combining assets in such a way that the random components are eliminated. In this case the portfolio will have holdings vector $` = (\theta_0, 2, -1, 1)^T$.

More generally σ will be dependent on the actual price process, e.g. $\mathrm{d}X_1 = \mathrm{d}t + X_2 \, \mathrm{d}B_1(t)$. In this case the portfolio is itself a stochastic process. Consider the following example.

Example 41 *Construct an arbitrage in the market*

$$\begin{aligned} \mathrm{d}X_0 &= 0 \\ \mathrm{d}X_1 &= 2 \, \mathrm{d}t + X_1(t) \, \mathrm{d}B_1(t) + X_1(t) \, \mathrm{d}B_2(t) \\ \mathrm{d}X_2 &= - \, \mathrm{d}t - X_2(t) \, \mathrm{d}B_1(t) - X_2(t) \, \mathrm{d}B_2(t) \end{aligned}$$

Proof. Solution The system of equations produced is

$$\begin{pmatrix} X_1 & X_1 \\ -X_2 & -X_2 \end{pmatrix} ` = \begin{pmatrix} 2 & -1 \end{pmatrix}$$

The matrix is clearly singular and it would seem obvious to construct the portfolio $` = (\theta_0, X_2, -X_1)^T$. However, from Itô's formula it is evident that including a term with the product $X_1(t)X_2(t)$ will give rise to other terms. Consider therefore $f(x, y) = xy$ with $x = X_1(t)$ and $y = X_2(t)$ then by Itô's formula

$$\begin{aligned} \mathrm{d}f &= X_2(t) \, \mathrm{d}X_1(t) + X_1(t) \, \mathrm{d}X_2(t) + \frac{1}{2} \, \mathrm{d}X_1(t) \, \mathrm{d}X_2(t) \\ &= X_2(t) \, \mathrm{d}X_1(t) + X_1(t) \, \mathrm{d}X_2(t) + \frac{1}{2} \left(2 \, \mathrm{d}t + X_1(t) \, \mathrm{d}B_1(t) + X_1(t) \, \mathrm{d}B_2(t)\right) \left(- \, \mathrm{d}t - X_2(t) \, \mathrm{d}B_1(t) - X_2(t) \, \mathrm{d}B_2(t)\right) \\ &= X_2(t) \, \mathrm{d}X_1(t) + X_1(t) \, \mathrm{d}X_2(t) - X_1(t)X_2(t) \, \mathrm{d}t \\ &= X_2(t) \left(2 \, \mathrm{d}t + X_1(t) \, \mathrm{d}B_1(t) + X_1(t) \, \mathrm{d}B_2(t)\right) + X_1(t) \left(- \, \mathrm{d}t - X_2(t) \, \mathrm{d}B_1(t) - X_2(t) \, \mathrm{d}B_2(t)\right) - X_1(t)X_2(t) \, \mathrm{d}t \\ &= \left(-X_1(t) + 2X_2(t) - X_1(t)X_2(t)\right) \, \mathrm{d}t \end{aligned}$$

Therefore the portfolio $` = (\theta_0, X_2, X_1)^T$ has value process

$$\mathrm{d}v_t = (X_1 + X_2 - 2X_2X_2) \, \mathrm{d}t$$

□

Completeness

Definition 42 *A market is complete if for all bounded $\mathcal{F}_T$ measurable functions $F(\omega)$, there exists a self financing portfolio $\mathbf{z}$ such that for all $A \in \mathbb{R}$*

$$F(\omega) = A + \int_0^T \mathbf{z}(t) \cdot \, \mathrm{d}\mathbf{X}(t)$$

Completeness really means that a portfolio can be constructed to attain any given value at any given time. Moreover, a market is complete if and only if there exists a unique equivalent martingale measure.

Theorem 43 *Let u be an m dimensional stochastic process on a time interval $[0, T]$ for which*

$$\sigma u = \mu - r\mathbf{x} \quad \text{and} \quad \mathbb{E}\left(\exp\left(\frac{1}{2}\int_0^T u^2(s,\omega)\ \mathrm{d}s\right)\right) < \infty$$

then the market $\mathbf{X}$ is complete if and only if there exists an $\mathcal{F}_t$ adapted matrix process $\Lambda(t,\omega)$ such that $\Lambda\sigma = I$.

Clearly this means that σ must be of full rank, which can provide a quick check for non-completeness.

(20.3.2) Pricing European Options

Definition 44 *A European option is an $\mathcal{F}_t$ measurable random function.*

- *A call option allows the holder to buy shares at a prescribed price K at a prescribed time T.*
- *A put option allows the holder to sell shares at a prescribed price K at a prescribed time T.*

From this it is immediately obvious that

- For the call option the payoff is the greatest of 0 (when the share price does not exceed K i.e. $F(\omega) < K$) and $F(\omega) - K$.
- For the put option the payoff is the greatest of 0 (when $F(\omega) < K$) and $K - F(\omega)$.

How much should one pay to buy such an option?

Option Pricing

From the buyer's point of view, an initial payment of y must be made, and a payoff of $F(\omega)$ is received at time T. In the interim time the buyer can trade his portfolio $\mathbf{z}$ and so

$$v_b(t) = -y + F(\omega) + \int_0^T \mathbf{z}(s)\ \mathrm{d}\mathbf{X}(s)$$

and this must be positive. Clearly the buyer wishes the price of the option to be as low as possible. However, interest lies in the largest price that will allow the buyer to break even. Hence define

$$p_b = \sup\left\{y \mid \exists\mathbf{z} \mid -y + \sum_{i=1}^{n}\int_0^T z_i(t)\ \mathrm{d}\overline{X}_i(t) \geqslant -F(\omega)\right\}$$

A more clear way to think about this is

$$p_b = \sup\left\{y \mid \exists\mathbf{z} \mid y \leqslant \sum_{i=1}^{n}\int_0^T z_i(t)\ \mathrm{d}\overline{X}_i(t) + F(\omega)\right\}$$

For the seller, a sum of Y is received for the option, which he can invest in a portfolio $\mathbf{Z}$. At time T the seller has liability $F(\omega)$, and even though the buyer may not exercise the option the seller must not assume this. Hence

$$v_s = y - F(\omega) + \sum_{i=1}^{n}\int_0^T Z_i(t)\ \mathrm{d}\overline{X}_i(t)$$

Clearly this must be a positive amount. The seller would like the price of the option to be as high as possible, but interest lies in the lowest value such that he does not make a loss. Hence define

$$p_s = \inf\left\{Y \mid \exists\mathbf{Z} \mid Y + \sum_{i=1}^{n}\int_0^T Z_i(t)\ \mathrm{d}\overline{X}_i(t) \geqslant F(\omega)\right\}$$

Generally $p_b \leqslant p_s$ and clearly a sale can only be made if these prices coincide at value p, say, at $t = 0$. For normalised price processes let $\xi_t = \frac{1}{X_0(t)}$ the results become

$$p_b = \sup \left\{ y \mid \exists \mathbf{z} \mid y \leqslant \sum_{i=1}^{n} \int_0^T z_i(t)\, \mathrm{d}\overline{X}_i(t) + \xi_T F(\omega) \right\} \tag{45}$$

$$p_s = \inf \left\{ Y \mid \exists \mathbf{Z} \mid Y + \sum_{i=1}^{n} \int_0^T Z_i(t)\, \mathrm{d}\overline{X}_i(t) \geqslant \xi_T F(\omega) \right\} \tag{46}$$

Theorem 47 *Suppose a measure Q can be defined such that*

$$\frac{\mathrm{d}Q}{\mathrm{d}P} = \exp\left(- \int_0^T \mathbf{u}(t,\omega)\, \mathrm{d}\mathbf{B}(t) - \frac{1}{2} \int_0^T (\mathbf{u}(t,\omega))^2\ \mathrm{d}t \right)$$

where $\sigma \mathbf{u} = \bar{} - r\mathbf{X}$ and $\mathbf{u}$ satisfies the Novikov§ *condition then*

$$p_b \leqslant \mathbb{E}\, Q\left(\xi_T F(\omega)\right) \leqslant p_s$$

Furthermore, if the market is complete, then these inequalities become equality.

Proof. Using Girsanov's theorem (Theorem 32) under the probability measure Q the price processes become

$$\mathrm{d}\overline{X}_0 = 0$$
$$\mathrm{d}\overline{X}_i = \xi_t \sigma_i(t)\, \mathrm{d}\tilde{B}(t) \quad \text{for } i \leqslant i \leqslant n$$

where $\tilde{B}(t)$ is a standard Brownian motion under Q. Hence for the buyer

$$\mathop{\mathbb{E}}_{Q}\left(\sum_{i=1}^{n} \int_0^T z_i(t)\, \mathrm{d}\overline{X}_i(t) \right) = 0$$

since this is a martingale under Q. Hence from equation (45),

$$p_b = y \leqslant \mathop{\mathbb{E}}_{Q}\left(\xi_T F(\omega)\right)$$

For the seller the gains process of the portfolio are also a martingale and hence from equation (46)

$$Y = p_s \geqslant \mathop{\mathbb{E}}_{Q}\left(\xi_T F(\omega)\right)$$

Hence the first part of the theorem.

Now, under completeness it is possible to construct a portfolio for the buyer such that

$$-y + \sum_{i=1}^{n} \int_0^T z_i\, \mathrm{d}\overline{X}_i = -\xi_T F(\omega)$$

Similarly a portfolio for the seller can be constructed such that

$$Y + \sum_{i=1}^{n} \int_0^T Z_i\, \mathrm{d}\overline{X}_i = \xi_T F(\omega)$$

§The Novikov condition on u is that $\mathbb{E}\left(\exp\left(\frac{1}{2}\int_0^T (u(t,\omega))^2\ \mathrm{d}t\right)\right) < \infty$.

Hence taking expectation under the measure Q it is evident that

$$p_b = \mathop{\mathbb{E}}_{Q}\left(\xi_T F(\omega)\right) = p_s \tag{48}$$

$\square$

Pricing At Some Intermediate Time

Using a similar method the value of the option at some time $0 < t < T$ can be found. Note in the above proof it is assumed that $\xi_0 = 1$ and so is ineffective when acting on y or Y. Now, at some time t for the buyer

$$-\xi_t y + \sum_{i=1}^{n} \int_t^T z_i(s)\, \mathrm{d}\overline{X}_i(s) + \xi_T F(\omega) \geqslant 0 \tag{49}$$

Since the gains process is a martingale under Q, at time t no information is available about the future. Hence taking expectation given $\mathcal{F}_t$ gives

$$\mathop{\mathbb{E}}_{Q}\left(\int_t^T z_i(s)\, \mathrm{d}\overline{X}_i(s) \mid \mathcal{F}_t\right) = \int_t^t z_i(s)\, \mathrm{d}\overline{X}_i(s) = 0$$

Therefore taking expectation under Q of equation (49) gives

$$\begin{aligned}\mathop{\mathbb{E}}_{Q}\left(\xi_t y \mid \mathcal{F}_t\right) &\leqslant \mathop{\mathbb{E}}_{Q}\left(\xi_T F(\omega) \mid \mathcal{F}_t\right) \\ p_b(t) = y &\leqslant \left(\xi_t\right)^{-1} \mathop{\mathbb{E}}_{Q}\left(\xi_T F(\omega) \mid \mathcal{F}_t\right)\end{aligned}$$

Similarly for the seller

$$\xi_t Y + \sum_{i=1}^{n} \int_t^T Z_i(s)\, \mathrm{d}\overline{X}_i(s) - \xi_T F(\omega) \geqslant 0$$

Again the conditional expectation under Q of the gains process is zero, and so

$$p_s(t) = Y \geqslant \left(\xi_t\right)^{-1} \mathop{\mathbb{E}}_{Q}\left(\xi_T F(\omega) \mid \mathcal{F}_t\right)$$

Under completeness portfolios for the buyer and seller can be constructed such as to allow equality and hence a mutually acceptable price can again be found. Putting $t = 0$ it is fairly clear these equations are consistent with those of Theorem 47.

Connection With Partial Differential Equations

Consider equation (48) in the context of the simple Black-Scholes economy

$$\begin{aligned}\mathrm{d}X_0(t) &= rX_0(t)\, \mathrm{d}t \\ \mathrm{d}X_1(t) &= \mu X_1(t)\, \mathrm{d}t + \sigma X_1(t)\, \mathrm{d}B_t\end{aligned}$$

Applying Girsanov's theorem (Theorem 32) produces the equivalent probability measure

$$\frac{\mathrm{d}Q}{\mathrm{d}P} = \exp\left(\frac{-(\mu - r)}{\sigma} B_t - \frac{(\mu - r)^2}{2\sigma^2} t\right)$$

So that $\tilde{B}_t = \frac{\mu - r}{\sigma} t + B_t$ is a Brownian motion under Q. Normalising and integrating gives

$$\overline{X}_1(t) = X_1(0) \exp\left(\frac{-\sigma^2}{2} t + \sigma \tilde{B}_t\right)$$

In equation (48) the function $F(\omega)$ represents the payoff of the option, and hence

$$F(\omega) = F(X_1(T)) = F(t, B_T) = \begin{cases} \max\left(X_1(T) - k, 0\right) & \text{for a call} \\ \max\left(k - X_1(T), 0\right) & \text{for a put} \end{cases}$$

Now using the equation for the option price at time $0 \leqslant t \leqslant T$,

$$\begin{aligned} p &= (\xi(t))^{-1} \mathop{\mathbb{E}}_{Q} \left(\xi(T) F(t, \tilde{B}_T) \mid \mathcal{F}_t\right) \\ &= (\xi(t))^{-1} \xi(T) \mathop{\mathbb{E}}_{Q} \left(F(t, \tilde{B}_T) \mid \mathcal{F}_t\right) \end{aligned}$$

Now use the Markov property, so $\mathbb{E}\left(Y \circ \theta_t \mid \mathcal{F}_t\right) = \mathbb{E}_{B_t} Y$.

$$\begin{aligned} &= e^{-r(T-t)} \mathop{\mathbb{E}_{\tilde{B}_t}}_{Q} F(t, \tilde{B}_{T-t}) \\ &= e^{-r\bar{t}} \mathop{\mathbb{E}_{\tilde{B}_t}}_{Q} F(t, \tilde{B}_{\bar{t}}) \quad \text{by changing variables} \end{aligned}$$

Changing back from the Brownian motion to the stock price, let

$$u(\bar{t}, X_1(\bar{t})) = e^{-r\bar{t}} \mathop{\mathbb{E}_{\tilde{B}_t}}_{Q} F(t, X_1(\bar{t})) \tag{50}$$

By Feynman-Kac (Theorem 30) u is the solution to the partial differential equation

$$\frac{\partial u}{\partial \bar{t}} = Au + cu$$

where A is the generator of the Itô diffusion that is $X_1(t)$, so putting $X_1(t) = x_1$ gives

$$A = \mu x_1 \frac{\partial}{\partial x_1} + \frac{\sigma^2 x_1^2}{2} \frac{\partial^2}{\partial x_1^2}$$

and $c(s) = -rs$. Hence

$$\frac{\partial u}{\partial \bar{t}} = \mu x_1 \frac{\partial u}{\partial x_1} + \frac{\sigma^2 x_1^2}{2} \frac{\partial^2 u}{\partial x_1^2} - ru$$

Clearly the variable $\bar{t}$ can be changed back to t to give

$$0 = \frac{\partial u}{\partial t} + \mu x_1 \frac{\partial u}{\partial x_1} + \frac{\sigma^2 x_1^2}{2} \frac{\partial^2 u}{\partial x_1^2} - ru \tag{51}$$

Here u gives the value of the option at time t when the underlying has value x_1. This is the Black-Scholes partial differential equation.

The objective now is to construct the portfolio $\mathbf{z}$ or $\mathbf{Z}$ mentioned in Section 20.3.2. In this simple case the portfolio will be simply $X_0(t) + \Delta(t) X_1(t)$ where $\Delta(t)$ is the quantity to be determined.

Theorem 52 $\Delta = \dfrac{\partial u}{\partial x_1}$.

Proof. Assuming completeness,

$$\xi(t) y = \int_t^T \Delta(t)\, \mathrm{d}\overline{X}_1(t) + \xi(T) F(\omega)$$

Now $y = u(T-t, X_1(t))$ and so applying Itô's formula on $\xi(t) y = e^{-rt} u(T-t, X_1(t))$ (using the product rule) gives

$$\mathrm{d}(\xi(t)y) = \left(\frac{\partial u}{\partial t}\,\mathrm{d}t + \frac{\partial u}{\partial x_1}\,\mathrm{d}x_1 + \frac{1}{2}\frac{\partial^2 u}{\partial x_1^2}(\,\mathrm{d}x_1)^2 - ru\,\mathrm{d}t\right)e^{-rt}$$

But $\mathrm{d}x_1 = \mu x_1\,\mathrm{d}t + \sigma x_1\,\mathrm{d}\tilde{B}_t$ and so

$$\begin{aligned}
&= \left(\frac{\partial u}{\partial t}\,\mathrm{d}t + \mu x_1\frac{\partial u}{\partial x_1}\,\mathrm{d}t + \sigma x_1\frac{\partial u}{\partial x_1}\,\mathrm{d}\tilde{B}_t + \frac{\sigma^2 x_1^2}{2}\frac{\partial^2 u}{\partial x_1^2}\,\mathrm{d}t - ru\,\mathrm{d}t\right)e^{-rt} \\
&= e^{-rt}\sigma x_1\frac{\partial u}{\partial x_1}\,\mathrm{d}\tilde{B}_t \quad \text{from equation (51)} \\
&= \xi(t)\sigma x_1\frac{\partial u}{\partial x_1}\,\mathrm{d}\tilde{B}_t \\
\xi(t)y &= \xi(T)F(\omega) + \int_t^T \xi(t)\frac{\partial u}{\partial x_1}\sigma X_1(t)\,\mathrm{d}\tilde{B}_t
\end{aligned}$$

□

(20.3.3) Pricing American Options

Option Pricing

Unlike European options, American options have no expiry time. The exercise time for an option is therefore a stopping time $t(\omega)$. Hence

$$F(\omega) = \begin{cases} \max\left(X_1(t(\omega)) - k, 0\right) & \text{for a call} \\ \max\left(k - X_1(t(\omega)), 0\right) & \text{for a put} \end{cases}$$

The cases for the buyer and seller can be considered in a similar way to the European option, but the option price is now dependent on ω.

For the buyer

$$p_b = \sup\left\{ y \mid \exists \mathbf{z}\ \exists t(\omega) \mid y \leqslant F(\omega) + \sum_{i=1}^{n}\int_0^{t(\omega)} z_i\,\mathrm{d}\overline{X}_i(t)\right\} \quad \text{for fixed } t(\omega)$$

Now use Girsanov's theorem to find a probability measure Q such that $\overline{X}_i$ are martingales. Taking expectation gives

$$y \leqslant \mathop{\mathbb{E}}_{Q}\left(\xi(t(\omega))F(\omega)\right) \quad \forall y$$

and hence letting $t(\omega)$ vary,

$$p_b \leqslant \sup_{t(\omega)} \mathop{\mathbb{E}}_{Q}\left(\xi(t(\omega))F(\omega)\right) \tag{53}$$

For the seller

$$p_s = \inf\left\{ Y \mid \exists \mathbf{Z}\ \exists t(\omega) \mid Y \geqslant F(\omega) - \sum_{i=1}^{n}\int_0^{t(\omega)} Z_i\,\mathrm{d}\overline{X}_i(t)\right\} \quad \text{for fixed } t(\omega)$$

Using the discounted portfolio and taking expected values under Q,

$$Y \geqslant \mathop{\mathbb{E}}_{Q}\left(\xi(t(\omega))F(\omega)\right) \quad \forall Y$$

and so letting $t(\omega)$ vary, all possibilities must be covered and so

$$p_s \geqslant \sup_{t(\omega)} \mathbb{E}_Q \left(\xi(t(\omega))F(\omega)\right) \tag{54}$$

Hence assuming completeness, equations (53) and (54) give

$$p_b = \mathbb{E}_Q \left(\xi(t(\omega))F(\omega)\right) = p_s \tag{55}$$

It follows that at any time t the option has price

$$p(t) = \sup_{\tau \geqslant t} \left(\xi(t)\right)^{-1} \mathbb{E}_Q \left(\xi(\tau)F(\tau) \mid \mathcal{F}_t\right)$$

Optimal Stopping

When to exercise an American option is of course just as big a problem as what to pay for one. The problem is to find when to stop the process so as to maximise the reward, but to do so without knowledge of the future of the process. This has many applications.

Definition 56 *Let $\{X_t\}$ be a finite sequence of random variables. The Snell envelope of this process is the stochastic process Z_t defined as*

$$\begin{aligned} Z_T &= X_T \\ Z_{t-1} &= \max\left(X_{t-1}, \mathbb{E}\left(Z_t \mid \mathcal{F}_{t-1}\right)\right) \end{aligned}$$

Theorem 57 *The Snell envelope is a supermartingale that dominates $\{X_t\}$ for all t. Moreover, where*

$$\tau = \min\{t \mid Z_t = X_t\}$$

then the stopped process $Z_{t\wedge\tau}$ is a martingale and τ is the optimal stopping time.

Definition 58 *Let f be a measurable function. f is called supermeanvalued with respect to the Itô process X_t if for all stopping times τ*

$$f(x) \geqslant \mathbb{E}_x f(X_\tau)$$

Furthermore, if f is lower semicontinuous then it is called superharmonic.

Lemma 59 *f is superharmonic if and only if $Af \leqslant 0$ where A is the generator of X_t.*

Definition 60 *Let h be a measurable function. The function f is a supermeanvalued majorant of h if*

1. *f is supermeanvalued.*
2. *$f \leqslant h$.*

Furthermore,

- *$\overline{h} = \inf_f\{f(x)\}$ is the least supermeanvalued majorant of h.*
- *Similarly $\hat{h}$ is the least superharmonic majorant of h.*

Theorem 61 *Let g be a continuous function representing the reward of some process. Where g^* is the optimal reward, $g^* = \hat{g}$.*

$\hat{g}$, the least superharmonic majorant of g can be constructed as the limit of a sequence of functions. Say

$$g_0 = g \quad \text{and} \quad g_n = \sup_t \mathop{\mathbb{E}}_x g_{n-1}(X_t)$$

Consider the set $D = \{x \mid g(x) < \hat{g}(x)\}$ then the first exit time of this set is of interest, τ_D say—this is an optimal stopping time. Further, let $U = \{x \mid Ag > 0\}$ where A is the generator of X. Then it is never optimal to stop before the process has exited U. This is fairly intuitive, since Ag is some sort of a derivative and so being positive g must still be 'increasing'.

European options give rise to partial differential equations, and similarly American options give rise to free boundary problem partial differential equations.

To reiterate the optimal stopping problem, the task is to optimally stop a process of the form $\sup_t \mathbb{E}_x g(X_t)$. This may be stated as the task to find Φ and $t \leqslant T$ (T is the first exit time from some domain V) such that

$$\begin{aligned}
\Phi(y) &= \sup_{t \leqslant T} J^t(y) = J^{t^*}(y) \\
J^t(y) &= \mathop{\mathbb{E}}_y \left(\int_0^t f(Y_\tau)\, \mathrm{d}\tau + g(Y_t) \right) \\
\mathrm{d}Y_t &= b(Y_t)\, \mathrm{d}t + \sigma(Y_t)\, \mathrm{d}B_t
\end{aligned}$$

Theorem 62 *Let Y be a stochastic process with generator*

$$L = \sum_{i=1}^{n} b_i(y) \frac{\partial}{\partial y_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} (\sigma\sigma^T)_{ij} \frac{\partial^2}{\partial y_i \partial y_j}$$

then where ϕ is a function that satisfies the conditions

$$\begin{aligned}
L\phi + f &\leqslant 0 && y \in V \setminus \overline{D} \\
L\phi + f &= 0 && y \in D \\
\phi &\geqslant g && y \in V \\
\phi &= g && y \in \partial V \\
\text{where } D &= \{x \in V \mid \phi(x) > g(x)\}
\end{aligned}$$

then $\phi(y) = \Phi(y)$ and the optimal stopping time is

$$t^* = t_D = \inf\{t > 0 \mid y_t \notin D\}$$

This being a free boundary value problem (D depends on ϕ) it is usually very difficult to solve. A further condition may be given to ensure the solution is C^1 on $\mathbb{R}^n$.