MASM006 FINANCIAL MATHEMATICS

(5) BROWNIAN MOTION

So far, we have modelled the behaviour of a share price using discrete time stochastic processes with a small timestep δt. As we take smaller and smaller timesteps δt, our model becomes closer to a process varying continuously in time.

We now adopt a mathematically more sophisticated viewpoint, and work directly with stochastic processes in continuous time. Our basic example of such a process is **Brownian motion**. The stochastic processes we shall need are all obtained by applying various transformations to Brownian motion. The behaviour of these new processes is related to the behaviour of the Brownian motion driving them by means of **stochastic differential equations**. In order to work with these, we will need the tools of **stochastic calculus**, in particular, **Itô's Lemma**. This will enable us to derive the Black-Scholes differential equation.

A **continuous time stochastic process** is simply a family of random variables $(X_t)_{t\geq 0}$, one for each real number $t \geq 0$. The variable t corresponds to time, so that as t varies, X_t represents the path taken by some randomly changing quantity X.

We have seen that for (discrete time) random walks $(S_n)_{n\geq 0}$, the probability distribution of S_n is approximately normal for large n, with mean and variance proportional to n. (This is a consequence of the Central Limit Theorem.) A Brownian motion is a continuous time stochastic process in which this normality is built in:

Definition

A **Brownian motion** (or **Wiener process**) is a continuous time stochastic process $(W_t)_{t\geq 0}$ with the following properties:

(i) there is a constant $\sigma > 0$ such that, for any real numbers s, $t \geq 0$, the random variable $W_{t+s} - W_s$ has normal distribution $\mathcal{N}(0, \sigma^2 t)$;

(ii) given any sequence of times $0 \leq t_0 \leq t_1 \leq \ldots \leq t_n$, the random variables $W_{t_r} - W_{t_{r-1}}$, for $1 \leq r \leq n$, are independent;

(iii) $W_0 = 0$;

(iv) W_t is a continuous function of t with probability 1 (i.e. in almost all possible paths).

The quantity σ^2 is called the **variance** of the Brownian motion. If $\sigma = 1$, the process is called a **standard** Brownian motion. □

Notice that the increments $W_{t+s} - W_s$ of a Brownian motion are **stationary** (that is, they have expectation 0) by (i). This means that the process is "driftless". In particular, we have $\mathbb{E}[W_t] = 0$ for all t. Also, the increments are independent on *disjoint* time intervals, by (ii).

We can easily start from some value a, instead of the initial value 0 imposed in (iii), and build in a drift of μ, by taking the process

$$X_t = a + \mu t + W_t.$$

Simply making the above definition does not, of course, guarantee the existence of a stochastic process with these properties. However, one can construct such a process by making rigorous the idea of taking limits of a sequence of random walks. For a sketch of this, see A. Etheridge, *A Course in Financial Calculus*, §3.2.

We should really say that the process $(W_t)_{t\geq 0}$ is a Brownian motion *with respect to a particular probability measure* $\mathbb{P}$. Just as in the discrete binomial model, where it was convenient to replace the actual market probability p with the risk-neutral probability p_*, we will allow ourselves the freedom to reweight the probabilities, keeping the same set of possible paths for the process. Thus $(W_t)_{t\geq 0}$ may be a Brownian motion with respect to the original probability measure $\mathbb{P}$, but not with respect to some other probability measure $\mathbb{Q}$; for instance, we may have $\mathbb{E}_{\mathbb{Q}}[W_t] \neq 0$. We shall say some more about the formalism of probability measures below.

Simulating Brownian Motion

To simulate Brownian motion in MATLAB, we must of course use an approximation in discrete time. If we fix a small timestep δt and write S_n for our approximation to $W_{n\delta t}$, then we should take

$$S_0 = 0; \qquad S_n = S_{n-1} + \sigma\sqrt{\delta t}\xi_n \quad \text{for } n \geq 1,$$

where the ξ_i are i.i.d. random variables from a standard normal distribution $\mathcal{N}(0, 1)$.

Some Properties of Brownian Motion

(1) (Invariance under scaling.)

If $(W_t)_{t\geq 0}$ is a standard Brownian motion, then so is $(cW_{t/c^2})_{t\geq 0}$ for any $c \neq 0$. This says that the process $(W_t)_{t\geq 0}$ has the fractal-like property, that a typical path of the process will look similar if it is scaled up. For instance, the process $(10W_{t/100})$ should look similar to the original process $(W_t)_{t\geq 0}$ itself. This can be illustrated using MATLAB.

Notice that we are not claiming that the scaled-up path is identical to the original path, only that it is qualitatively similar. This is because the invariance property asserts that the processes $(cW_{t/c^2})_{t\geq 0}$ and $(W_t)_{t\geq 0}$ have the same probability distributions, not that the actual paths they take will be the same.

Proof of (1): We check that the process $(V_t)_{t\geq 0}$ given by $V_t = cW_{t/c^2}$ satisfies the 4 properties in the definition of Brownian motion (with $\sigma = 1$). To do so, we use the fact that $(W_t)_{t\geq 0}$ itself satisfies these properties.

(i) For s, $t > 0$, we have that

$$W_{(t+s)/c^2} - W_{s/c^2} \sim \mathcal{N}(0, t/c^2),$$

since $(W_t)_{t\geq 0}$ is a standard Brownian motion, so that

$$V_{t+s} - V_s = c\left(W_{(t+s)/c^2} - W_{s/c^2}\right) \sim \mathcal{N}(0, c^2t/c^2) = \mathcal{N}(0, t).$$

(ii) Given a sequence of times $0 \leq t_0 \leq t_1 \leq \ldots \leq t_n$, the random variables $W_{t_{r+1}/c^2} - W_{t_r/c^2}$ are independent since $(W_t)_{t\geq 0}$ is a Brownian motion. Hence the random variables $V_{t_{r+1}} - V_{t_r} = c\left(W_{t_{r+1}/c^2} - W_{t_r/c^2}\right)$ are independent.

(iii) $V_0 = cW_0 = 0$.

(iv) As the functions $t \mapsto t/c^2$ and $W \mapsto cW$ are continuous, and $t \mapsto W_t$ is continuous with probability 1, their composite $t \mapsto cW_{t/c^2} = V_t$ is also continuous with probability 1. □

(2) (W_t is nowhere differentiable.)

More precisely, the function $t \mapsto W_t$ is nowhere differentiable, with probability 1.

We will not give a rigorous proof of this, but the following heuristic argument at least makes it clear that we cannot expect W_t to be differentiable. (This should also be clear from the jaggedness of the simulated paths.)

To say that the function is differentiable at some point $t = t_0$ means that the limit

$$\lim_{\delta t\to 0} \frac{W_{t_0+\delta t} - W_{t_0}}{\delta t}$$

exists. But $W_{t_0+\delta t} - W_{t_0} \sim \mathcal{N}(0, \delta t)$, so that $|W_{t_0+\delta t} - W_{t_0}|$ is typically of size $\sqrt{\delta t}$. Thus $|(W_{t_0+\delta t} - W_{t_0})/\delta t|$ will typically be of size $\sqrt{\delta t}/\delta t = (\delta t)^{-1/2}$, which tends to ∞ as $\delta t \to 0$.

(3) (W_t eventually hits any given value.)

More precisely, for any $a > 0$, we have

$$\mathbb{P}[W_t < a \text{ for all } t] = 0,$$

and for any $a < 0$, we have

$$\mathbb{P}[W_t > a \text{ for all } t] = 0.$$

Proof of (3). Suppose that $a > 0$. (The case $a < 0$ is similar.) We begin by defining the **first hitting time**

$$T_a = \sup\{t \mid W_{t'} < a \text{ for all } t' < t\}.$$

Here sup means *supremum* (that is, *least upper bound*), so by the continuity of $t \mapsto W_t$, T_a is the first time t for which $W_t = a$. If there is no t with $W_t = a$ then $W_t < a$ for all t (by continuity, and using $W_0 = 0 < a$), and by convention $T_a = \infty$. We must show that $\mathbb{P}[T_a = \infty] = 0$.

We will show in a moment that

$$\mathbb{P}[T_a < t] = 2\mathbb{P}[W_t > a]. \tag{1}$$

Assuming (1) for now, set $X_t = W_t/\sqrt{t}$ for $t > 0$. As $W_t \sim \mathcal{N}(0, t)$ we have $X_t \sim \mathcal{N}(0, 1)$, and

$$\mathbb{P}[W_t > a] = \mathbb{P}[X_t > a/\sqrt{t}] \to 1/2 \quad \text{as } t \to \infty.$$

Thus $\mathbb{P}[T_a < t] \to 1$ as $t \to \infty$, so that $\mathbb{P}[T_a \geq t] \to 0$ as $t \to 0$. This shows that $\mathbb{P}[T_a = \infty] = 0$, so that $\mathbb{P}[W_t < a$ for all $t] = 0$, as required.

It remains to prove (1). The idea is that, after hitting a at time T_a, the Brownian motion does not "remember" how it reached a, and at any subsequent time is equally likely to be above or below a. To express this formally, we use a conditional probability argument. From the rule $\mathbb{P}[A \mid B]\ \mathbb{P}[B] = \mathbb{P}[A \cap B]$, we have

$$\begin{aligned} \mathbb{P}[W_t - W_{T_a} > 0 \mid t > T_a]\ \mathbb{P}[t > T_a] &= \mathbb{P}[W_t - W_{T_a} > 0 \text{ and } t > T_a] \\ &= \mathbb{P}[W_t - W_{T_a} > 0], \end{aligned}$$

since if $W_t > W_{T_a}$ then automatically $t > T_a$ by continuity. Now the increments in W_t after time T_a are independent of those up to time T_a: more precisely, the stochastic process $(W'_\tau)_{\tau \geq 0}$ defined by $W'_\tau = W_{\tau + T_a} - W_{T_a}$ is a standard Brownian motion. (This works because T_a is a **stopping time**: for any t we can determine whether or not $t \leq T_a$ by examining the path of the Brownian motion up to time t.) Thus for $t > T_a$ we have

$$\mathbb{P}[W_t - W_{T_a} > 0 \mid t > T_a] = \mathbb{P}[W'_{t - T_a} > 0] = 1/2.$$

Thus

$$\frac{1}{2}\mathbb{P}[T_a < t] = \mathbb{P}[W_t - W_{T_a} > 0] = \mathbb{P}[W_t > a],$$

giving (1).

Formal Framework for Probability

Before discussing further properties of Brownian motion, we need to be a bit more precise in the way we formulate probabilistic statements. A **probability triple** $(\Omega, \mathcal{F}, \mathbb{P})$ consists of

- a **sample space** Ω; this is the set of all possible outcomes (for example, all possible histories of the the stock market).
- a collection $\mathcal{F}$ of **events** (subsets of Ω to which a probability can be assigned). The point here is that there can be "weird" sets of outcomes to which it is not possible to assign a probability; such sets will not be members of $\mathcal{F}$.

  We assume that $\mathcal{F}$ is a **σ-algebra** (also called a **σ-ring** or **σ-field**), i.e., $\mathcal{F}$ satisfies the following properties

- $\Omega \in \mathcal{F}$;
- $\mathcal{F}$ is closed under complements:

$$\text{if } A \in \mathcal{F} \text{ then } A' = \{\omega \in \Omega \mid \omega \notin A\} \in \mathcal{F};$$

- Ω is closed under countable unions: if $A_1, A_2, \ldots$ is an infinite sequence of sets in $\mathcal{F}$, then

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}.$$

  (In particular, $\mathcal{F}$ is closed under finite unions: take all but finitely many of the A_n to be the empty set.) Using complements, it then follows that $\mathcal{F}$ is also closed under countable intersections.

- a **probability measure** $\mathbb{P}$, assigning to each $A \in \mathcal{F}$ a probability $\mathbb{P}[A]$ so that the following probability axioms hold:
  - $0 \leq \mathbb{P}[A] \leq 1$ for all $A \in \mathcal{F}$;
  - $\mathbb{P}[\Omega] = 1$;
  - $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ if A and B are disjoint;
  - for an increasing sequence $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$ of sets in $\mathcal{F}$,

$$\mathbb{P}[A_n] \to \mathbb{P}\left[\bigcup_{m=1}^{\infty} A_m\right] \quad \text{as } n \to \infty.$$

A random variable X is then just a function $X\colon \Omega \longrightarrow \mathbb{R}$ (where $\mathbb{R}$ is the set of real numbers) for which $\mathbb{P}[a \leq X \leq b]$ is defined whenever $a \leq b$ (more formally, the set

$$\{\omega \in \Omega \mid a \leq X(\omega) \leq b\}$$

is in $\mathcal{F}$). We then say that X is **$\mathcal{F}$-measurable**.

A **filtration** $\{\mathcal{F}_t\}_{t \geq 0}$ is a family of σ-algebras, one for each real number $t \geq 0$, such that $\mathcal{F}_s \subseteq \mathcal{F}_t$ whenever $s \leq t$, and $\mathcal{F}_t \subseteq \mathcal{F}$ for all t. (Think of t as time, and $\mathcal{F}_t$ as all events which are already known by time t.) We write $(\mathcal{F}_t^X)_{t \geq 0}$ for the **natural filtration** associated to the stochastic process $(\mathcal{F}_t)_{t \geq 0}$: an event belongs to $\mathcal{F}_t^X$ if and only if we can determine whether it occurs by examining the path $(X_s)_{0 \leq s \leq t}$ of the process up to time t.

We say that the stochastic process $(X_t)_{t \geq 0}$ is **adapted** to the filtration $(\mathcal{F}_t)_{t \geq 0}$ if, for each $t \geq 0$, the random variable X_t is $\mathcal{F}_t$-measurable. If we think of $\mathcal{F}_t$ as the history of the system up to time t, then $(X_t)_{t \geq 0}$ is adapted to $(\mathcal{F}_t)_{t \geq 0}$ if and only if the value taken by X_t is determined by time t.

Examples:

(1) $(X_t)_{t \geq 0}$ is adapted to its natural filtration $(\mathcal{F}_t^X)_{t \geq 0}$, by the definition of $\mathcal{F}_t^X$.

(2) The stochastic processes $(Y_t)_{t\geq 0}$, $(M_t)_{t\geq 0}$, $(Z_t)_{t\geq 0}$, defined by

$$Y_t = \int_0^t X_s \mathrm{d}s, \qquad M_t = \max_{0\leq s\leq t} X_s, \qquad Z_t = X_t^3 - X_{t/2},$$

are all adapted to the natural filtration $(\mathcal{F}_t^X)_{t\geq 0}$, but the stochastic process $(V_t)_{t\geq 0}$, defined by

$$V_t = X_{2t} + X_t$$

is not. (The value of V_t is not determined until time $2t$ is reached.)

Conditional Expectations

Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration of the σ-algebra $\mathcal{F}$. For a random variable X, we write $\mathbb{E}[X \mid \mathcal{F}_t]$ for the conditional expectation of X relative to $\mathcal{F}_t$. Then $\mathbb{E}[X \mid \mathcal{F}_t]$ is a random variable which is $\mathcal{F}_t$-measurable. If we think of $\mathcal{F}_t$ as the history up to time t, then we should interpret $\mathbb{E}[X \mid \mathcal{F}_t]$ as the expectation of X given the history up to time t; it is a random variable whose value is determined at time t.

Just as for discrete time stochastic processes, we have the 3 rules for conditional expectations:

(1) (Time 0 Rule).

$$\mathbb{E}[Z \mid \mathcal{F}_0] = \mathbb{E}[Z].$$

(2) (Tower Law). For $t > s$

$$\mathbb{E}[\mathbb{E}[Z \mid \mathcal{F}_t] \mid \mathcal{F}_s] = \mathbb{E}[Z \mid \mathcal{F}_s].$$

In particular $\mathbb{E}[\mathbb{E}[Z \mid \mathcal{F}_t]] = \mathbb{E}[Z]$.

(3) (Taking Out a Known Factor). If the random variable Y is $\mathcal{F}_t$-measurable then

$$\mathbb{E}[YZ \mid \mathcal{F}_t] = Y\mathbb{E}[Z \mid \mathcal{F}_t].$$

□

Martingales

We define continuous time martingales just as in the discrete time case, except that we build in some extra flexibility by explicitly mentioning the filtration.

Definition

A continuous time stochastic process $(M_t)_{t\geq 0}$ is called a **martingale** with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ if

(i) $\mathbb{E}[M_t \mid \mathcal{F}_s] = M_s$ for all $t > s$;

(ii) $\mathbb{E}[\,|M_t|\,] < \infty$ for all t.

As before the main point is (i), which says that a martingale is "driftless". Condition (ii) is a technical restriction: a martingale is not allowed to "get too big too quickly".

The terms $\mathbb{E}[M_t \mid \mathcal{F}_s]$ and $\mathbb{E}[\,|M_t|\,]$ are defined with respect to a particular probability measure $\mathbb{P}$ on the underlying sample space Ω. When we need to emphasize this, we say that $(M_t)_{t\geq 0}$ is a martingale with respect to $\mathbb{P}$ (and with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$), and we write the above conditions as

$$\mathbb{E}_{\mathbb{P}}[M_t \mid \mathcal{F}_s] = M_s \quad \text{for all } t > s; \qquad \mathbb{E}_{\mathbb{P}}[\,|M_t|\,] < \infty \quad \text{for all } t.$$

Just as in the discrete case, we can often start with a stochastic process which is not a martingale (with respect to the given probability measure $\mathbb{P}$), and turn it into a martingale by "reweighting the probabilities", that is, by replacing $\mathbb{P}$ by a suitably chosen new probability measure $\mathbb{P}_*$.

The following result should not come as a surprise, given our examples of discrete time martingales.

Lemma.
Let $(W_t)_{t\geq 0}$ be a standard Brownian motion, and let $(\mathcal{F}_t)_{t\geq 0}$ be its associated filtration. Then

(i) $(W_t)_{t\geq 0}$ is a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$;

(ii) the stochastic process $(W_t^2 - t)_{t\geq 0}$ is a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ (but $(W_t^2)_{t\geq 0}$ is not).

Proof. (i) We have to show that $\mathbb{E}[W_t \mid \mathcal{F}_s] = W_s$ whenever $t > s$, and that $\mathbb{E}[\,|W_t|\,] < \infty$ for all t.

Now

$$\mathbb{E}[W_t \mid \mathcal{F}_s] = \mathbb{E}[(W_t - W_s) + W_s \mid \mathcal{F}_s] = \mathbb{E}[W_t - W_s \mid \mathcal{F}_s] + W_s.$$

Since the increments in the Brownian motion are stationary, and the increments after time s are independent of the history up to time s, we have

$$\mathbb{E}[W_t - W_s \mid \mathcal{F}_s] = \mathbb{E}[W_t - W_s] = 0.$$

Thus

$$\mathbb{E}[W_t \mid \mathcal{F}_s] = 0 + W_s = W_s,$$

as required.

We must also check that $\mathbb{E}[\,|W_t|\,] < \infty$. But $W_t \sim \mathcal{N}(0,t)$, and for a random variable X with distribution $\mathcal{N}(0,t)$, we have

$$\mathbb{E}[\,|X|\,] = \sqrt{\frac{2t}{\pi}}.$$

(ii) Writing $M_t = W_t^2 - t$, we have for $t > s$ that

$$\begin{aligned}\mathbb{E}[M_t \mid \mathcal{F}_s] &= \mathbb{E}[(W_t - W_s)^2 + 2W_s(W_t - W_s) + W_s^2 - t \mid \mathcal{F}_s] \\ &= \mathbb{E}[(W_t - W_s)^2 \mid \mathcal{F}_s] + 2W_s\mathbb{E}[W_t - W_s \mid \mathcal{F}_s] + W_s^2 - t.\end{aligned}$$

Now $\mathbb{E}[W_t - W_s \mid \mathcal{F}_s] = 0$ since we have already shown that $(W_t)_{t\geq 0}$ is a martingale. The behaviour of the Brownian motion from time s on is again a Brownian motion: more precisely, the process $(W'_\tau)_{\tau\geq 0}$ defined by $W'_\tau = W_{\tau+s} - W_s$ is itself a standard Brownian motion. Thus

$$\mathbb{E}[(W_t - W_s)^2 \mid \mathcal{F}_s] = \mathbb{E}[{W'_{t-s}}^2] = \mathrm{Var}[W'_{t-s}] = t - s.$$

Putting all this together,

$$\mathbb{E}[M_t \mid F_s] = (t - s) + 0 + W_s^2 - t = W_s^2 - s = M_s.$$

This shows that the process $(M_t)_{t\geq 0}$ satisfies the first condition in the definition of a martingale. It also shows that

$$\mathbb{E}[W_t^2 \mid \mathcal{F}_s] = W_s^2 + (t - s) \neq W_s^2,$$

so that the process $(W_t^2)_{t\geq 0}$ cannot be a martingale (it has a positive drift).

Finally, we check that

$$\mathbb{E}[\,|M_t|\,] = \mathbb{E}[\,|W_t^2 - t|\,] \leq \mathbb{E}[\max(W_t^2, t)] = \max(\mathbb{E}[W_t^2], t) = \max(t, t) = t < \infty.$$

□

As an application of the martingale property for the standard Brownian motion $(W_t)_{t\geq 0}$, we calculate the covariances of this process. Recall that for any random variables X and Y we have

$$\mathrm{Cov}[X, Y] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Lemma
Let $(W_t)_{t\geq 0}$ be a standard Brownian motion. Then

$$\mathrm{Cov}[W_s, W_t] = \min(s, t).$$

Proof. Let $(\mathcal{F}_t)_{t\geq 0}$ be the natural filtration associated to $(W_t)_{t\geq 0}$. Without loss of generality, we suppose $s \leq t$. We know $\mathbb{E}[W_s] = \mathbb{E}[W_t] = 0$. Using the fact that W_s is $\mathcal{F}_s$-measurable, we calculate

$$\begin{aligned}\mathrm{Cov}[X, Y] &= \mathbb{E}[W_s W_t] - 0 \\ &= \mathbb{E}[\mathbb{E}[W_s W_t \mid \mathcal{F}_s]] \quad \text{(Tower Law.)} \\ &= \mathbb{E}[W_s \mathbb{E}[W_t \mid \mathcal{F}_s]] \quad \text{(Taking out } W_s\text{.)} \\ &= \mathbb{E}[W_s W_s] \quad \text{(as } (W_t)_{t\geq 0} \text{ is a martingale.)} \\ &= \mathrm{Var}[W_s] \\ &= s.\end{aligned}$$

□

Nigel Byott

February 2007