Brownian Motion and Ito's Lemma

Jacob Lundbeck Serup

September 16. 2024

A continuous time dynamic process

Definition (Continuous time process)

A continuous time process is a sequence $\{X_0, ..., X_t, ...\}$ of variables indexed by time t for all $t > 0$. If the variables X_t are random, we say that X_t follows a stochastic continuous time process and if the variables X_t are deterministic for all $t > 0$, we say that X_t follows a deterministic continuous time process.

To be able to distinguish between a stochastic process and a function of time, we will denote a stochastic process in time t by X_t - that is with t in the subscript t and a differentiable function of time by $X(t)$.

However, some care must be taken so to not confuse X_t with a partial derivative in t and different notation will be used for partial derivatives.

A continuous time dynamic process

When working with a stochastic process, X_t, we will write its dynamics as something like

$$dX_t = \mu_t(\cdot)d(\cdot) + \sigma_t(\cdot)d(\cdot), \quad t \geq 0. \tag{1}$$

where μ_t and σ_t could be constant, deterministic processes or even stochastic processes. You can loosely interpret dX_t as *the increment of X_t for an infinitesimally small change in time*. However, most of our processes will be continuous and therefore

$$\lim_{\delta \searrow 0} X_{t+\delta} - X_t = 0. \tag{2}$$

The best and most correct way to understand dX_t is in the *Fundamental Theorem of Calculus sense* as

$$X_t = X_s + \int_s^t dX_u$$

A continuous time dynamic process

Many times, we will be presented with the dynamics of some stochastic process and we will look for a 'solution' corresponding to those dynamics.

That is, we will try to find a stochastic process X_t that is consistent with our dynamics and be able to write an expression for X_t of the form

$$X_t = \dots \tag{3}$$

where the right hand side does not depend on X_t. We can always integrate the dynamics of X_t and write something like

$$X_t = X_0 + \int_0^t dX_u. \tag{4}$$

but that is typically not what we mean by a 'solution' and certainly not if the right hand side involves X_u.

Dynamics of the bank account

As an example of one of the more simple stochastic processes we will encounter, we can look at the bank account.

The bank account follows a stochastic process with dynamics given by

$$\begin{aligned} dB_t &= r_t B_t dt, \quad t \geq 0 \\ B_0 &= b_0 \end{aligned} \tag{5}$$

where r_t follows a stochastic process and b_0 a known initial value.

We can ingrate the dynamics of the bank account and write

$$B_t = b_0 + \int_0^t dB_u = b_0 + \int_0^t r_u B_u du, \quad t \geq 0 \tag{6}$$

But here the right hand side contains B_u and this is not a solution

Dynamics of the bank account

In the case of the bank account, it will turn out that even if r_t follows a stochastic process, there exists a solution $B(t)$ to (5) that is differentiable with respect to t. This solution $B(t)$ solves the ODE

$$\begin{aligned}\frac{\partial B(t)}{\partial t} &= r_t B(t) \\ B(0) &= b_0 \end{aligned} \tag{7}$$

To the naked eye, it might seem that all we did was to 'divide' by dt on both sides of (5) and then somehow turned 'division by dt' into a partial derivative.

This is however not what happened and 'tricks' like this will often fail.

One example of a process that is not differentiable and for which *dividing by dt* will not work is when the process is driven by Brownian motion since Brownian motion is not differentiable. More about this very soon.

Dynamics of the bank account

Having written the bank account as satisfying an *Ordinary Differential Equation* (ODE), we can find the solution to be

$$B(t) = b_0 \exp\left(\int_0^t r_u du\right) \tag{8}$$

To convince yourself that (8) is indeed the solution to (7), simply differentiate (8) with respect to t using the chain rule and Leibniz rule for differentiation involving integrals.

$$\begin{aligned}\frac{\partial B(t)}{\partial t} &= b_0 \exp\left(\int_0^t r_u du\right) \frac{\partial}{\partial t}\left(\int_0^t r_u du\right) = b_0 \exp\left(\int_0^t r_u du\right) r_t \\ &= r_t B(t) \end{aligned} \tag{9}$$

Probability Spaces

Definition (Probability Space)

A probability space consists of the following three elements

1) *The sample space* Ω.
2) *The* σ*-algebra* $\mathcal{F}$ *of all events under consideration.*
3) *The probability measure* $\mathbb{P}$ *associated with all the events in* $\mathcal{F}$.

The Sample Space Ω

Depending on the situation, what is meant by sample space will differ.

If we consider a single random variable, the sample space is simply *all* the possible values the random variable can take. If for example $X \sim N(\mu, \sigma)$, then $\Omega_X = (-\infty, \infty)$.

If on the other hand we consider a stochastic process, the sample space includes all the possible trajectories that the stochastic process can follow. If for example X_t is a simple random walk that only takes two steps, then $\Omega = \{[0,1,2],[0,1,0],[0,-1,0],[0-1,-2]\}$.

The σ-algebra $\mathcal{F}$

A σ-algebra is simply a collection of sets where each set consists of a subset of the sample space.

A σ-algebra always contains the empty set, $\varnothing$, and the entire sample space, Ω, but typically also contains a vast number of other non-trivial sets.

A σ-algebra is closed under countable union and countable intersection. This implies that if sets $A, B \in \mathcal{F}$ then also $A \cup B \in \mathcal{F}$ and $A \cap B \in \mathcal{F}$.

In this course, you can simply think of a σ-algebra as a huge collection of sets containing all the sets we could possibly wish to compute the probability of.

The Probability measure $\mathbb{P}$

The probability measure $\mathbb{P}$ assigns a non-negative number to all sets of events contained in $\mathcal{F}$.

Measure $\mathbb{P}$ assigns measure 1 to Ω, $\mathbb{P}(\Omega) = 1$ and measure 0 to $\varnothing$, $\mathbb{P}(\varnothing) = 0$.

The probability measure satisfies all the standard rules of probability. If for example A and B are disjoint sets, then the measure of the union of A and B is simply the sum of the measure of A and the measure of B.

Probability Spaces

In this course, you will need a basic understanding of probability spaces and in particular the concept of a *filtration generated by a stochastic process*.

The filtration of $\mathcal{F}_T$ is a collection in T of σ-algebras.

$\mathcal{F}_T$ for T fixed can be interpreted as an object containing *all information available at time* T. Knowing $\mathcal{F}_T$, we know the outcome of random variables and stochastic processes prior to T of variables adapted to the filtration.

Knowing $\mathcal{F}_T$ will not allow us to determine the out come of variables after T but the conditional distribution of $X_u|\mathcal{F}_T$ for $u > T$ will depend on T.

Likewise, $\mathcal{F}_T(X_t)$ can be interpreted as a collection of σ-algebras containing all the information one can extract by observing the stochastic process X_t up to time T.

Filtrations

We will say that a stochastic process X_t is adapted to a certain filtration $\mathcal{F}_T(\,\cdot\,)$ if, given the filtration, we can evaluate X_t for all $t \leq T$.

In addition, knowing $\mathcal{F}_T(\,\cdot\,)$, also changes how we view probabilities of future events. For example, knowing $\mathcal{F}_T(\,\cdot\,)$ also changes our assessment of $P(X_u \leq a)$ for $u > T$.

Another question that will arise is whether one stochastic process is adapted to the filtration generated by another stochastic process. If for example we know $\mathcal{F}_T(X_t)$, we may or may not be able to evaluate Y_u for $u \leq T$ and knowing $\mathcal{F}_T(X_t)$, may or may not influence how we compute probabilities of events involving Y_u for $u > T$.

When addressing these questions, we will need to use probability spaces and the notation introduced here.

First order variation

Definition (First Order Variation)

The first order variation, $\mathrm{FV}_T(X_t)$*, of a process* X_t *on* $[0, T]$ *is given by*

$$\mathrm{FV}_T(X_t) = \lim_{\|\Pi\| \nearrow 0} \sum_{j=0}^{n-1} |X_{t_{j+1}} - X_{t_j}| \qquad (10)$$

where $\Pi = [0 = t_0 < t_1, ..., t_n = T]$ *is a partition of* $[0, T]$ *that may or may not be equally spaced but will have maximum step size given by*

$$\|\Pi\| = \max_{j=0,\ldots,n-1} (t_{j+1} - t_j)$$

FVof a differential function or stochastic process

Corollary

if $X(t)$ is a differentiable function, the first order variation of $X(t)$ from 0 *to T can be computed as*

$$\mathrm{FV}_T(X(t)) = \int_0^T |X'(t)| dt$$

To prove this very simple result, simply use that $X(t)$ is differentiable and the *Mean Value Theorem*.

Quadratic variation

Definition (Quadratic variation)

The quadratic variation, $\mathrm{QV}_T(X_t) = [X_t, X_t](T)$*, of a process on* $[0, T]$ *is given by*

$$\mathrm{QV}_T(X_t) = \lim_{\|\Pi\| \nearrow 0} \sum_{j=0}^{n-1} \left| X t_{j+1} - X_{t_j} \right|^2 \tag{11}$$

where $\Pi = [0 = t_0 < t_1, ..., t_n = T]$ *is a partition of* $[0, T]$ *that may or may not be equally spaced but will have maximum step size given by*

$$\|\Pi\| = \max_{j=0,\dots,n-1} (t_{j+1} - t_j)$$

Quadratic variation

Lemma

If the process X_t is continuous and has strictly positive quadratic variation, the process must have infinite first order variation.

Quadratic variation

Proof:

The proof is by contraposition (if $B^c \Rightarrow A^c$ then $A \Rightarrow B$). Suppose X_t is continuous and also suppose that $\mathrm{FV}_T(X)$ is finite

$$\sum_{j=0}^{n-1} \left|X_{t_{j+1}} - X_{t_j}\right|^2 \leq \max_{\Pi}\left|X_{t_{j+1}} - X_{t_j}\right| \cdot \sum_{j=0}^{n-1} \left|X_{t_{j+1}} - X_{t_j}\right| \tag{12}$$

and since X_t is continuous process and $\mathrm{FV}_T(X_t)$ is finite, we have that

$$\mathrm{QV}_T(X) \leq \lim_{\|\Pi\| \nearrow 0} \max_{|s-t| \leq \|\Pi\|} \left|X_t - X_s\right| \cdot \mathrm{FV}_T(X_t) = 0 \tag{13}$$

So, for a continuous process, we have the if $\mathrm{FV}_T(X_t)$ is finite, the quadratic variation is 0. We have thus proven that if the quadratic variation of a continuous process is strictly positive, the first order variation is infinite.

□

The simple random walk

Let us consider the simple discrete time random walk

$$M_t = \sum_{j=1}^{t} X_j, \qquad X_j = \begin{cases} 1, & \text{with probability } \frac{1}{2} \\ -1, & \text{with probability } \frac{1}{2} \end{cases} \tag{14}$$

where $t = 1, 2, 3, \ldots$ and X_j are independent (and identically distributed).

Clearly, we have that

$$\mathrm{E}[M_t] = 0, \quad \mathrm{Var}[M_t] = t, \quad \mathrm{FV}(M_t) = t, \quad \mathrm{QV}(M_t) = t, \tag{15}$$

Also, the increments of M_t are stationary and independent.

The scaled simple random walk

To create a continuous time process that is suitable as the diffusion of the price of a financial asset or interest rates, we will

- Increase the number of steps
- Decrease the size of each step
- Insure that the variance and quadratic variation still equals time t

These considerations lead is to define the scaled simple random walk

$$W_t^{(n)} = \frac{1}{\sqrt{n}} M_{nt} = \frac{1}{\sqrt{n}} \sum_{j=1}^{nt} X_j, \quad \text{for } nt \text{ an integer} \tag{16}$$

The scaled simple random walk

The scaled simple random walk has stationary independent increment. Also, we have that

$$\begin{aligned}
\mathrm{E}\left[W_t^{(n)}\right] &= \frac{1}{\sqrt{n}}\sum_{j=1}^{nt}\mathrm{E}\left[X_j\right] = 0 \\
\mathrm{Var}\left[W_t^{(n)}\right] &= \mathrm{Var}\left[\frac{1}{\sqrt{n}}\sum_{j=1}^{nt}X_j\right] = \frac{1}{n}\sum_{j=1}^{nt}\mathrm{Var}\left[X_j\right] = t \\
\mathrm{FV}\left(W_t^{(n)}\right) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{nt}\left|X_j\right| = t\sqrt{n} \\
\mathrm{QV}\left(W_t^{(n)}\right) &= \frac{1}{n}\sum_{j=1}^{nt}X_j^2 = t \qquad (17)
\end{aligned}$$

for nt a positive integer.

The scaled simple random walk

To turn the scaled simple random walk into a process we can use to create randomness in our stochastic processes, we still have to

- Interpolate so that the process becomes a continuous time process,
- Send n to infinity.

The scaled simple random walk is *not* differentiable with respect to t for any finite n. As it turns out, it will also not be differentiable in the limit.

The first order variation diverges as $n \nearrow \infty$ but the quadratic variation is independent of n and stays equal to t.

The scaled simple random walk

If we look at the increment of the scaled simple random walk for $s < t$, and ns and nt integers

$$W_t^{(n)} - W_s^{(n)} = \frac{1}{\sqrt{n}} \sum_{j=1}^{nt} X_j - \frac{1}{\sqrt{n}} \sum_{j=1}^{ns} X_j = \frac{1}{\sqrt{n}} \sum_{j=ns+1}^{nt} X_j \tag{18}$$

then this is the sum of independent and identically distributed random variables so, by the Central Limit Theorem, we have that

$$W_t^{(n)} - W_s^{(n)} \stackrel{\mathcal{D}}{\to} N(0, t-s), \quad \text{as } n \nearrow \infty \tag{19}$$

That is, the increments of the scaled simple random walk converge in distribution to that of a normal random variable with mean 0 and variance equal to the time frame of the increment.

Brownian Motion

Definition (Brownian motion)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Suppose there exists a continuous sequence W_t of $t \geq 0$ for each $\omega \in \Omega$, that satisfies $W_0 = 0$ and that depends on ω. Then W_t is a Brownian motion if for all $0 = t_0 < t_1 < \ldots, t_m$, the increments

$$W_{t_1} = W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \ldots, W_{t_m} - W_{t_{m-1}}$$

are independent and each of these increments is normally distributed with

$$\mathrm{E}\left[W_{t_{i+1}} - W_{t_i}\right] = 0$$

$$\mathrm{Var}\left[W_{t_{i+1}} - W_{t_i}\right] = t_{i+1} - t_i$$

Distribution of Brownian Motion

Because the increments of Brownian motion are independent and normally distributed, the random variables $W_{t_1}, W_{t_2}, ..., W_{t_m}$ are jointly normally distributed.

$$\mathrm{E}[W_t] = 0 \tag{20}$$

$$\mathrm{Cov}[W_s, W_t] = s, \quad s \leq t \tag{21}$$

The covariance can be computed using the fact that Brownian motion has independent increments.

$$\begin{aligned}
\mathrm{Cov}[W_s, W_t] &= \mathrm{E}[W_s W_t] = \mathrm{E}[W_s(W_t - W_s + W_s)] = \mathrm{E}[\mathrm{E}[W_s(W_t - W_s + W_s)|\mathcal{F}_s]] \\
&= \mathrm{E}[\mathrm{E}[W_s(W_t - W_s)|\mathcal{F}_s]] + \mathrm{E}[\mathrm{E}[W_s^2|\mathcal{F}_s]] \\
&= \mathrm{E}[\mathrm{E}[W_s|\mathcal{F}_s] \cdot \mathrm{E}[(W_t - W_s)|\mathcal{F}_s]] + \mathrm{E}[\mathrm{E}[W_s^2|\mathcal{F}_s]] = 0 + s = s
\end{aligned} \tag{22}$$

Brownian motion as a martingale

Theorem (Martingale property for Brownian motion)

Brownian motion is a martingale.

To see that Brownian motion is a martingale, we can perform the following very simple computation.

Let $0 \leq s \leq t$. Then

$$\mathrm{E}\left[W_t | \mathcal{F}_s\right] = \mathrm{E}\left[\left(W_t - W_s\right) + W_s | \mathcal{F}_s\right] = \mathrm{E}\left[\left(W_t - W_s\right) | \mathcal{F}_s\right] + \mathrm{E}\left[W_s | \mathcal{F}_s\right] = W_s$$

Hence, Brownian motion has no drift and therefore neither has a tendency to rise nor fall over time.

Theorem (Markov property for Brownian motion)

Brownian motion is a Markov process.

The Martingale property says that

$$\mathrm{E}\big[f(W_t)\big|\mathcal{F}_s\big] = g(W_s), \quad s \leq t \tag{23}$$

for f and g some measurable functions.

To evaluate the expected value of a function of Brownian motion at time t in the future given the filtration up to time s, the only information that matters is the value W_s of the Brownian motion at time s. The path of Brownian motion up to time s is irrelevant given W_s.

Brownian motion as a Markov process

Proof:

We need to show that for $0 \leq s \leq t$ and f a measurable function, there is another measurable function g such that

$$\mathrm{E}\big[f(W_t)\big|\mathcal{F}_s\big] = g(W_s) \tag{24}$$

To do this we write

$$\mathrm{E}\big[f(W_t - W_s + W_s)\big|\mathcal{F}_s\big] \tag{25}$$

Now, replace the last W_s with a dummy variable X and compute the conditional expectation using the PDF of a normal random variable with $W_t - W_s$ as the argument. Then you are left with some function g of the dummy variable X, so simply replace X by W_s and we have the result.

$\square$

Quadratic variation of Brownian motion

Theorem

Let W_t be a Brownian motion for $t \geq 0$, then $\mathrm{QV}_T(W_t) = T$ for all $T > 0$ with probability one.

Proof:

To compute the quadratic variation of a Brownian motion, we use the partition $\Pi = \{t_0 = 0, t_1, t_2, ..., t_n = T\}$ and define the sample quadratic variation corresponding to this partition to be

$$\mathrm{Q}_\Pi = \sum_{j=0}^{n-1} \left(W_{t_{j+1}} - W_{t_j}\right)^2 \tag{26}$$

Quadratic variation of Brownian motion

Computing the mean of Q_Π gives us

$$\mathrm{E}[\mathrm{Q}_\Pi] = \sum_{j=0}^{n-1} \mathrm{E}\left[\left(W_{t_{j+1}} - W_{t_j}\right)^2\right] = \sum_{j=0}^{n-1} (t_{j+1} - t_j) = T \tag{27}$$

Moreover

$$\begin{aligned} \mathrm{Var}\left[\left(W_{t_{j+1}} - W_{t_j}\right)^2\right] &= \mathrm{E}\left[\left(W_{t_{j+1}} - W_{t_j}\right)^4\right] - \left(\mathrm{E}\left[\left(W_{t_{j+1}} - W_{t_j}\right)^2\right]\right)^2 \\ &= 3\left(t_{j+1} - t_j\right)^2 - \left(t_{j+1} - t_j\right)^2 = 2\left(t_{j+1} - t_j\right)^2 \end{aligned} \tag{28}$$

Thus

$$\begin{aligned} \mathrm{Var}[\mathrm{Q}_\Pi] &= \sum_{j=0}^{n-1} \mathrm{Var}\left[\left(W_{t_{j+1}} - W_{t_j}\right)^2\right] = \sum_{j=0}^{n-1} 2\left(t_{j+1} - t_j\right)^2 \\ &\leq \sum_{j=0}^{n-1} 2\|\Pi\|\left(t_{j+1} - t_j\right) = 2\|\Pi\| T \underset{\|\Pi\|\to 0}{\to} 0 \end{aligned} \tag{29}$$

Quadratic variation of Brownian motion

The sample quadratic variation is linked to the quadratic variation in that

$$\mathrm{QV}_T(W_t) = \lim_{\|\Pi\| \to 0} \mathrm{Q}_\Pi \tag{30}$$

However as $n \nearrow \infty$, the variance of the random variable Q_Π goes to 0 as the partition becomes finer and we conclude that

$$\mathrm{QV}_T(W_t) = \lim_{\|\Pi\| \to 0} \mathrm{Q}_\Pi = \mathrm{E}[\mathrm{Q}_\Pi] = T \tag{31}$$

to give us the result.

□

Cross variation of Brownian motion with time

The covariation of Brownian with time can be found from

$$
\begin{aligned}
\mathrm{QV}_T(W_t, t) &= \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \left(W_{t_{j+1}} - W_{t_j}\right)\left(t_{j+1} - t_j\right) \\
&\leq \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \left|\left(W_{t_{j+1}} - W_{t_j}\right)\left(t_{j+1} - t_j\right)\right| \\
&\leq \lim_{\|\Pi\|\to 0} \max_{0\leq k\leq n-1} \left|W_{t_{k+1}} - W_{t_k}\right| \sum_{j=0}^{n-1} \left(t_{j+1} - t_j\right) \\
&= \lim_{\|\Pi\|\to 0} \max_{0\leq k\leq n-1} \left|W_{t_{k+1}} - W_{t_k}\right| T = 0
\end{aligned}
\tag{32}
$$

where the last equality comes from the fact that W_t is a continuous process implying that even the largest increment of a path of Brownian motion goes to zero almost surely.

Quadratic variation time

In a similar spirit to the previous calculations, we can compute the 'quadratic variation of time' as follows

$$\mathrm{QV}_T(t) = \lim_{\|\Pi\|\to 0} \sum_{j=1}^{n-1} \left(t_{j+1} - t_j\right)^2 \leq \lim_{\|\Pi\|\to 0} \|\Pi\| \sum_{j=1}^{n-1} \left(t_{j+1} - t_j\right) = \lim_{\|\Pi\|\to 0} \|\Pi\| T = 0 \tag{33}$$

Quadratic variation of Brownian motion

The results we have just seen regarding the variation of Brownian motion with respect to itself and to time can be used to say something very important about a number of integrals that will be important to understand Ito's lemma. In particular, we have that

$$\int_0^T dW_t dW_t = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \left(W_{t_{j+1}} - W_{t_j}\right)^2 = T \tag{34}$$

It is not hard to write T as an integral to give us

$$\int_0^T \left(dW_t\right)^2 = T = \int_0^T dt \tag{35}$$

Quadratic variation of Brownian motion

Now, this is true for any $T > 0$ no matter how small a T and it is extremely tempting to conclude that the integrand of the two integrals must be the same in which case

$$dW_t dW_t = dt \tag{36}$$

We are saying that the squared infinitesimal increments to a highly stochastic process are equal to a deterministic time increment. Can this really be?

In a strict mathematical sense it is only true in the integral sense. However, the relation (36) can be used informally as long as we know to use it only when the squared increments dW_t appear in the integrand of a stochastic integral satisfying certain conditions.

Quadratic variation of Brownian motion

Using similar arguments, we can say that

$$\int_0^T dW_t dt = \mathrm{QV}_T(W_t, t) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} \left(W_{t_{j+1}} - W_{t_j}\right)\left(t_{j+1} - t_j\right) = 0 \tag{37}$$

and that

$$\int_0^T (dt)^2 = \mathrm{QV}_T(t, t) = \lim_{\|\Pi\| \to 0} \sum_{j=0}^{n-1} (t_{j+1} - t_j)^2 = 0 \tag{38}$$

We thus have the following informal table to be used 'in the integral sense'

$$dW_t dW_t = dt, \quad dW_t dt = 0, \quad dtdt = 0 \tag{39}$$

Properties of Brownian motion

We can now conclude about Brownian motion and say that W_t, for $t \geq 0$

1) is a continuous time stochastic process with continuous paths,
2) has stationary independent increments,
3) has normally distributed increments, $W_t - W_s \sim N(0, t-s)$
4) has infinite first order variation,
5) has quadratic variation the grows linearly in time, $\int_0^T dW_t^2 = T$,
6) is a martingale,
7) is a Markov process.

Ito processes

Definition (Ito process)

Let W_t, $t \geq 0$ be a Brownian motion and let $\mathcal{F}_t$ be the filtration associated with W_t. An Ito process is a stochastic process of the form

$$X_T = x_0 + \int_0^T \mu_t dt + \int_0^T \sigma_t dW_t \tag{40}$$

where x_0 is non-random, and μ_t and σ_t are adapted processes. The dynamics of the Ito process can equivalently be written

$$\begin{aligned} dX_t &= \mu_t dt + \sigma_t dW_t, \\ X_0 &= x_0 \end{aligned} \tag{41}$$

Existence and uniqueness of the solution to an SDE

The Stochastic Differential equation (SDE) given in (41) will have a unique solution provided some conditions are met.

These conditions will generally be met in this course and you need not worry too much about them.

In particular, the SDE in (41) will have a solution that is unique, continuous in t and adapted, provided that the drift μ_t and the diffusion σ_t follow adapted, continuous, square-integrable stochastic processes.

Ito calculus

We have now defined an Ito process or equivalently an SDE driven by Brownian motion.

Very often in this course, we will work with some 'object' that is a sufficiently differentiable function $f = f(t, X)$ of time and the Ito process.

The 'object' could be the value of a complex derivative with maturity T.

By the Fundamental Theorem of Calculus, can write

$$f(T, X_T) = f(0, X_0) + \int_0^T df(t, X_t) \tag{42}$$

where the integrand $df(t, X_t)$ is the dynamics of $f(t, X)$.

To make sense of the integral and the integrand, we will need Ito calculus.

Ito calculus

Before we handle the case where X_t follows a stochastic process, let us remember what traditional calculus would imply in the simpler case when $X(t)$ is a nice sufficiently differentiable function of t.

If we use a Taylor expansion, we can write the differential of $f(t, X(t))$ as

$$df(t, X(t)) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX(t) + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x} dt dX(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX(t))^2 + \dots.$$

Integrating from $t = 0$ to $t = T$ will then allow us to write something like

$$\begin{aligned} f(T, X(T)) = f(0, X(0)) + \int_0^T \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX(t) \\ + \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x} dt dX(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX(t))^2 + \dots \end{aligned} \tag{43}$$

Ito calculus

Integrating each of the terms on the right hand side of (43), traditional calculus dictates that *all* higher order terms vanish and we are left with

$$f(T, X(T)) = f(0, X(0)) + \int_0^T \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX(t) \tag{44}$$

This is true for all $T > 0$ and taking a limit as $T \nearrow 0$ allows us to recover the so called 'Total differential' as the sum of partial derivatives

$$df(t, X(t)) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX(t) \tag{45}$$

Then, if we were to take this a step further, the chain rule would give us

$$df(t, X(t)) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} \frac{\partial X}{\partial t} dt \tag{46}$$

Ito calculus

Now, let us use similar arguments as before for the case when X_t follows a stochastic process. The Taylor series expansion, again assuming f is sufficiently differentiable, then allows us to write the following expression for the differential $df(t, X_t)$

$$df(t, X_t) = \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} (dt)^2 + \frac{\partial^2 f}{\partial t \partial x} dt dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX_t)^2 + ... \quad (47)$$

We can as before integrate from $t = 0$ to $t = T$ and write

$$\begin{aligned} f(T, X_T) = f(0, X_0) + \int_0^T \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t \\ + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} (dt)^2 + \frac{\partial^2 f}{\partial t \partial x} dt dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX_t)^2 + ... \end{aligned} \quad (48)$$

In the following, we will look at the terms on the RHS one by one and as we will see, most of the higher order terms will vanish but not all!

One of them will remain and be of an *unusual* form.

Ito calculus

The first term in (48)

$$\int_0^T \frac{\partial f}{\partial t} dt$$

is a regular time integral and will pose no challenge.

The send term becomes

$$\int_0^T \frac{\partial f}{\partial x} (\mu_t dt + \sigma_t dW_t)$$

The time integral will once again not be a problem, but the integral with respect to Brownian motion will.

Ito integral

We will look at the Ito integral for a general integrand σ_t given by

$$I(T) = \int_0^T \sigma_t dW_t \tag{49}$$

where σ_t is an adapted stochastic process.

It would be tempting to define the integral $I(T)$ using ordinary calculus and the chain rule to give us

$$I(T) = \int_0^T \sigma_t \left(\frac{dW}{dt} \right) dt \tag{50}$$

but this will *not* work since Brownian motion is not differentiable.

Ito integral

To make sense of the Ito integral $I(T)$ as a function of T, we will use a partition Π of $[0, T]$

$$\Pi = \{t_0, t_1, ..., t_n\}, \quad 0 = t_0 \leq t_1 \leq ... \leq t_n = T \tag{51}$$

and approximate the integral $I(T)$ using left endpoints.

The approximation $\hat{I}_n(T)$ will be

$$\hat{I}_n(T) = \sum_{j=0}^{n-1} \sigma_{t_j} \left[W_{t_{j+1}} - W_{t_j}\right] \tag{52}$$

Ito integral

In the approximation we are considering, we approximate the continuously varying process σ_t, which serves as the integrand in the Ito integral, with a simple process $\sigma_t^{(n)}$ that is piecewise constant.

Now, the trick will be to choose a sequence of partitions such that $\sigma_t^{(n)}$ converges to σ_t in a sensible way by which we mean

$$\lim_{n \to \infty} \mathrm{E} \int_0^T \left| \sigma_t^{(n)} - \sigma_t \right|^2 dt = 0 \tag{53}$$

This will require some heavy lifting that we will not do here.

We will be content to know that the limit exists and that the Ito integral is well-defined.

Ito integral

Definition (Ito integral)

The Ito integral will be defined as

$$\int_0^T \sigma_t dW_t = \lim_{n\to\infty} \int_0^T \sigma_t^{(n)} dW_t \tag{54}$$

where $\sigma_t^{(n)}$ *is a sequence of simple functions that converges to the continuously varying* σ_t *in the sense that*

$$\lim_{n\to\infty} \mathrm{E} \int_0^T \left|\sigma_t^{(n)} - \sigma_t\right|^2 dt = 0 \tag{55}$$

Continuity, adaptivity and linearity of the Ito integral

The Ito integral is continuous in its upper limit which follows from the fact that W_t is a continuous process and that σ_t is adapted.

For each $T \geq 0$, the Ito integral $I(T)$ is adapted which follows from the fact that both W_t and σ_t follow adapted stochastic processes.

The Ito integral is linear in the sense that if $I(T) = \int_0^T \sigma_t dW_t$ and $J(T) = \int_0^T \phi_t dW_t$ are Ito integrals and c a constant, then

$$
\begin{aligned}
I(T) + J(T) &= \int_0^T \sigma_t dW_t + \int_0^T \phi_t dW_t = \int_0^T (\sigma_t + \phi_t) dW_t \\
cI(T) &= \int_0^T c\sigma_t dW_t
\end{aligned}
\tag{56}
$$

which follows from the definition of the Ito integral.

The Ito integral is a martingale

The Ito integral is a martingale and in particular

$$\mathrm{E}\big[I(T)\big] = \mathrm{E}\big[I(T)\big|\mathcal{F}_0\big] = 0 \tag{57}$$

To see that the Ito integral is a martingale, assume $0 \leq S \leq T$ we look at

$$\begin{aligned}\mathrm{E}\big[I(T)\big|\mathcal{F}_S\big] =& \mathrm{E}\Big[\int_0^T \sigma_t dW_t \Big|\mathcal{F}_S\Big] = \mathrm{E}\Big[\int_0^S \sigma_t dW_t \Big|\mathcal{F}_S\Big] + \mathrm{E}\Big[\int_S^T \sigma_t dW_t \Big|\mathcal{F}_S\Big] \\ =& \int_0^S \sigma_t dW_t + \int_S^T \sigma_t \mathrm{E}\big[dW_t\big|\mathcal{F}_S\big] = \int_0^S \sigma_t dW_t = I(S)\end{aligned} \tag{58}$$

where we used the definition of the integral to split at s, the fact that the integral is adapted and the fact that future increments of Brownian motion are independent of the past and have mean zero.

Ito isometry

Theorem (Ito isometry)

The variance of the Ito integral is for $T > 0$ given by

$$\mathrm{Var}\left[I(T)\right] = \mathrm{E}\left[I^2(T)\right] = \int_0^T \sigma_t^2 dt \tag{59}$$

To justify Ito Isometry, we will use a heroic but strictly speaking not completely rigorous argument.

$$\begin{aligned} I^2(T) &= \left(\int_0^T \sigma_t dW_t\right)^2 = \int_0^T \int_0^T \sigma_u \sigma_s dW_u dW_s \\ &= \int_0^T \int_s^T \sigma_u \sigma_s dW_u dW_s + \int_0^T \int_0^s \sigma_u \sigma_s dW_u dW_s + \int_0^T \sigma_s^2 (dW_s)^2 \end{aligned} \tag{60}$$

Ito isometry

Notice that in the first integral, we have $u > s$ and in the second we have $s > u$. Taking expectations and conditioning gives us

$$
\begin{aligned}
\mathrm{E}\big[I^2(T)\big] &= \mathrm{E}\Big[\mathrm{E}\big[\int_0^T \int_s^T \sigma_u\sigma_s dW_u dW_s \big| \mathcal{F}_s\big]\Big] + \mathrm{E}\Big[\mathrm{E}\big[\int_0^T \int_0^s \sigma_u\sigma_s dW_u dW_s \big| \mathcal{F}_u\big]\Big] \\
&\quad + \mathrm{E}\Big[\int_0^T \sigma_s^2 (dW_s)^2\Big] \\
&= \mathrm{E}\Big[\int_0^T \int_s^T \sigma_s dW_s \mathrm{E}\big[\sigma_u dW_u \big| \mathcal{F}_s\big]\Big] + \mathrm{E}\Big[\int_0^T \int_0^s \sigma_u dW_u \mathrm{E}\big[\sigma_s dW_s \big| \mathcal{F}_u\big]\Big] \\
&\quad + \int_0^T \sigma_t^2 dt = \int_0^T \sigma_t^2 dt
\end{aligned}
\tag{61}
$$

Quadratic variation of the Ito integral

Theorem (Quadratic variation of the Ito integral)

The quadratic variation of the Ito integral up to time T will be

$$\mathrm{QV}_T\big(I(t)\big) = \int_0^T \sigma_t^2 dt \tag{62}$$

Again, we will give a very simple but not entirely rigorous argument using that informally $dW_t dW_t = dt$

$$\mathrm{QV}_T\big(I(t)\big) = \lim_{\|\Pi\|\to 0} \sum_{j=0}^{n-1} \sigma_{t_j}^2 \big(W_{t_{j+1}} - W_{t_j}\big) = \int_0^T \sigma_t^2 (dW_t)^2 = \int_0^T \sigma_t^2 dt \tag{63}$$

Ito calculus

We will now return to expression from (48) that we are trying to make sense of restated below

$$f(T, X_T) = f(0, X_0) + \int_0^T \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t$$
$$+ \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2 + \frac{\partial^2 f}{\partial t \partial x} dX_t dt + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 + \ldots$$

We have now taken care of the two first integrals involving dt and dX_t and in particular discussed that

$$\int_0^T \frac{\partial f}{\partial x} dX_t = \int_0^T \frac{\partial f}{\partial x}\mu_t dt + \int_0^T \frac{\partial f}{\partial x}\sigma_t dW_t \tag{64}$$

is well-defined.

Ito calculus

The third integral involving $(dt)^2$

$$\int_0^T \frac{1}{2}\frac{\partial^2 f}{\partial t^2}(dt)^2$$

will, as it turns out, vanish.

We will not give a rigorous argument but simply state that it is true appealing to intuition from ordinary calculus and also recalling our result from earlier that

$$\int_0^T (dt)^2 = 0$$

The fourth integral

$$\int_0^T \frac{\partial^2 f}{\partial t \partial x} dX_t dt = \int_0^T \frac{\partial^2 f}{\partial t \partial x} \mu_t (dt)^2 + \int_0^T \frac{\partial^2 f}{\partial t \partial x} \sigma_t dW_t dt \tag{65}$$

will also vanish and again, we will not give a rigourous argument but remember our informal argument from earlier that

$$\int_0^T dW_t dt = 0 \tag{66}$$

Ito calculus

The fifth integral

$$\int_0^T \frac{\partial^2 f}{\partial x^2}(dX_t)^2 = \int_0^T \frac{\partial^2 f}{\partial x^2}\mu_t(dt)^2 + \int_0^T \frac{\partial^2 f}{\partial x^2}\mu_t\sigma_t dW_t dt + \int_0^T \frac{\partial^2 f}{\partial x^2}\sigma_t^2(dW_t)^2 \quad (67)$$

is the truly interesting one.

The first two integrals in (67) will vanish, but the third will give us

$$\int_0^T \frac{\partial^2 f}{\partial x^2}\sigma_t^2(dW_t)^2 = \int_0^T \frac{\partial^2 f}{\partial x^2}\sigma_t^2 dt \quad (68)$$

where, for intuition, we appeal to the argument given when finding the quadratic variation of the Ito process and also that

$$\int_0^T (dW_t)^2 = \int_0^T dt \quad (69)$$

Ito calculus

We have now handled the first and second order terms of (48) and what remains is to deal with integrals of the form

$$\int_0^T \frac{\partial^{m+n} f}{\partial t^m \partial x^n} (dt)^m (dX_t)^n \tag{70}$$

where $m + n > 2$. However, by appealing to the informal table in (39) and restated below

$$dW_t dW_t = dt, \quad dW_t dt = 0, \quad dtdt = 0$$

we conclude that all higher order terms vanish as well.

We are thus in position to give Ito's formula.

Theorem (Ito's formula)

Let X_t be an Ito process with dynamics

$$dX_t = \mu_t dt + \sigma_t dW_t$$

where μ_t and σ_t follow adapted stochastic processes and let $f(t, x)$ be a function for which the partial derivatives $\frac{\partial f}{\partial t}, \frac{\partial f}{\partial x}$ and $\frac{\partial^2 f}{\partial x^2}$ are well-defined and continuous. Then for every $T \geq 0$, we have

$$\begin{aligned} f(T, X_T) &= f(0, X_0) + \int_0^T \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX_t)^2 \\ &= f(0, X_0) + \int_0^T \frac{\partial f}{\partial t} dt + \int_0^T \frac{\partial f}{\partial x} \mu_t dt + \int_0^T \frac{\partial f}{\partial x} \sigma_t dWt + \frac{1}{2} \int_0^T \frac{\partial^2 f}{\partial x^2} \sigma_t^2 dt \end{aligned}$$

Ito's formula

In differential form, Ito's formula becomes

$$
\begin{aligned}
df(t, X_t) &= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} (dX_t)^2 \\
&= \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} \mu_t dt + \frac{\partial f}{\partial x} \sigma_t dWt + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} \sigma_t^2 dt,
\end{aligned}
\tag{71}
$$

where we have once again used the informal rules saying that

$$
dW_t dW_t = dt, \qquad dW_t dt = 0, \qquad dtdt = 0.
$$

Ito's formula

To illustrate the difference between Ito calculus and regular calculus, we will compute the following integral

$$\int_0^T W_t dW_t. \tag{72}$$

Let us for a moment consider what we would get if Brownian motion could be differentiated. Suppose $X(t)$ is differentiable and that $X(0) = 0$, then

$$\int_0^T X(t)dX(t) = \int_0^T X(t)\frac{\partial X}{\partial t}dt = \frac{1}{2}\left[X^2(t)\right]_0^T = \frac{1}{2}X^2(T). \tag{73}$$

Ito's formula

To compute $\int_0^T W_t dW_t$, we apply Ito to $f(t,x) = x^2$ with $dX_t = dW_t$

$$d(W_t^2) = 2W_t dW_t + (dW_t)^2 = 2W_t dW_t + dt. \tag{74}$$

Integration from 0 to T gives us that

$$W_T^2 - W_0^2 = \int_0^T d(W_t^2) = 2\int_0^T W_t dW_t + T, \tag{75}$$

and we get that

$$\int_0^T W_t dW_t = \frac{1}{2}W_T^2 - \frac{1}{2}T. \tag{76}$$

Theorem (Ito's formula in the bivariate case)

Let X_t and Y_t be an Ito processes each with drift and diffusion following adapted processes and let $f(t,x,y)$ be a function for which the partial derivatives $\frac{\partial f}{\partial t}$, $\frac{\partial f}{\partial x}$, $\frac{\partial^2 f}{\partial x^2}$, $\frac{\partial f}{\partial y}$, $\frac{\partial^2 f}{\partial y^2}$ and $\frac{\partial^2 f}{\partial y \partial x}$ are well-defined and continuous. Then for every $T \geq 0$, we have

$$f(T, X_T, Y_T) = f(0, X_0, Y_0) + \int_0^T \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{\partial f}{\partial y} dY_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 + \frac{\partial^2 f}{\partial x \partial y} dX_t dY_t + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(dY_t)^2.$$

In differential form, we have

$$df(t, X_t, Y_t) = \frac{\partial f}{\partial x} dX_t + \frac{\partial f}{\partial y} dY_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 + \frac{\partial^2 f}{\partial x \partial y} dX_t dY_t + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(dY_t)^2.$$

Geometric Brownian motion

The stochastic process S_t follows a Geometric Brownian motion and has dynamics given by

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t, \\ S_0 &= s_0, \end{aligned} \tag{77}$$

where μ and σ are constant.

We can of course always integrate S_t straight away to get

$$S_T = s_0 + \int_0^T \mu S_t dt + \int_0^T \sigma S_t dW_t, \tag{78}$$

But, we still have S_t on the right hand side and though true, this does not represent a solution.

However, a very neat solution can indeed be found.

Geometric Brownian motion

Let us define X_t as $X_t = \log(S_t)$ and find the dynamics of X_t.

Hence, let us find $dX_t = df(t,S) = d\ln(S_t)$ by applying Ito's formula where $f(t,S) = \log S$ and S_t is a Geometric Brownian motion

$$\begin{aligned} dX_t &= \frac{\partial f}{\partial S} dS_t + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}(dS_t)^2 = \frac{1}{S_t}(\mu S_t dt + \sigma S_t dW_t) - \frac{1}{2}\frac{1}{S_t^2}(\sigma S_t dW_t)^2 \\ &= (\mu - \frac{1}{2}\sigma^2)dt + \sigma dW_t. \end{aligned} \tag{79}$$

Now, there is no X_t on the right hand side and we can integrate to get

$$X(T) = x_0 + \int_0^T (\mu - \frac{1}{2}\sigma^2)dt + \int_0^T \sigma dW_t = x_0 + (\mu - \frac{1}{2}\sigma^2)T + \sigma W_T. \tag{80}$$

This is a solution for X_T and the solution for S_T becomes

$$S(T) = s_0 \exp\left([\mu - \frac{1}{2}\sigma^2]T + \sigma W_T\right). \tag{81}$$

From the solution for X_T, we see that

$$X_T \sim N\Big(x_0 + [\mu - \frac{1}{2}\sigma^2]\, T, \sigma^2 T\Big), \tag{82}$$

and it follows that S_T follows a log normal distribution. To find the expected value of S_T, we look at the moment generating function $M_{W_T}(\omega)$ of W_T.

$$M_{W_T}(\omega) = \mathrm{E}\big[e^{\omega W_T}\big], \tag{83}$$

where ω is the argument of the MGF and note that

$$\mathrm{E}\big[W_T\big] = \frac{\partial}{\partial\omega} M_{W_T}(\omega)\Big|_{\omega=0}, \quad \mathrm{E}\big[W_T^2\big] = \frac{\partial^2}{\partial\omega^2} M_{W_T}(\omega)\Big|_{\omega=0}, \quad \mathrm{E}\big[W_T^k\big] = \frac{\partial^k}{\partial\omega^k} M_{W_T}(\omega)\Big|_{\omega=0} \tag{84}$$

Geometric Brownian motion

Now, we know that W_T follows a normal distribution and recall that if $X \sim N(\mu, \sigma^2)$, then it's moments generating function is given by

$$M_X(\omega) = \mathrm{E}\left[e^{\omega X}\right] = e^{u\omega + \frac{1}{2}\sigma^2\omega^2}. \tag{85}$$

The expected value of S_T thus becomes

$$\begin{aligned} \mathrm{E}\left[S_T\right] &= \mathrm{E}\left[s_0 \exp\left([\mu - \frac{1}{2}\sigma^2]T + \sigma W_T\right)\right] = s_0 \exp\left([\mu - \frac{1}{2}\sigma^2]T\right)\mathrm{E}\left[e^{\sigma W_T}\right] \\ &= s_0 \exp\left([\mu - \frac{1}{2}\sigma^2]T\right) \exp\left(\frac{1}{2}\sigma^2 T\right) = s_0 e^{\mu T}. \end{aligned} \tag{86}$$

Simulation of an Ito Process

Later when pricing more complicated derivatives, we will need to rely on simulation of the stochastic process or processes of the underlying asset. In order to do so with reasonable accuracy, we will to develop a range of methods.

Suppose we wish to simulate the stochastic process S_t, $t \geq 0$ on the time interval from $t = 0$ to $t = T$. Furthermore, assume that we will do so by taking N steps on a discrete grid with uniform mesh to cover the range from 0 to T. The relation between the range in time T, the number of steps N and the mesh δ is thus

$$\delta = \frac{T}{N}. \tag{87}$$

The times of the steps will be denoted by t

$$t \in \left[0, \delta, ..., n\delta, ..., N\delta = T\right] \tag{88}$$

Simulation of an Ito Process - Euler discretization

The Euler discretization is the simplest method and assumes allows for simulation of a stochastic process S_t driven by Brownian motion where the coefficients are allowed to depend on both time t and the value of the process at time t. The dynamics of S_t are of the form

$$dS_t = \mu(t, S_t)dt + \sigma(t, S_t)dW_t. \tag{89}$$

Integrating from t to $t + \delta$ gives us that

$$S_{t+\delta} = S_t + \int_t^{t+\delta} \mu(u, S_u)du + \int_t^{t+\delta} \sigma(u, S_u)dW_u. \tag{90}$$

Simulation of an Ito Process - Euler discretization

Using the left end-point, we can approximate the integrals in (90) by

$$\int_t^{t+\delta} \mu(u, S_u)du \approx \mu(t, S_t) \int_t^{t+\delta} du = \mu(t, S_t)\delta. \tag{91}$$

and by

$$\int_t^{t+\delta} \sigma(u, S_u)dW_u \approx \sigma(t, S_t) \int_t^{t+\delta} dW_u = \sigma(t, S_t)\big(W_{t+\delta} - W_t\big) = \sigma(t, S_t)\sqrt{\delta}Z_t. \tag{92}$$

where $Z_t \sim N(0, 1)$ is a standard normal random variable. The Euler discretization then becomes

$$S_{t+\delta} \approx S_t + \delta\mu(t, S_t) + \sigma(t, S_t)\sqrt{\delta}Z_t. \tag{93}$$

Simulation of an Ito Process - the Milstein scheme

The Milstein scheme is an improvement of the Euler discretization but, at least in the version we will discuss here, only applies if the coefficients of the stochastic process only depend on S_t and *not* on t. The dynamics of S_t are in other words of the form

$$dS_t = \mu(S_t)dt + \sigma(S_t)dW_t = \mu_t dt + \sigma_t dW_t. \tag{94}$$

Integrating from t to $t+\delta$ gives us that

$$S_{t+\delta} = S_t + \int_t^{t+\delta} \mu_u du + \int_t^{t+\delta} \sigma_u dW_u \tag{95}$$

Next, we will proceed towards expressing μ_u and σ_u in (95) as integrals themselves in terms of the differentials $d\mu_t$ and $d\sigma_t$ using Ito's formula.

Simulation of an Ito Process

Using the notation that

$$\mu_t' = \frac{\partial \mu(S_t)}{\partial S_t}, \qquad \mu_t'' = \frac{\partial^2 \mu(S_t)}{\partial S_t^2}, \qquad \sigma_t' = \frac{\partial \sigma(S_t)}{\partial S_t}, \qquad \sigma_t'' = \frac{\partial^2 \sigma(S_t)}{\partial S_t^2}. \tag{96}$$

We can then write the differentials $d\mu_t$ and $d\sigma_t$ using Ito's formula as

$$d\mu_t = \mu_t' dS_t + \frac{1}{2}\mu_t'' \left(dS_t\right)^2 = \mu_t' \mu_t dt + \mu_t' \sigma_t dW_t + \frac{1}{2}\mu_t'' \sigma_t^2 dt, \tag{97}$$

and as

$$d\sigma_t = \sigma_t' dS_t + \frac{1}{2}\sigma_t'' \left(dS_t\right)^2 = \sigma_t' \mu_t dt + \sigma_t' \sigma_t dW_t + \frac{1}{2}\sigma_t'' \sigma_t^2 dt, \tag{98}$$

Simulation of an Ito Process - the Milstein scheme

Using the differentials for $d\mu_t$ and $d\sigma_t$, we can express μ_u and σ_u from (95) as shown in the following below

$$\mu_u = \mu_t + \int_t^u d\mu_v = \mu_t + \int_t^u \mu'_v \mu_v dv + \mu'_v \sigma_v dW_v + \frac{1}{2}\mu''_v \sigma_v^2 dv, \tag{99}$$

and the similarly

$$\sigma_u = \sigma_t + \int_t^u d\sigma_v = \sigma_t + \int_t^u \sigma'_v \mu_v dv + \sigma'_v \sigma_v dW_v + \frac{1}{2}\sigma''_v \sigma_v^2 dv \tag{100}$$

Substituting back into (95) and rearranging slightly gives us

$$\begin{aligned} S_{t+\delta} = S_t &+ \int_t^{t+\delta} \left(\mu_t + \int_t^u \left(\mu'_v \mu_v dv + \frac{1}{2}\mu''_v \sigma_v^2 dv\right)\right) du + \int_t^{t+\delta}\int_t^u \mu'_v \sigma_v dW_v du \\ &+ \int_t^{t+\delta} \left(\sigma_t + \int_t^u \left(\sigma'_v \mu_v dv + \frac{1}{2}\sigma''_v \sigma_v^2 dv\right)\right) dW_u + \int_t^{t+\delta}\int_t^u \sigma'_v \sigma_v dW_v dW_u \end{aligned} \tag{101}$$

Simulation of an Ito Process - the Milstein scheme

Now, terms involving $dvdu$ are of order $\mathcal{O}(\delta^2)$, terms involving $dW_v du$ or $dvdW_u$ are of order $\mathcal{O}(\delta^{3/2})$ but terms involving $dW_v dW_u$ are of order $\mathcal{O}(\delta)$.

Retaining only terms of order $\mathcal{O}(\delta)$ or less gives the following approximation

$$S_{t+\delta} \approx S_t + \mu_t \int_t^{t+\delta} du + \sigma_t \int_t^{t+\delta} dW_u + \int_t^{t+\delta} \int_t^u \sigma'_v \sigma_v dW_v dW_u \tag{102}$$

The only thing that now remains is to look at the integral in (102). First, we will approximate using left endpoints to give us that

$$\begin{aligned} I &= \int_t^{t+\delta} \int_t^u \sigma'_v \sigma_v dW_v dW_u \approx \sigma'_t \sigma_t \int_t^{t+\delta} \int_t^u dW_v dW_u = \sigma'_t \sigma_t \int_t^{t+\delta} (W_u - W_t) dW_u \\ &= \sigma'_t \sigma_t \int_t^{t+\delta} W_u dW_u - \sigma'_t \sigma_t W_t \int_t^{t+\delta} dW_u \end{aligned} \tag{103}$$

Simulation of an Ito Process - the Milstein scheme

To proceed, let us recall the result from earlier that

$$\int_0^t W_u dW_u = \frac{1}{2} W_t^2 - \frac{1}{2} t. \tag{104}$$

To give us that

$$\int_t^{t+\delta} W_u dW_u = \frac{1}{2} W_{t+\delta}^2 - \frac{1}{2}(t+\delta) - \frac{1}{2} W_t^2 + \frac{1}{2} t = \frac{1}{2}\left(W_{t+\delta}^2 - W_t^2 - \delta\right) \tag{105}$$

Inserting these results into (103) gives us

$$I = \sigma_t' \sigma_t \int_t^{t+\delta} W_u dW_u - \sigma_t' \sigma_t W_t \int_t^{t+\delta} dW_u = \frac{1}{2} \sigma_t' \sigma_t \left[\left(W_{t+\delta} - W_t\right)^2 - \delta \right] \tag{106}$$

Simulation of an Ito Process - the Milstein scheme

The Milstein discretization then looks as follows

$$S_{t+\delta} \approx S_t + \mu_t \int_t^{t+\delta} du + \sigma_t \int_t^{t+\delta} dW_u + \frac{1}{2}\sigma_t' \sigma_t \left[\left(W_{t+\delta} - W_t \right)^2 - \delta \right] \tag{107}$$

When simulating using this method, we will for each time step draw an standard normal random variable $Z_t \sim N(0,1)$ and compute the value corresponding to the next step as

$$S_{t+\delta} = S_t + \mu_t \delta + \sigma_t \sqrt{\delta} Z_t + \frac{1}{2}\sigma_t' \sigma_t \delta \left(Z_t - 1 \right) \tag{108}$$

The Black-Scholes model

We need to be able to price complex derivatives and in order to do that, we need to understand how these derivatives can be priced in a complete market where it is possible to replicate the derivative.

We will begin by looking at the simplest case where we have one risk free asset - the bank account, and one risky asset. We will refer to the underlying asset as 'the stock' but it could also be a number of other assets, a bond for example.

In addition, we will have a European derivative and assume that the only stochastic component affecting the value of the derivative is the stock price. Recall that for European call- and put options, the value of the derivative at maturity is determined solely by the value of the underlying stock.

The Black-Scholes model

The dynamics of the bank account will be of the form

$$dB_t = rB_t dt, \tag{109}$$

where r is a constant. The short rate is thus *not* stochastic but known in advance for all future points in time. The case of a time-varying but deterministic short rate is equivalent as we will see later.

The underlying stock has price dynamics under the objective measure $\mathbb{P}$

$$dS_t = \mu(t, S_t)S_t dt + \sigma(t, S_t)S_t dW_t^{\mathbb{P}}. \tag{110}$$

where μ and $\sigma(t, S_t)$ are known functions of time and the stock price.

The Black-Scholes model

The derivative will as mentioned be of European type implying that the derivative only pays one cash flow at maturity T.

The cashflow, χ, paid at maturity is assumed to be a function Φ of the stock price at maturity

$$\chi = \Phi(S_T). \tag{111}$$

Furthermore, we assume that the derivative can be traded in the market and that its price process Π_t is a function F of time t and value of the stock, S_t, at time t

$$\Pi_t = F(t, S_t). \tag{112}$$

The Black-Scholes model

Recall that in order to be sure we can replicate a contingent claim, we will need the market to be complete in the appropriate sense.

What exactly that means in general, we will discuss later but lets note that we only have one stochastic component in the model, the Brownian motion W_t, and one traded underlying asset, the stock. Now, we will:

1) Construct a self-financing portfolio consisting of the stock and the risky derivative in such a way that it has no stochastic component.

2) Argue that since the portfolio has no stochastic component, it must have the risk free rate as its infinitesimal.

3) Impose on the self-financing portfolio that it must have the risk free rate as its return and from that condition, find an equation for the value of the derivative.

Let us denote derivatives of the function F with subscripts $\frac{\partial F}{\partial t} = F_t$ and find the dynamics, $d\Pi_t$, of the derivative using Ito's formula

$$\begin{aligned} d\Pi_t &= F_t dt + F_s dS_t + \frac{1}{2} F_{ss} (dS_t)^2 = F_t dt + \mu S_t F_s dt + \frac{1}{2}\sigma^2 S_t^2 F_{ss} dt + \sigma S_t F_s dW_t^{\mathbb{P}} \\ &= \mu_\pi \Pi_t dt + \sigma_\pi \Pi dW_t^{\mathbb{P}}, \end{aligned} \tag{113}$$

where

$$\mu_\pi = \frac{F_t + \mu S_t F_s + \frac{1}{2}\sigma^2 S_t^2 F_{ss}}{F}, \qquad \sigma_\pi = \frac{\sigma S_t F_s}{F}. \tag{114}$$

The replicating portfolio will consist of h_s stocks and h_π derivative contracts. The value of the replicating portfolio will be V_t where

$$V_t = h_s S_t + h_\pi \Pi_t = V_t \left(\frac{h_s S_t}{V_t} + \frac{h_\pi \Pi_t}{V_t} \right) = V_t (u_s + u_\pi). \tag{115}$$

Here, u_s is the fraction of the portfolio invested in the stock and u_π the fraction of V_t invested in the derivative such that $u_s + u_\pi = 1$.

The Black-Scholes model

The dynamics of the replicating portfolio then becomes

$$
\begin{aligned}
dV_t &= h_s dS_t + h_\pi d\Pi_t = h_s\left(\mu S_t dt + \sigma S_t dW_t^{\mathbb{P}}\right) + h_\pi\left(\mu_\pi \Pi_t dt + \sigma_\pi \Pi dW_t^{\mathbb{P}}\right) \\
&= V_t\left(u_s[\mu dt + \sigma dW_t^{\mathbb{P}}] + u_\pi[\mu_\pi dt + \sigma_\pi dW_t^{\mathbb{P}}]\right) = \\
&= V_t\left[u_s\mu + u_\pi\mu_\pi\right]dt + \left[u_s\sigma + u_\pi\sigma_\pi\right]dW_t^{\mathbb{P}}.
\end{aligned} \tag{116}
$$

From this expression, we can eliminate the stochastic element of dV_t. Also, we know that the portfolio weights u_s and u_π must sum to one and we get the following set of equations to solve

$$
\left.\begin{aligned}
u_s\sigma + u_\pi\sigma_\pi &= 0 \\
u_s + u_\pi &= 1
\end{aligned}\right\} \Rightarrow
\begin{aligned}
u_s &= \tfrac{\sigma_\pi}{\sigma_\pi - \sigma} = \tfrac{S_t F_s}{S_t F_s - F} \\
u_\pi &= \tfrac{-\sigma}{\sigma_\pi - \sigma} = \tfrac{-F}{S_t F_s - F}
\end{aligned}\;. \tag{117}
$$

The Black-Scholes model

Choosing u_s and u_π as above will eliminate the stochastic component from the replicating portfolio, and we are left with a portfolio that is risk-free.

In a complete market, any risk-free asset or portfolio must yield the risk free rate of return to prevent an arbitrage from occurring. Hence we must impose that

$$dV_t = V_t\left[u_s\mu + u_\pi\mu_\sigma\right]dt = rV_t dt. \tag{118}$$

Inserting the expressions for u_s and u_π from (117) into the no arbitrage condition in (118) gives us

$$\begin{aligned}
r &= \frac{S_t F_s}{S_t F_s - F}\mu + \frac{-F}{S_t F_s - F}\frac{F_t + \mu S_t F_s + \frac{1}{2}\sigma^2 S_t^2 F_{ss}}{F} \quad \Rightarrow \\
r\left(S_t F_s - F\right) &= \mu S_t F_s - F_t - \mu S_t F_s - \frac{1}{2}\sigma^2 S_t^2 F_{ss} \quad \Rightarrow \\
& F_t + rS_t F_s + \frac{1}{2}\sigma^2 S_t^2 F_{ss} - rF = 0. \qquad (119)
\end{aligned}$$

The Black-Scholes model

We have now seen that the price of a European derivative in this model must for $t < T$ solve the Black-Scholes equation (119).

At maturity when $t = T$, we have that the value of the derivative will be given by the contract function

$$\Pi(T) = F(T, S_T) = \Phi(S_T). \tag{120}$$

We have now derived one of the most important results in finance.

Theorem (Black-Scholes equation)

Assume the dynamics of the risk-free asset and the risky stock are

$$dB(t) = rB(t)dt,$$

$$dS_t = \mu(t, S_t)S_t dt + \sigma(t, S_t)S_t dW_t^{\mathbb{P}}.$$

where r *is constant, and* $\mu(t, S_t)$ *and* $\sigma(t, S_t)$ *are known functions. Assume that we want to price a European contingent claim with contract function*

$$\chi = \Phi(S_T).$$

Then, the only pricing function $F = F(t, S_t)$ *consistent with the absence of arbitrage solves the following boundary value problem on* $[0, T] \times R_+$

$$F_t + rS_t F_s + \frac{1}{2}\sigma^2 S_t^2 F_{ss} - rF = 0,$$

$$F(T, S_T) = \Phi(S_T). \quad (121)$$

The Black-Scholes equation

The pricing function $\Pi(t;\chi) = F(t, S_t)$ depends on the price of the underlying asset and hence, the derivative is priced relative to the price of the underlying asset.

The price of the derivative is to be understood as the only price consistent with that of the underlying asset.

The boundary value problem solved by the pricing function does not depend on the local rate of return and it does not matter if the underlying asset has a local rate of return of 20 % or 5%, the price of the derivative remains the same.

This might seem counterintuitive but stems from the fact that we are pricing the derivative in terms of the underlying asset.

Risk-neutral valuation

We now know that the pricing function $\Pi(t) = F(t, S_t)$ must satisfy Black-Scholes equation and solves a very specific boundary value problem. However, we have not yet solved this problem.

To solve the Black-Scholes equation, we note that the F_t, F_s and F_{ss} terms look like the terms you would get by applying Ito's formula to a function $F(t, S_t)$ where S_t has dynamics

$$dS_t = rS_t dt + \sigma(t, S_t) S_t dW_t^{\mathbb{Q}}. \tag{122}$$

However, we also need to produce a term $-rF$. This last term must come from the time derivative and we are therefore tempted to look at

$$H(t, S_t) = e^{-rt} F(t, S_t). \tag{123}$$

Risk-neutral valuation

The strategy we will pursue is to assume that $F(t, S_t)$ solves the Black-Scholes equation, that the stock has dynamics given by (122) and then apply Ito's formula to $H(t, S_t)$ given in (123). Hopefully, this procedure will allow us to characterize the solution to the Black-Scholes equation. Applying Ito gives us

$$d\left(e^{-rt}F\right) = -re^{-rt}Fdt + e^{-rt}F_t dt + e^{-rt}rS_t F_s dt + e^{-rt}\sigma S_t F_s dW_t + \frac{1}{2}e^{-rt}\sigma^2 S_t^2 F_{ss} dt.$$

Using that $F(t, S_t)$ solves the Black-Scholes equation leaves us with

$$d\left(e^{-rt}F\right) = e^{-rt}\sigma(t, S_t)S_t F_s dW_t^{\mathbb{Q}}.$$

Now integrating from t to T gives us

$$\int_t^T d\left(e^{-ru}F(u, S_u)\right) = e^{-rT}F(T, S_T) - e^{-rt}F(t, S_t) = \int_t^T e^{-ru}\sigma(u, S_u)S_u F_s(u, S_u)dW_u.$$

Risk-neutral valuation

We now have an expression for the time t price $\Pi(t) = F(t, S_t)$ of our European derivative

$$F(t, S_t) = e^{-r(T-t)}\Phi(S_T) - \int_t^T e^{r(t-u)}\sigma(u, S_u)S_u F_s(u, S_u)dW_u, \tag{124}$$

and we can also get rid of the Ito integral by taking expectations.

To emphasize that the expectation is to be taken under the assumption that the dynamics of the stock price has r as it's drift, we put a $\mathbb{Q}$ in the expectation operator

$$F(t, S_t) = e^{-r(T-t)}\mathrm{E}^{\mathbb{Q}}\big[\Phi(S_T)\big|\mathcal{F}_t\big]. \tag{125}$$

Risk-neutral valuation

Theorem (Risk-neutral valuation)

Assume the market consists of a risk-free asset with constant return r*, a risky stock,* S_t*, and a European derivative with contract function* $\chi = \Phi(S_T)$*. The stock and the derivative are both traded in the market and driven by the same Brownian motion. Then the arbitrage free price of the claim* $\Phi(S_T)$ *is given by* $\Pi_t = F(t, S_T)$ *where* $F(t, S_t)$ *is given by*

$$F(t, S_t) = e^{-r(T-t)} \mathrm{E}^{\mathbb{Q}}\left[\Phi(S_T) \middle| \mathcal{F}_t\right],$$

and the expectation is taken under the risk-neutral measure $\mathbb{Q}$ *where the dynamics of* S_t *are of the form*

$$dS_t = rS_t dt + \sigma(t, S_t) S_t dW_t^{\mathbb{Q}}.$$

Risk-neutral valuation

Let us summarize our findings and try interpret the results.

When constructing the Black-Scholes model, we assumed that

- There is a risk-free asset with constant return r.
- The risky stock has $\mathbb{P}$-dynamics $dS_t = \mu(t, S_t)S_t dt + \sigma(t, S_t)S_t dW_t^{\mathbb{P}}$
- There is a European derivative with contract function $\chi = \Phi(S_T)$ traded in the market and it has pricing function of the form $\Pi(t) = F(t, S_t)$.

We then showed that there exist a strategy that both replicates the European derivative but has no stochastic component.

Since this strategy is risk-less, we imposed that it must have the risk-free rate as its local rate of return.

Risk-neutral valuation

From the no-arbitrage condition we were able to deduce that the pricing function $\Pi(t) = F(t, S_t)$ solves a specific boundary value problem.

We were then able to recognize some of the terms of the PDE in the boundary value problem and could conclude that the solution to the boundary value problem can be found by taking an expectation *not* under the objective probability measure $\mathbb{P}$ but rather under the risk-neutral probability $\mathbb{Q}$ where the stock has dynamics

$$dS_t = rS_t dt + \sigma(t, S_t) S_t dW_t^{\mathbb{Q}}. \tag{126}$$

So what really happened? Why did we have to change the dynamics of the stock?

Risk-neutral valuation

What happened to the dynamics of stock is *not* that the drift changed but rather, we changed the probability measure under which we are computing prices.

This change was a direct consequence of the fact that we are operating in a complete market where the derivative can be replicated.

But how do asset prices move and relate under the risk-neutral measure?

As we will see, it turns out that all risky assets when scaled by the risk-free asset become martingales under the risk neutral measure.

The risk-neutral measure

Let us look at the dynamics of the risky assets under $\mathbb{Q}$ and in particular investigate if $\frac{S_t}{B_t}$ is a martingale under the risk neutral measure.

In particular, we will find the dynamics of $Y_t = \frac{S_t}{B_t}$

$$\begin{aligned} dY_t &= \frac{1}{B_t} dS_t - \frac{S_t}{B_t^2} dB_t + \frac{S_t}{B_t^3} (dB_t)^2 = \frac{1}{B_t} \left(rS_t dt + \sigma S_t dW_t^{\mathbb{Q}} \right) - \frac{S_t}{B_t^2} rB_t dt + \frac{S_t}{B_t^3} r^2 B_t^2 dt^2 \\ &= \frac{1}{B_t} \sigma S_t dW_t^{\mathbb{Q}}. \end{aligned} \tag{127}$$

Integrating for $T > t$ gives us that

$$Y_T = Y_t + \int_t^T \sigma S_u dW_u^{\mathbb{Q}}. \tag{128}$$

Hence

$$\mathrm{E}\left[Y_T \middle| \mathcal{F}_t \right] = Y_t \tag{129}$$

and we conclude that $Y_t = \frac{S_t}{B_t}$ is indeed a martingale.

The risk-neutral measure

Next, we look at whether the derivative is a martingale under $\mathbb{Q}$ and look at the ratio the ratio $\frac{\Pi_t}{B_t}$.

Since the risk-free rate is constant, we have for $T > t$ that

$$B_T = B_t e^{r(T-t)}. \tag{130}$$

Let us also recall that

$$\Pi_t = e^{-r(T-t)}\mathrm{E}\left[\Phi(S_T)\middle|\mathcal{F}_t\right] = e^{-r(T-t)}\mathrm{E}\left[\Pi_T\middle|\mathcal{F}_t\right] = \frac{B_t}{B_T}\mathrm{E}\left[\Pi_T\middle|\mathcal{F}_t\right], \tag{131}$$

and hence

$$\mathrm{E}\left[\frac{\Pi_T}{B_T}\middle|\mathcal{F}_t\right] = \frac{\Pi_t}{B_t}, \tag{132}$$

and indeed, $\frac{\Pi_t}{B_t}$ is a martingale.

The risk-neutral measure

Definition (The risk neutral measure)

The risk neutral measure $\mathbb{Q}$ *is such that*

i) *the sets of measure zero are as under the objective measure* $\mathbb{P}$,

ii) *all risky assets have the risk-free rate as their local rate of return,*

iii) *when scaled by the risk-free asset, all risky assets become martingales.*

Market Completeness

The derivation of the Black-Scholes equation depended crucially on the fact that we were able to replicate the derivative.

That is, we relied on the market being complete.

Hence, it seems we need an updated version of our definition of market completeness.

Having updated the notion of market completeness, it would seem natural to also revisit the first and second fundamental theorems of asset pricing.

Market Completeness

Definition (Market Completeness)

The financial market is complete if every contingent claim is reachable. That is, the market is complete if every contingent claim can be replicated using a portfolio consisting solely of traded assets.

Theorem (First Fundamental Theorem of Asset Pricing)

Assume the market consists of a risk-free asset B_t and N traded risky assets following stochastic processes $S_t^{(1)}, S_t^{(2)}, \ldots, S_t^{(N)}$. Then the market is arbitrage free if and only if there exists a measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ such that the processes

$$\frac{S_t^{(1)}}{B_t}, \frac{S_t^{(2)}}{B_t}, \ldots, \frac{S_t^{(N)}}{B_t}$$

are martingales under $\mathbb{Q}$.

Second Fundamental Theorem of Asset Pricing

Theorem (Second Fundamental Theorem of Asset Pricing)

Assume the market is complete. Then the market is arbitrage free if and only if the risk-neutral measure $\mathbb{Q}$ *is unique.*

Completeness and arbitrage in multivariate models

Theorem (Meta-theorem)

Let N denote the number of linearly independent assets traded in the market and let M denote the number of independent random sources (Brownian motions). Then we have the following heuristic relations.

1) *The market is complete if and only if $N \geq M$.*
2) *If $N = M$ the market is complete and arbitrage free.*
3) *The market is arbitrage free if $N \leq M$.*