Probability Distributions and Stochastic Processes

At the heart of modern finance theory lies the uncertain movement of financial quantities. For modelling purposes we are concerned with the evolution of random events through time. *Stochastic Processes* are a family of random variables parametrized by time. Financial assets can be considered as stochastic processes.

A *diffusion process* is one that is continuous in space, while a *random walk* is a process that is discrete. The random path followed by the process is called a realization. Hence when referring to the path traced out by a financial variable will be termed as an asset price realization.

The mathematics can be achieved by the concept of a transition density function and is the connection between probability theory and differential equations.

Starting with a binomial random walk which is discrete we can obtain a continuous time process to obtain a partial differential equation for the transition probability density function (i.e. a time dependent PDF).

A simple symmetric random walk models the dynamics of a random variable, with value y at time t. The probability of an up/down move is $\alpha = 1/2$.

The Transition Probability Density Function

The transition pdf is denoted by

$$p\left(y, t; y', t'\right)$$

We can gain information such as the centre of the distribution, where the random variable might be in the long run, etc. by studying its probabilistic properties. So the density of particles diffusing from (y, t) to (y', t').

Think of (y, t) as current (or backward) variables and (y', t') as futures ones.

The more basic assistance it gives is with

$$\mathbb{P}\left(a < y' < b \text{ at } t' \middle| \, y \text{ at } t\right) = \int_a^b p\left(y, t; y', t'\right) dy'$$

i.e. the probability that the random variable y' lies in the interval a and b,at a future time t', given it started out at time t with value y.

$p\left(y, t; y', t'\right)$ satisfies two equations:

Forward equation involving derivatives with respect to the future state (y', t'). Here (y, t) is a starting point and is 'fixed'.

Backward equation involving derivatives with respect to the current state (y, t). Here (y', t') is a future point and is 'fixed'. The backward equation tells us the probability that we were at (y, t) given that we are now at (y', t'), which is fixed.

The mathematics: Start out at a point (y, t). We want to answer the question, what is the probability density function of the position y' of the diffusion at a later time t'?

This is known as the **transition density function** written $p\left(y, t; y', t'\right)$ and represents the density of particles diffusing from (y, t) to (y', t').

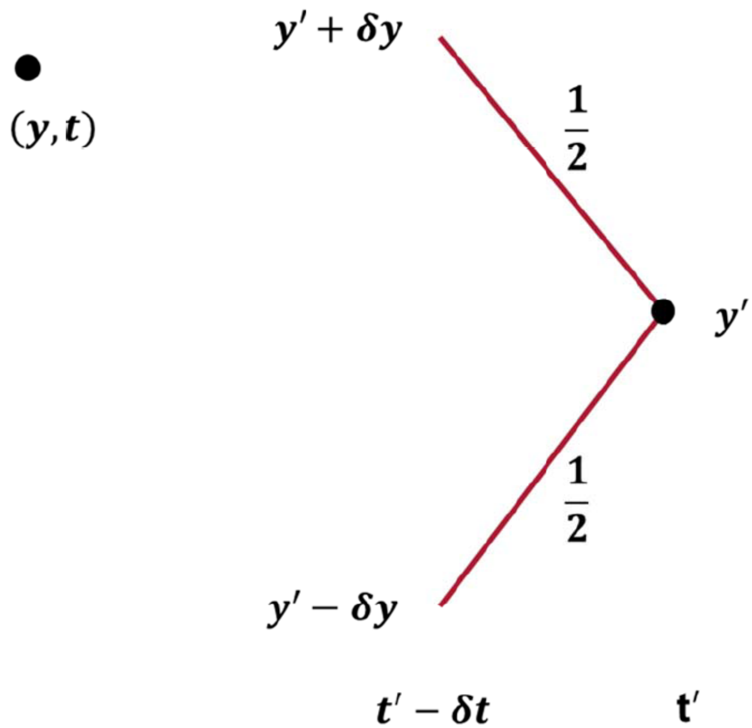
How can we find p?

Well there are a number of ways to do this. Start by considering the following (two step) binomial random walk. So the random variable can either rise or fall with equal probability.

y is the random variable and δt is a time step. δy is the size of the move in y.

$\mathbb{P}[\delta y] = \mathbb{P}[-\delta y] = 1/2.$

Suppose we are at (y', t'), how did we get there? At the previous step time step we must have been at one of $(y' + \delta y, t' - \delta t)$ or $(y' - \delta y, t' - \delta t)$.



So

$$p(y', t') = \tfrac{1}{2} p(y' + \delta y, t' - \delta t) + \tfrac{1}{2} p(y' - \delta y, t' - \delta t)$$

Taylor series expansion gives

$$\begin{aligned} p(y' + \delta y, t' - \delta t) &= p(y', t') - \frac{\partial p}{\partial t'}\delta t + \frac{\partial p}{\partial y'}\delta y + \tfrac{1}{2}\frac{\partial^2 p}{\partial y'^2}\delta y^2 + \ldots \\ p(y' - \delta y, t' - \delta t) &= p(y', t') - \frac{\partial p}{\partial t'}\delta t - \frac{\partial p}{\partial y'}\delta y + \tfrac{1}{2}\frac{\partial^2 p}{\partial y'^2}\delta y^2 + \ldots \end{aligned}$$

Substituting into the above

$$\begin{aligned} p(y', t') &= \tfrac{1}{2}\left(p(y', t') - \frac{\partial p}{\partial t'}\delta t + \frac{\partial p}{\partial y'}\delta y + \tfrac{1}{2}\frac{\partial^2 p}{\partial y'^2}\delta y^2\right) \\ &\quad + \tfrac{1}{2}\left(p(y', t') - \frac{\partial p}{\partial t'}\delta t - \frac{\partial p}{\partial y'}\delta y + \tfrac{1}{2}\frac{\partial^2 p}{\partial y'^2}\delta y^2\right) \end{aligned}$$

$$0 = -\frac{\partial p}{\partial t'}\delta t + \tfrac{1}{2}\frac{\partial^2 p}{\partial y'^2}\delta y^2$$

$$\frac{\partial p}{\partial t'} = \tfrac{1}{2}\frac{\delta y^2}{\delta t}\frac{\partial^2 p}{\partial y'^2}$$

Now take limits. This only makes sense if $\frac{\delta y^2}{\delta t}$ is $O(1)$, i.e. $\delta y^2 \sim O(\delta t)$ and letting $\delta y,\ \delta t \longrightarrow 0$ gives the equation

$$\frac{\partial p}{\partial t'} = \tfrac{1}{2}\frac{\partial^2 p}{\partial y'^2}$$

This is called the **forward Kolmogorov equation.** Also called Fokker Planck equation.

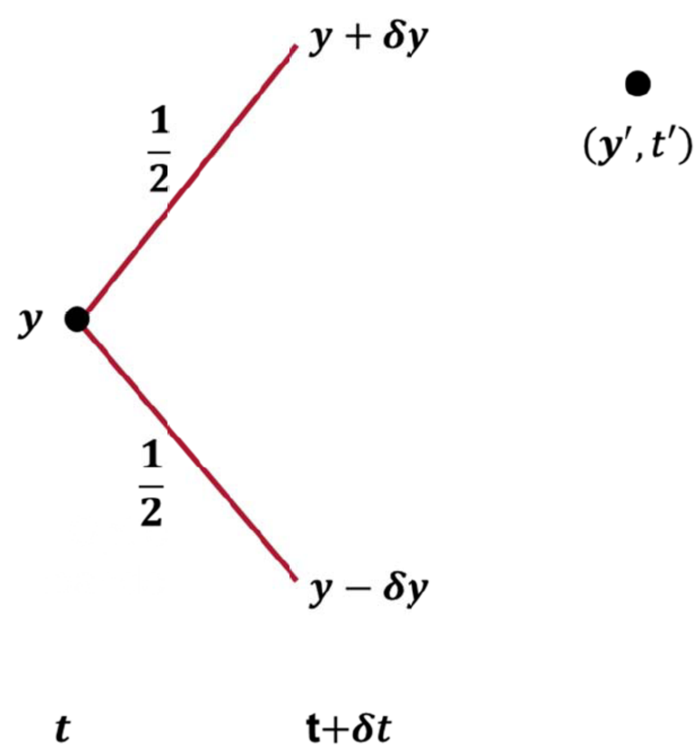
It shows how the probability density of future states evolves, starting from (y, t).

A particular solution of this is

$$p(y, t; y', t') = \frac{1}{\sqrt{2\pi(t'-t)}}\exp\left(-\frac{(y'-y)^2}{2(t'-t)}\right)$$

At $t' = t$ this is equal to $\delta(y'-y)$. The particle is known to start from (y, t) and its density is normal with mean y and variance $t'-t$.

The **backward equation** tells us the probability that we are at (y, t) given that we are at (y', t') in the future. So (y', t') are now fixed and (y, t) are variables. So the probability of being at (y, t) given we are at y' at t' is linked to the probabilities of being at $(y+\delta y, t+\delta t)$ and $(y-\delta y, t+\delta t)$.



$$p(y, t; y', t') = \tfrac{1}{2}p(y+\delta y, t+\delta t; y', t') + \tfrac{1}{2}p(y-\delta y, t+\delta t; y', t')$$

Since (y', t') do not change, drop these for the time being and use a TSE on the right hand side

$p(y,t) =$

$$\tfrac{1}{2}\left(p(y,t) + \frac{\partial p}{\partial t}\delta t + \frac{\partial p}{\partial y}\delta y + \tfrac{1}{2}\frac{\partial^2 p}{\partial y^2}\delta y^2 + ...\right) +$$
$$\tfrac{1}{2}\left(p(y,t) + \frac{\partial p}{\partial t}\delta t - \frac{\partial p}{\partial y}\delta y + \tfrac{1}{2}\frac{\partial^2 p}{\partial y^2}\delta y^2 + ...\right)$$

which simplifies to

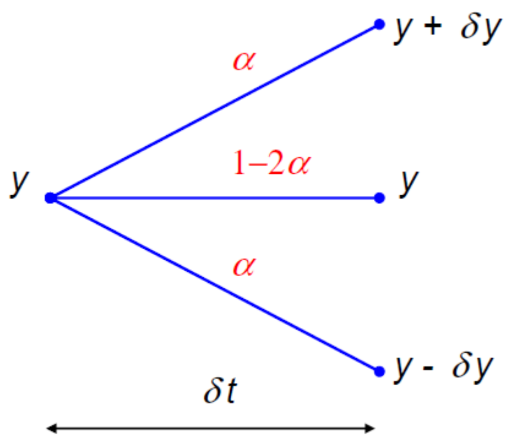
$$0 = \frac{\partial p}{\partial t} + \tfrac{1}{2}\frac{\delta y^2}{\delta t}\frac{\partial^2 p}{\partial y^2}.$$

Putting $\frac{\delta y^2}{\delta t} = O(1)$ and taking limit gives the **backward equation**

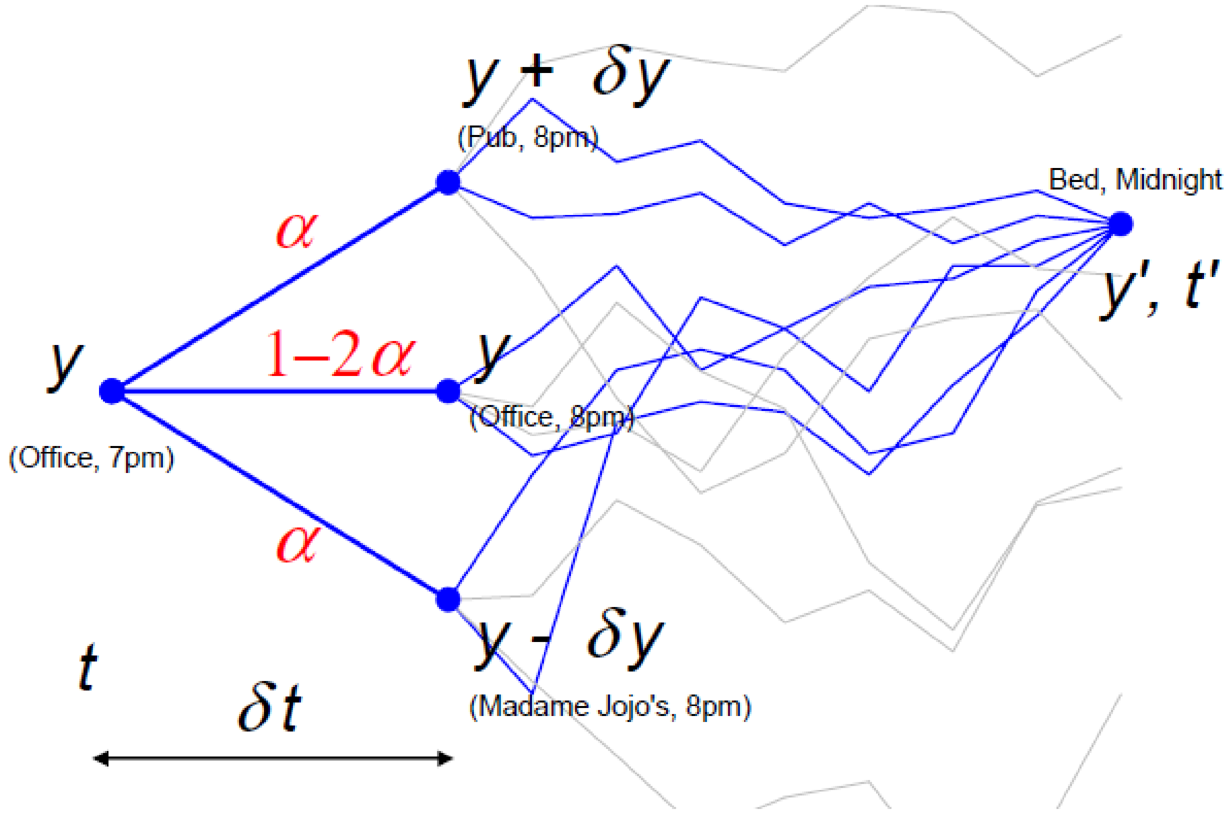
$$-\frac{\partial p}{\partial t} = \tfrac{1}{2}c^2\frac{\partial^2 p}{\partial y^2}.$$

or commonly written as $\dfrac{\partial p}{\partial t} + \dfrac{1}{2}\dfrac{\partial^2 p}{\partial y^2} = 0.$

The backward equation is particularly important in the context of finance, but also a source of much confusion. Illustrate with the 'real life' example that Wilmott uses. Wilmott uses a *Trinomial* Random Walk



So 3 possible states at the next time step. Here $\alpha < 1/2$.



- At 7pm you are at the office - this is the point (y, t)
- At 8pm you will be at one of three places:

  § The Pub - the point $(y + \delta y, t + \delta t)$;

  § Still at the office - the point $(y, t + \delta t)$;

  § Madame Jojo's - the point $(y - \delta y, t + \delta t)$

We are interested in the probability of being tucked up in bed at midnight (y', t'), given that we were at the office at 7pm.
Looking at the earlier figure, we can only get to bed at midnight via either

- the pub
- the office
- Madame Jojo's

at 8pm. What happens after 8pm doesn't matter - we don't care, you may not even remember! We are only concerned with being in bed at midnight.The earlier figure shows many different paths, only the ones ending up in 'our' bed are of interest to us.

In words: The probability of going from the office at 7pm to bed at midnight is

- the probability of going to the pub from the office and then to bed at midnight plus
- the probability of staying in the office and then going to bed at midnight plus
- the probability of going to Madame Jojo's from the office and then to bed at midnight.

The above can be expressed mathematically as

$$\begin{aligned} p\left(y,t;y',t'\right) &= \alpha p\left(y+\delta y,t+\delta t;y',t'\right)+\left(1-2\alpha\right)p\left(y,t+\delta t;y',t'\right)+ \\ &\quad \alpha p\left(y-\delta y,t+\delta t;y',t'\right). \end{aligned}$$

Performing a Taylor expansion gives dropping y',t'

$$\begin{aligned} p\left(y,t\right) &= \alpha\left(p+\frac{\partial p}{\partial t}\delta t+\frac{\partial p}{\partial y}\delta y+\frac{1}{2}\frac{\partial^2 p}{\partial y^2}\delta y^2+..\right) \\ &\quad +\left(1-2\alpha\right)\left(p-\frac{\partial p}{\partial t}\delta t+..\right) \\ &\quad \alpha\left(p+\frac{\partial p}{\partial t}\delta t-\frac{\partial p}{\partial y}\delta y+\frac{1}{2}\frac{\partial^2 p}{\partial y^2}\delta y^2+..\right). \end{aligned}$$

Most of the terms cancel and leave

$$0=\delta t\frac{\partial p}{\partial t}+\alpha\delta y^2\frac{\partial^2 p}{\partial y^2}+...$$

which becomes

$$0=\frac{\partial p}{\partial t}+\alpha\frac{\delta y^2}{\delta t}\frac{\partial^2 p}{\partial y^2}+...$$

and letting $\alpha\frac{\delta y^2}{\delta t}=c^2$ where c is non-zero and finite as $\delta t,\delta y\longrightarrow 0$, we have

$$\frac{\partial p}{\partial t}+c^2\frac{\partial^2 p}{\partial y^2}=0$$

Use of a trinomial random walk for the Forward Equation would have resulted in

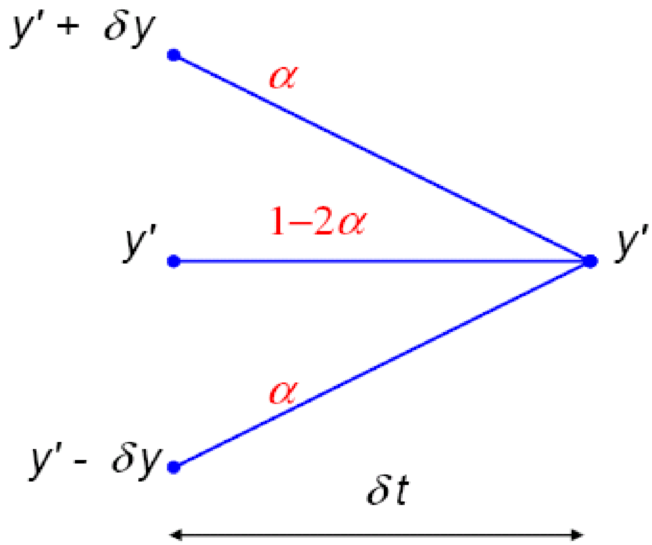
$$\frac{\partial p}{\partial t'}=c^2\frac{\partial^2 p}{\partial y'^2}.$$

This can be derived using the followingwhere the random walker is at y at an earlier time $t<t'$. The probability of being at y' at time t' is given by

$$p\left(y,t\right)=\alpha p\left(y+\delta y,t-\delta t\right)+\left(1-2\alpha\right)p\left(y,t-\delta t\right)+\alpha p\left(y-\delta y,t-\delta t\right)$$

Taylor series expansion gives

$$\begin{aligned} p\left(y+\delta y,t-\delta t\right) &= p\left(y,t\right)-\frac{\partial p}{\partial t}\delta t+\frac{\partial p}{\partial y}\delta y+\tfrac{1}{2}\frac{\partial^2 p}{\partial y^2}\delta y^2+... \\ p\left(y,t-\delta t\right) &= p\left(y,t\right)-\frac{\partial p}{\partial t}\delta t+... \\ p\left(y-\delta y,t-\delta t\right) &= p\left(y,t\right)-\frac{\partial p}{\partial t}\delta t-\frac{\partial p}{\partial y}\delta y+\tfrac{1}{2}\frac{\partial^2 p}{\partial y^2}\delta y^2+... \end{aligned}$$



Substituting into (1.1)

$$\begin{aligned} p(y,t) &= \alpha\left(p(y,t) - \frac{\partial p}{\partial t}\delta t + \frac{\partial p}{\partial y}\delta y + \tfrac{1}{2}\frac{\partial^2 p}{\partial y^2}\delta y^2 + ...\right) \\ &\quad + (1-2\alpha)\left(p(y,t) - \frac{\partial p}{\partial t}\delta t + ...\right) \\ &\quad + \alpha\left(p(y,t) - \frac{\partial p}{\partial t}\delta t - \frac{\partial p}{\partial y}\delta y + \tfrac{1}{2}\frac{\partial^2 p}{\partial y^2}\delta y^2 + ...\right) \end{aligned}$$

$$\frac{\partial p}{\partial t}\delta t = \alpha\frac{\partial^2 p}{\partial y'^2}\delta y^2$$

$$\frac{\partial p}{\partial t'} = \alpha\frac{\delta y^2}{\delta t}\frac{\partial^2 p}{\partial y'^2}$$

Now take limits. This only makes sense if $\frac{\delta y^2}{\delta t}$ is $O(1)$, i.e. $\delta y^2 \sim O(\delta t)$ and letting the constant $c^2 = \alpha\frac{\delta y^2}{\delta t}$, gives the equation

$$\frac{\partial p}{\partial t'} = c^2\frac{\partial^2 p}{\partial y'^2},$$

Solving the Forward Equation

The equation is

$$\frac{\partial p}{\partial t'} = c^2 \frac{\partial^2 p}{\partial y'^2}$$

for the unknown function $p = p(y', t')$. The idea is to obtain a solution in terms of Gaussian curves. Let's drop the primed notation (for convenience).

We assume a solution of the following form exists:

$$p(y,t) = t^a f\left(\frac{y}{t^b}\right)$$

where a, b are constants to be determined. In some textbooks start by specifying the value of a and b.

So create a new variable, from combining t and y

$$\xi = \frac{y}{t^b} = yt^{-b},$$

which is a dimensionless. We have the following derivatives

$$\frac{\partial \xi}{\partial y} = t^{-b}; \quad \frac{\partial \xi}{\partial t} = -byt^{-b-1}$$

we can now say

$$p(y,t) = t^a f(\xi)$$

therefore

$$\frac{\partial p}{\partial y} = \frac{\partial p}{\partial \xi}\frac{\partial \xi}{\partial y} = t^a f'(\xi) .t^{-b} = t^{a-b} f'(\xi)$$

$$\begin{aligned}
\frac{\partial^2 p}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial p}{\partial y}\right) = \frac{\partial}{\partial y}\left(t^{a-b} f'(\xi)\right) \\
&= \frac{\partial \xi}{\partial y}\frac{\partial}{\partial \xi}\left(t^{a-b} f'(\xi)\right) \\
&= t^{a-b}\frac{1}{t^b}\frac{\partial}{\partial \xi} f'(\xi) = t^{a-2b} f''(\xi)
\end{aligned}$$

$$\frac{\partial p}{\partial t} = t^a \frac{\partial}{\partial t} f(\xi) + at^{a-1} f(\xi)$$

we can use the chain rule to write

$$\frac{\partial}{\partial t} f(\xi) = \frac{\partial f}{\partial \xi} \cdot \frac{\partial \xi}{\partial t} = -byt^{-b-1} f'(\xi)$$

so we have

$$\frac{\partial p}{\partial t} = at^{a-1} f(\xi) - byt^{a-b-1} f'(\xi)$$

and then substituting these expressions in to the pde gives

$$at^{a-1} f(\xi) - byt^{a-b-1} f'(\xi) = c^2 t^{a-2b} f''.$$

We know from ξ that

$$y = t^b \xi$$

hence the equation above becomes

$$at^{a-1}f(\xi) - b\xi t^{a-1}f'(\xi) = c^2 t^{a-2b} f''.$$

For the similarity solution to exist we require the equation to be independent of t, i.e. $a-1 = a-2b \Longrightarrow b = 1/2$, therefore

$$af - \frac{1}{2}\xi f' = c^2 f''$$

thus we have so far

$$p = t^a f\left(\tfrac{y}{\sqrt{t}}\right)$$

which gives us a whole family of solutions dependent upon the choice of a.

We know that p represents a pdf, hence

$$\int_{\mathbb{R}} p(y,t)\,dy = 1 = \int_{\mathbb{R}} t^a f\left(\tfrac{y}{\sqrt{t}}\right) dy$$

change of variables $u = y/\sqrt{t} \longrightarrow du = dy/\sqrt{t}$ so the integral becomes

$$t^{a+1/2}\int_{-\infty}^{\infty} f(u)\,du = 1$$

which we need to normalize independent of time t. This is only possible if $a = -1/2$.

So the D.E becomes

$$-\frac{1}{2}(f + \xi f') = c^2 f''.$$

We have an exact derivative on the lhs, i.e. $\dfrac{d}{d\xi}(\xi f) = f + \xi f'$, hence

$$-\frac{1}{2}\frac{d}{d\xi}(\xi f) = c^2 f''$$

and we can integrate once to get

$$-\frac{1}{2}(\xi f) = c^2 f' + K.$$

We obtain K from the following information about a probability density, as $\xi \to \infty$

$$\begin{aligned} f(\xi) &\rightarrow 0 \\ f'(\xi) &\rightarrow 0 \end{aligned}$$

hence $K = 0$ in order to get the correct solution, i.e.

$$-\frac{1}{2}(\xi f) = c^2 f'$$

which can be solved as a simple first order variable separable equation:

$$\begin{aligned} -\tfrac{1}{2}(\xi f) &= c^2\frac{df}{d\xi} \rightarrow \frac{df}{f} = -\tfrac{1}{2c^2}\xi d\xi \rightarrow \int \tfrac{df}{f} = -\tfrac{1}{2c^2}\xi d\xi \\ \log f &= -\tfrac{1}{4c^2}\xi^2 + C \end{aligned}$$

Taking exponentials of both sides

$$f(\xi) = e^{-\frac{1}{4c^2}\xi^2 + C} = e^C e^{-\frac{1}{4c^2}\xi^2};\ \text{now write } e^C = A$$

$$f(\xi) = A \exp\left(-\tfrac{1}{4c^2}\xi^2\right).$$

A is a normalizing constant, so write

$$A \int_{\mathbb{R}} \exp\left(-\tfrac{1}{4c^2}\xi^2\right) d\xi = 1.$$

Now substitute $x = \xi/2c$, so $2cdx = d\xi$

$$2cA \underbrace{\int_{\mathbb{R}} \exp\left(-x^2\right) dx}_{=\sqrt{\pi}} = 1,$$

which gives $A = 1/2c\sqrt{\pi}$. Returning to

$$p(y,t) = t^{-1/2} f(\xi)$$

becomes

$$p(y',t') = \frac{1}{2c\sqrt{\pi t'}} \exp\left(-\frac{y'^2}{4t'c^2}\right).$$

This is a pdf for a variable y that is normally distributed with mean zero and standard deviation $c\sqrt{2t}$, which we ascertained by the following comparison:

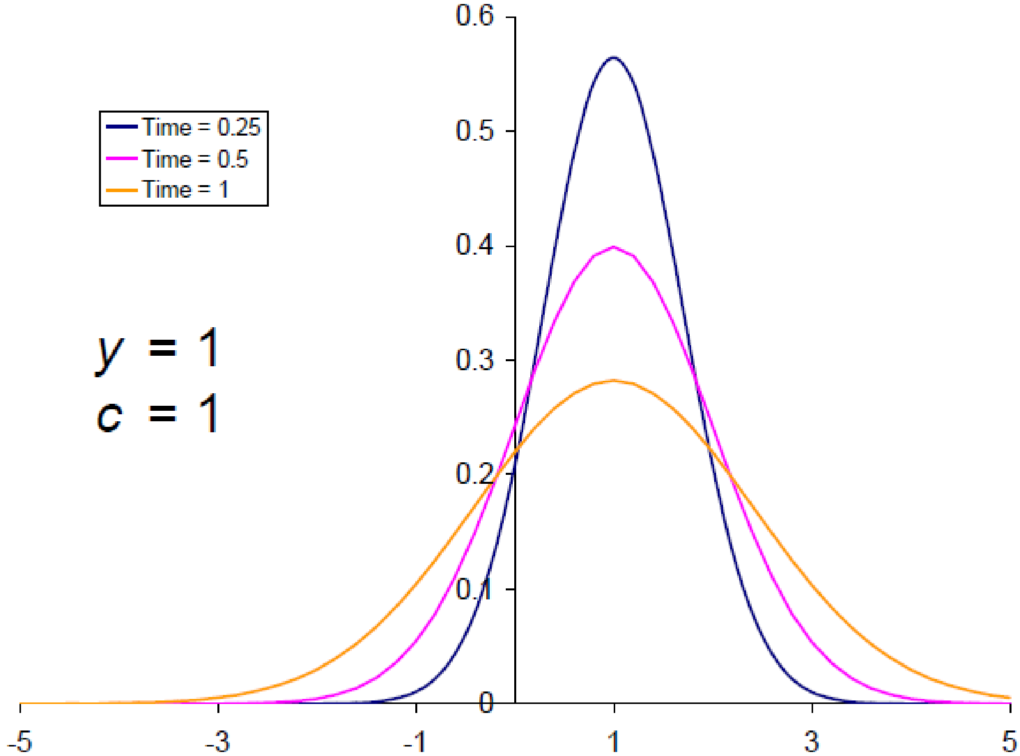
$$-\frac{1}{2}\frac{y'^2}{2t'c^2} : -\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}$$

i.e. $\mu \equiv 0$ and $\sigma^2 \equiv 2t'c^2$.

This solution is also called the **Source Solution** or **Fundamental Solution**.

If the random variable y' has value y at time t then we can generalize to

$$p(y,t;y',t') = \frac{1}{2c\sqrt{\pi(t'-t)}} \exp\left(-\frac{(y'-y)^2}{4c^2(t'-t)}\right)$$



At $t' = t$ this is now a Dirac delta function $\delta(y' - y)$. This particle is known to start from (y, t) and diffuses out to (y', t') with mean y and variance $(t' - t)$

Recall this behaviour of decay away from one point y, unbounded growth at that point and constant area means that $p(y, t; y', t')$ has turned in to a **Dirac delta function** $\delta(y' - y)$ as $t' \to t$.

Applied Stochastic Calculus

Stochastic Process

The evolution of financial assets is random and depends on time. They are examples of *stochastic processes* which are random variables indexed (parameterized) with time.

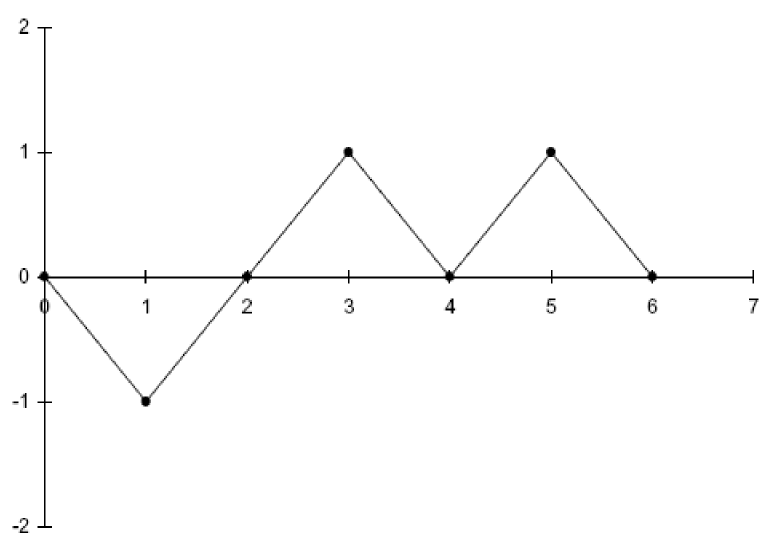
If the movement of an asset is discrete it is called a *random walk*. A continuous movement is called a *diffusion process*. We will consider the asset price dynamics to exhibit continuous behaviour and each random path traced out is called a *realization*.

We need a defintion and set of properties for the randomness observed in an asset price realization, which will be *Brownian Motion*.

This is named after the Scottish Botanist who in 1827, while examining grains of pollen of the plant Clarkia pulchella suspended in water under a microscope, observed minute particles, ejected from the pollen grains, executing a continuous fidgety motion. In 1900 Louis Bachelier was the first person to model the share price movement using Brownian motion as part of his PhD. Five years later Einstein used Brownian motion to study diffusions. In 1920 Norbert Wiener, a mathematician at MIT provided a mathematical construction of Brownian motion together with numerous results about the properties of Brownian motion - in fact he was the first to show that Brownian motion exists and is a well defined entity! Hence Wiener process is also used as a name for this.

Construction of Brownian Motion

Brownian Motion can be constructed as a carefully scaled limit of a symmetric random walk, in the context of a simple gambling game. Consider the coin tossing experiment



where we define the random variable

$$R_i = \begin{cases} +1 & \text{if } H \\ -1 & \text{if } T \end{cases}$$

and examine the statistical properties of R_i.

Firstly the mean

$$\mathbb{E}[R_i] = (+1)\frac{1}{2} + (-1)\frac{1}{2} = 0$$

and secondly the variance

$$\begin{aligned}\mathbb{V}[R_i] &= \mathbb{E}\left[R_i^2\right] - \underbrace{\mathbb{E}^2[R_i]}_{=0} \\ &= \mathbb{E}\left[R_i^2\right] = 1\end{aligned}$$

Suppose we now wish to keep a score of our winnings after the n^{th} toss - we introduce a new random variable

$$W_n = \sum_{i=1}^{n} R_i$$

This allows us to keep a track of our total winnings. This represents the position of a marker that starts off at the origin (no winnings). So starting with no money means

$$W_0 = R_0 = 0$$

Now we can calculate expectations of W_n

$$\mathbb{E}[W_n] = \mathbb{E}\left[\sum_{i=1}^{n} R_i\right] = \sum_{i=1}^{n} \mathbb{E}[R_i] = 0$$

$\mathbb{E}[X_n^2] =$

$$\begin{aligned}\mathbb{E}\left[X_n^2\right] &= \mathbb{E}\left[R_1^2 + R_2^2 + ...R_n^2 + 2R_1R_2 + ... + 2R_{n-1}R_n\right] \\ &= \mathbb{E}\left[\sum_{i=1}^{n} R_i^2\right] + \mathbb{E}\left[\sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n} R_iR_j\right] = \sum_{i=1}^{n}\mathbb{E}\left[R_i^2\right] + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n}\mathbb{E}[R_i]\,\mathbb{E}[R_j] \\ &= n\times 1 + 0\times 0 = n\end{aligned}$$

A Note on Variations

Consider a function f_t, where $t_i = i\dfrac{t}{n}$, we can define different measures of how much f_t varies over time as

$$V^N = \sum_{i=1}^{n} \left|f_{t_i} - f_{t_{i-1}}\right|^N$$

The cases $N = 1, 2$ are important.

$$V = \sum_{i=1}^{n} \left|f_{t_i} - f_{t_{i-1}}\right| \quad \textit{variation} \text{ of trajectory - sum of absolute changes}$$

$$V^2 = \sum_{i=1}^{n} \left|f_{t_i} - f_{t_{i-1}}\right|^2 \quad \textit{quadratic variation} \text{ - sum of squared changes}$$

Now look at the *quadratic variation* of the random walk. After each toss, we have won or lost \$1. That is

$$W_n - W_{n-1} = \pm 1 \Longrightarrow |W_n - W_{n-1}| = 1$$

Hence

$$\sum_{i=1}^{n} \underbrace{(W_i - W_{i-1})^2}_{=1} = n$$

Let's now extend this by introducing time dependence. Perform six tosses of a coin in a time t. So each toss must be performed in time $t/6$, and a bet size of $\sqrt{t/6}$ (and not \$1), i.e. we win or lose $\sqrt{t/6}$ depending on the outcome.

Let's examine the quadratic variation for this experiment

$$\begin{aligned} & \sum_{i=1}^{6} (W_i - W_{i-1})^2 \\ = & \sum_{i=1}^{6} \left(\pm\sqrt{t/6} \right)^2 \\ = & \ 6 \times \frac{t}{6} = t \end{aligned}$$

Now speed up the game. So we perform n tosses within time t with each bet being $\sqrt{t/n}$. Time for each toss is t/n.

$$W_i - W_{i-1} = \pm\sqrt{t/n}$$

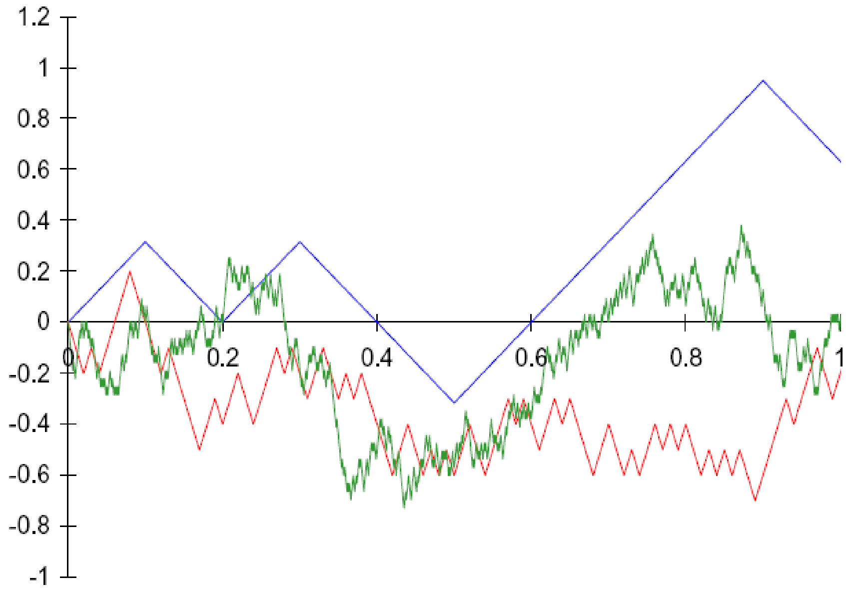
The quadratic variation is

$$\begin{aligned} \sum_{i=1}^{n} (W_i - W_{i-1})^2 &= n \times \left(\pm\sqrt{t/n} \right)^2 \\ &= t \end{aligned}$$

As n becomes larger and larger, time between subsequent tosses decreases and the bet sizes become smaller. The time and bet size decrease in turn like

$$\begin{aligned} \text{time decrease} &\sim O\left(\frac{1}{n}\right) \\ \text{bet size} &\sim O\left(\frac{1}{\sqrt{n}}\right) \end{aligned}$$

The diagram below shows a series of coin tossing experiments.



The scaling we have used has been chosen carefully to both keep the random walk finite and also not becoming zero. i.e. In the limit $n \longrightarrow \infty$, the random walk stays finite. It has an expectation conditional on a starting value of zero, of

$$\begin{aligned} \mathbb{E}[W_t] &= \mathbb{E}\left[\lim_{n\to\infty}\sum_{i=1}^{n} R_i\right] \\ &= \lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}[R_i] = n\cdot 0 \\ \text{Mean of } W_t &= 0 \end{aligned}$$

$$\begin{aligned} \mathbb{E}\left[W_t^2\right] &= \mathbb{E}\left[\lim_{n\to\infty}\sum_{i=1}^{n} R_i^2\right] \\ &= \lim_{n\to\infty}\sum_{i=1}^{n}\mathbb{E}\left[R_i^2\right] = \lim_{n\to\infty} n\cdot\left(\sqrt{t/n}\right)^2 \\ \mathbb{V}[W_t] &= \mathbb{E}\left[W_t^2\right] = t \end{aligned}$$

This limiting process as dt tends to zero is called Brownian Motion and denoted W_t.

Alternative notation for Brownian motion/Wiener process is X_t or B_t.

Mean Square Convergence

Consider a function $F(X)$. If

$$\mathbb{E}\left[(F(X)-l)^2\right] \longrightarrow 0$$

then we say that $F(X) = l$ in the *mean square limit*, also called *mean square convergence*. We present a full derivation of the mean square limit. Starting with the quantity:

$$\mathbb{E}\left[\left(\sum_{j=1}^{n}\left(W(t_j)-W(t_{j-1})\right)^2 - t\right)^2\right]$$

where $t_j = \dfrac{jt}{n} = j\Delta t$.

Hence we are saying that *up to mean square convergence,*

$$dW^2 = dt.$$

This is the symbolic way of writing this property of a Wiener process, as the partitions Δt become smaller and smaller.

Developing the terms inside the expectation

First, we will simplify the notation in order to deal more easily with the outer (right most) squaring. Let $Y(t_j) = \left(W(t_j) - W(t_{j-1})\right)^2$, then we can rewrite the expectation as:

$$\mathbb{E}\left[\left(\sum_{j=1}^{n} Y(t_j) - t\right)^2\right]$$

Expanding we have:

$$\mathbb{E}\left[\left(Y(t_1)+Y(t_2)+\ldots+Y(t_n)-t\right) \times \left(Y(t_1)+Y(t_2)+\ldots+Y(t_n)-t\right)\right]$$

The term inside the Expectation is equal to

$$\begin{aligned}
&Y(t_1)^2 + Y(t_1)Y(t_2) + \ldots + Y(t_1)Y(t_n) - Y(t_1)t \\
&+Y(t_2)^2 + Y(t_2)Y(t_1) + \ldots + Y(t_2)Y(t_n) - Y(t_2)t \\
&\vdots \\
&+Y(t_n)^2 + Y(t_n)Y(t_1) + \ldots + Y(t_n)Y(t_{n-1}) - Y(t_n)t \\
&-tY(t_1) - tY(t_2) - \ldots - tY(t_n) + t^2
\end{aligned}$$

Rearranging

$$\begin{aligned}
&Y(t_1)^2 + Y(t_2)^2 + \ldots + Y(t_n)^2 \\
&2Y(t_1)Y(t_2) + 2Y(t_1)Y(t_3) + \ldots + 2Y(t_{n-1})Y(t_n) \\
&-2Y(t_1)t - 2Y(t_2)t - \ldots - 2Y(t_n)t \\
&+t^2
\end{aligned}$$

We can now factorize to get

$$\sum_{j=1}^{n} Y(t_j)^2 + 2\sum_{i=1}^{n}\sum_{j<i} Y(t_i)Y(t_j) - 2t\sum_{j=1}^{n} Y(t_j) + t^2$$

Substituting back $Y(t_j) = (W(t_j) - W(t_{j-1}))^2$ and taking the expectation, we arrive at:

$$\begin{aligned}
&\mathbb{E}\,[\sum_{j=1}^{n} (W(t_j) - W(t_{j-1}))^4 \\
&+2\sum_{i=1}^{n}\sum_{j<i} (W(t_i) - W(t_{i-1}))^2 (W(t_j) - W(t_{j-1}))^2 \\
&-2t\sum_{j=1}^{n} (W(t_j) - W(t_{j-1}))^2 \\
&+t^2\,]
\end{aligned}$$

Computing the expectation

By linearity of the expectation operator, we can write the previous expression as:

$$\begin{aligned}
&\sum_{j=1}^{n} \mathbb{E}\left[(W(t_j) - W(t_{j-1}))^4\right] \\
&+2\sum_{i=1}^{n}\sum_{j<i} \mathbb{E}\left[(W(t_i) - W(t_{i-1}))^2 (W(t_j) - W(t_{j-1}))^2\right] \\
&-2t\sum_{j=1}^{n} \mathbb{E}\left[(W(t_j) - W(t_{j-1}))^2\right] \\
&+t^2
\end{aligned}$$

Now, since $Z(t_j) = W(t_j) - W(t_{j-1})$ follows a Normal distribution with mean 0 and variance $\frac{t}{n}\,(= dt)$, it follows (standard result) that its fourth moment is equal to $3\frac{t^2}{n^2}$. We will show this shortly.

Firstly we know that $Z(t_j) \sim N\left(0, \frac{t}{n}\right)$, i.e.

$$\mathbb{E}\left[Z(t_j)\right] = 0, \quad \mathbb{V}\left[Z(t_j)\right] = \frac{t}{n}$$

therefore we can construct its PDF. For any random variable $\psi \sim N(\mu, \sigma^2)$ its probability density is given by

$$p(\psi) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(\psi-\mu)^2}{\sigma^2}\right)$$

hence for $Z(t_j)$ the PDF is

$$p(z) = \frac{1}{\sqrt{t/n}\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{z^2}{t/n}\right)$$

$$\begin{aligned}
\mathbb{E}\left[(W(t_j) - W(t_{j-1}))^4\right] &= \mathbb{E}\left[Z^4\right] \\
&= 3\frac{t^2}{n^2} \qquad \text{for } j = 1, \ldots, n
\end{aligned}$$

So

$$\begin{aligned}
\mathbb{E}\left[Z^4\right] &= \int_{\mathbb{R}} Z^4 p(z)\,dz \\
&= \sqrt{\frac{n}{2t\pi}} \int_{\mathbb{R}} Z^4 \exp\left(-\frac{1}{2}\frac{z^2}{t/n}\right) dz
\end{aligned}$$

now put

$$u = \frac{z}{\sqrt{t/n}} \longrightarrow du = \sqrt{n/t}dz$$

Our integral becomes

$$\sqrt{\frac{n}{2t\pi}} \int_{\mathbb{R}} \left(\sqrt{\frac{t}{n}} u \right)^4 \exp\left(-\tfrac{1}{2}u^2\right) \sqrt{\frac{t}{n}} du$$

$$\begin{aligned}
&= \sqrt{\frac{1}{2\pi}\frac{t^2}{n^2}} \int_{\mathbb{R}} u^4 \exp\left(-\tfrac{1}{2}u^2\right) du \\
&= \frac{t^2}{n^2} . \sqrt{\frac{1}{2\pi}} \int_{\mathbb{R}} u^4 \exp\left(-\tfrac{1}{2}u^2\right) du \\
&= \frac{t^2}{n^2} . \mathbb{E}\left[u^4\right].
\end{aligned}$$

So the problem reduces to finding the fourth moment of a standard normal random variable. Here we do not have to explicitly calculate any integral. Two ways to do this.

Either use the MGF as we did earlier and obtained the fourth moment to be three.

Or the other method is to make use of the fact that the kurtosis of the standardised normal distribution is 3.

That is

$$\mathbb{E}\left[\frac{(\phi-\mu)^4}{\sigma^4}\right] = \mathbb{E}\left[\frac{(\phi-0)^4}{1^4}\right] = 3.$$

Hence $\mathbb{E}\left[u^4\right] = 3$ and we can finally write $3\frac{t^2}{n^2}$.

and

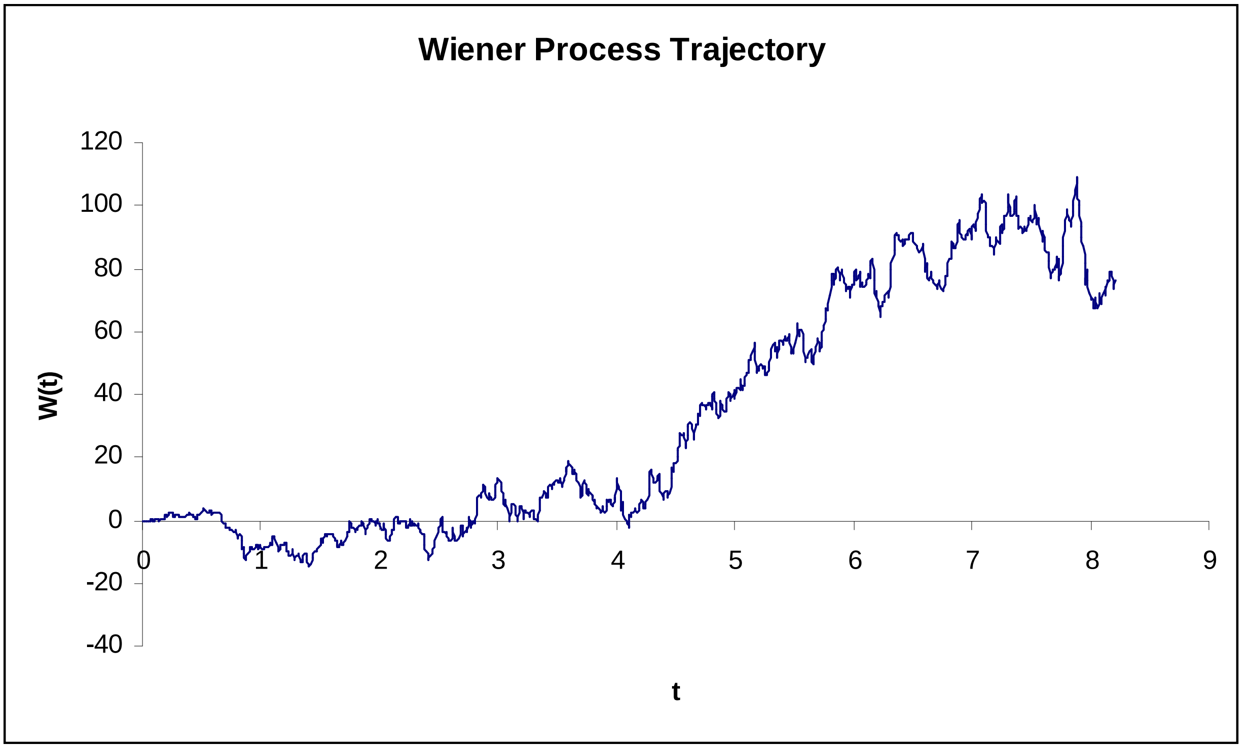
$$\mathbb{E}\left[\left(W(t_j) - W(t_{j-1})\right)^2\right] = \frac{t}{n} \qquad \text{for } j = 1, \ldots, n$$

Because of the single summation, the fourth moment and the variance multiplied by t actually recur n times. Because of the double summation, the product of variances occurs $\frac{n(n-1)}{2}$ times.

We can now conclude that the expectation is equal to:

$$\begin{aligned} & 3n\frac{t^2}{n^2} + n(n-1)\frac{t^2}{n^2} - 2tn\frac{t}{n} + t^2 \\ = \ & 3\frac{t^2}{n} + t^2 - \frac{t^2}{n} - 2t^2 + t^2 = 2\frac{t^2}{n} \\ = \ & O(\frac{1}{n}) \end{aligned}$$

So, as our partition becomes finer and finer and n tends to infinity, the quadratic variation will tend to t *in the mean square limit*.



The diagram above represents a realisation of a Wiener process, with $\Delta t = 0.0001$.

Numerical Scheme:

$$\begin{aligned} \text{Start} \ &: \ t_0, \ W_0 = 0; \ \text{define } \Delta t = T/n \\ \text{loop } i \ &= \ 1, 2, \cdots, n: \\ t_i \ &= \ t_{i-1} + \Delta t \\ \text{draw } \phi \ &\sim \ N(0,1) \\ W_i \ &= \ W_{i-1} + \phi\sqrt{\Delta t} \end{aligned}$$

Taylor Series and Itô

If we were to do a naive Taylor series expansion of F, completely disregarding the nature of W, and treating dW as a small increment in W, we would get

$$F(W + dW) = F(W) + \frac{dF}{dW}dW + \frac{1}{2}\frac{d^2F}{dW^2}dW^2,$$

ignoring higher-order terms.

We could argue that $F(W + dW) - F(W)$ was just the 'change in' F and so

$$dF = \frac{dF}{dW}dW + \frac{1}{2}\frac{d^2F}{dW^2}dW^2.$$

This is *almost* correct.

Because of the way that we have defined Brownian motion, and have seen how the quadratic variation behaves, it turns out that the dW^2 term isn't really random at all.

The dW^2 term becomes (as all time steps become smaller and smaller) the same as its average value, dt.

Taylor series and the 'proper' Itô are very similar. The only difference being that the correct Itô's lemma has a dt instead of a dW^2.

You can, with little risk of error, use Taylor series with the 'rule of thumb'

$$dW^2 = dt.$$

- and in practice you will get the right result.

We can now answer the question, "If $F = W^2$ what is dF?" In this example

$$\frac{dF}{dW} = 2W \text{ and } \frac{d^2F}{dW^2} = 2.$$

Therefore Itô's lemma tells us that

$$dF = dt + 2W dW.$$

This is an example of a **stochastic differential equation (SDE)**.

Now consider a slight extension. A function of a Wiener Process $f = f(t, W(t))$, so we can allow both t and $W(t)$ to change, i.e.

$$\begin{aligned} t &\longrightarrow t + dt \\ W &\longrightarrow W + dW. \end{aligned}$$

Using Taylor as before

$$\begin{aligned} f(t + dt, W + dW) &= f(t, W) + \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial W} dW + \tfrac{1}{2} \frac{\partial^2 f}{\partial W^2} dW^2 + .. \\ df &= f(t + dt, W + dW) - f(t, W) = \\ & \left(\frac{\partial f}{\partial t} + \tfrac{1}{2} \frac{\partial^2 f}{\partial W^2} \right) dt + \frac{\partial f}{\partial W} dW \end{aligned}$$

This gives another form of Itô:

$$df = \left(\frac{\partial f}{\partial t} + \tfrac{1}{2} \frac{\partial^2 f}{\partial W^2} \right) dt + \frac{\partial f}{\partial W} dW. \qquad (*)$$

This is also a SDE.

Examples:

1. Obtain a SDE for $f = te^{W(t)}$. We need $\frac{\partial f}{\partial t} = e^{W(t)}$; $\frac{\partial f}{\partial W} = te^{W(t)} = \frac{\partial^2 f}{\partial W^2}$, then substituting in $(*)$

$$df = \left(e^{W(t)} + \tfrac{1}{2} te^{W(t)} \right) dt + te^{W(t)} dW.$$

We can factor out $te^{W(t)}$ and rewrite the above as

$$\frac{df}{f} = \left(\tfrac{1}{t} + \tfrac{1}{2} \right) dt + dW.$$

2. Consider the function of a stochastic variable $f = t^2 W^n(t)$

$$\frac{\partial f}{\partial t} = 2tW^n; \ \frac{\partial f}{\partial W} = nt^2 W^{n-1}; \frac{\partial^2 f}{\partial W^2} = n(n-1) t^2 W^{n-2},$$

in $(*)$ gives

$$df = \left(2tW^n + \tfrac{1}{2} n(n-1) t^2 W^{n-2} \right) dt + nt^2 W^{n-1} dW.$$

Itô multiplication table:

$\times$	dt	dW
dt	$dt^2 = 0$	$dtdW = 0$
dW	$dWdt = 0$	$dW^2 = dt$

A Formula for Stochastic Integration

If we take the 2D form of Itô given by $(*)$, rearrange and integrate over $[0, t]$, we obtain a very nice formula for integrating functions of the form $f(t, W(t))$:

$$\int_0^t \frac{\partial f}{\partial W} dW = f(t, W(t)) - f(0, W(0)) - \int_0^t \left(\frac{\partial f}{\partial \tau} + \frac{1}{2} \frac{\partial^2 f}{\partial W^2} \right) d\tau$$

Example: Show that

$$\int_0^t \left(t + e^W\right) dW = tW + e^W - 1 - \int_0^t \left(W_\tau + \tfrac{1}{2} e^{W_\tau}\right) d\tau.$$

Comparing this to the stochastic integral formula above, we see that $\frac{\partial f}{dW} \equiv t + e^W \Longrightarrow f = tW + e^W$. Also

$$\frac{\partial^2 f}{\partial W^2} = e^{W_t}, \quad \frac{\partial f}{\partial t} = W_t.$$

Substituting all these terms in to the formula and noting that $f(0, W(0)) = 1$ verifies the result.

Naturally if $f = f(W(t))$ then the integral formula simply collapses to

$$\int_0^t \frac{df}{dW} dW = f(W(t)) - f(W(0)) - \frac{1}{2} \int_0^t \frac{d^2 f}{dW^2} d\tau$$