Stochastic Calculus and Itô's Lemma

In this lecture. . .

* Construction of a B.M

* Properties

* Itô and SDE

* Applications

By the end of this lecture you will be able to

- understand where Brownian motion and diffusion processes come from

  B.M.

- manipulate functions of random variables

Introduction

Stochastic calculus is very important in the mathematical modeling of financial processes. This is because of the underlying (assumed) random nature of financial markets.

Stochastic calculus: A motivating example

Toss a coin. Every time you throw a head you receive \$1, every time you throw a tail you pay out \$1. In the experiment below the sequence was THHTHT, and you finished even.

at start of game – zero winnings. $p(H) = p(T) = \frac{1}{2}$

No money

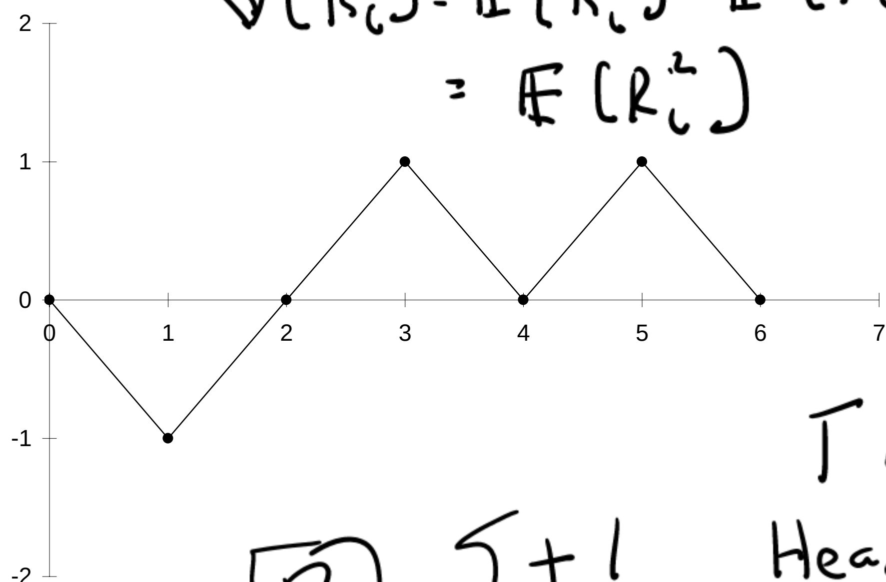
$V(R_i) = \mathbb{E}[R_i^2] - \mathbb{E}^2(R_i)$

$= \mathbb{E}(R_i^2)$

$p(H) \longrightarrow £1$

$p(T) \longrightarrow -£1$

0



$p(+1) = \frac{1}{2}$

$p(-1) = \frac{1}{2}$

THHTHT $\longrightarrow 0$

$\boxed{R_i} = \begin{cases} +1 & \text{Head} \\ -1 & \text{Tail} \end{cases}$

$\mathbb{E}(R_i) = \sum x_i p_i$

$= (+1)\frac{1}{2} + (-1)\frac{1}{2} = 0$

- R_i denotes the random amount, either \$1 or -\$1, you make on the ith toss:

$$E[R_i] = 0, \quad \underbrace{E[R_i^2]}_{\text{variance}} = 1 \quad \text{and} \quad \overbrace{E[R_i R_j] = 0.}^{} \quad \text{indep.}$$

$= (+1)^2 \frac{1}{2} + (-1)^2 \frac{1}{2}$

In this example it doesn't matter whether or not these expectations are conditional on the past. In other words, if you threw five heads in a row it does not affect the outcome of the sixth toss.

- Introduce S_i to mean the total amount of money you have won up to and including the ith toss so that

New R.V. $$S_i = \sum_{j=1}^{i} R_j.$$ subject to $S_0 = 0$

Later on it will be useful if we have $S_0 = 0$, i.e., you start with no money.

If we now calculate expectations of S_i it *does* matter what information we have. If we calculate expectations of future events before the experiment has even begun then

$$E[S_i] = 0 \quad \text{and} \quad E[S_i^2] = E[R_1^2 + 2R_1R_2 + \cdots] = i.$$

On the other hand, suppose there have been five tosses already, can we use this information and what can we say about expectations for the sixth toss?

$$\mathbb{E}\Big[\sum_{j=1}^{i} R_j\Big] = \sum_{j=1}^{i} \mathbb{E}(R_j) = i \times 0 = 0$$

- This is the **conditional expectation**.

$$\mathbb{V}(S_i) = \mathbb{E}[S_i^2] = \mathbb{E}\Big[\Big(\sum_{j=1}^{i} R_j\Big)^2\Big] = \mathbb{E}\big[R_1^2 + R_2^2 + \cdots + R_i^2 + 2R_1R_2 + \cdots\big]$$

The expectation of S_6 conditional upon the previous five tosses gives

$$= \mathbb{E}\Big[\sum_{j=1}^{i} R_j^2\Big]$$

$$S_5$$

$$E[S_6|R_1, \ldots, R_5] = S_5.$$

$$= \sum_{j=1}^{i} \mathbb{E}(R_j^2) = i \times 1 = i$$

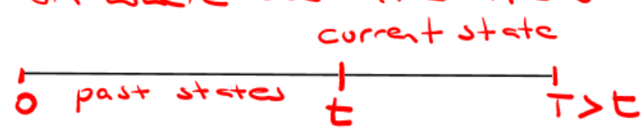
The Markov property Memoryless

- The distribution of the value of the random variable S_i conditional upon all of the past events *only depends on the previous value* S_{i-1}. This is the **Markov property**.

We say that the random walk has no memory beyond where it is now. Note that it doesn't have to be the case that the expected value of the random variable S_i is the same as the previous value.

This can be generalized to say that given information about S_j for some values of $1 \leq j < i$ then the only information that is of use to us in estimating S_i is the value of S_j for the largest j for which we have information.

Markov: where future depends at most on where we are now



The martingale property (Special case of Markov)

The coin tossing experiment possesses another property that can be important in finance. You know how much money you have won after the fifth toss. Your expected winnings after the sixth toss, and indeed after any number of tosses if we keep playing, is just the amount you already hold.

- The conditional expectation of your winnings at any time in the future is just the amount you already hold:

$$E_j[S_i|S_j, j < i] = S_j.$$

Expectation of future value

current value

This is called the **martingale property**. Expect future value to be current value.

Quadratic variation

measure of variation of random walk

The **quadratic variation** of the random walk is defined by

$$Q = \sum_{j=1}^{i} \left(S_j - S_{j-1} \right)^2 . \quad \pm 1$$

Because you either win or lose an amount \$1 after each toss, $|S_j - S_{j-1}| = 1$. Thus the quadratic variation is always i:

$$\sum_{j=1}^{i} \left(S_j - S_{j-1} \right)^2 = i. \qquad \boxed{i}$$

We are going to use the coin-tossing experiment for one more demonstration. And that will lead us to a continuous-time random walk.

Brownian motion

$S_0 = 0$

$\frac{t}{6}$

Change the rules of the coin-tossing experiment slightly.

$\pm\sqrt{\frac{t}{6}}$

First of all restrict the time allowed for the six tosses to a period t, so each toss will take a time $t/6$. Second, the size of the bet will not be \$1 but $\sqrt{t/6}$.

$n=6$

$p(H) = p\left(+\sqrt{t/6}\right) = \frac{1}{2}$

This new experiment clearly still possesses both the Markov and martingale properties, and its quadratic variation measured over the whole experiment is

$p(T) = p\left(-\sqrt{\frac{t}{6}}\right) = \frac{1}{2}$

$$\sum_{j=1}^{6} \underbrace{\left(S_j - S_{j-1}\right)^2}_{\left(\pm\sqrt{\frac{t}{6}}\right)^2} = 6 \times \left(\sqrt{\frac{t}{6}}\right)^2 = t.$$

$\pm\sqrt{\frac{t}{n}}$

finite

Change the rules again, to speed up the game. $\pm\sqrt{\frac{t}{n}} \pm \sqrt{t}\boxed{\frac{1}{\sqrt{n}}}$

- We will have n tosses in the allowed time t, with an amount $\sqrt{t/n}$ riding on each throw. time: $\frac{t}{n}$ bet-size $\pm\sqrt{t/n}$

$O(\frac{1}{n})$ $O(\frac{1}{\sqrt{n}})$

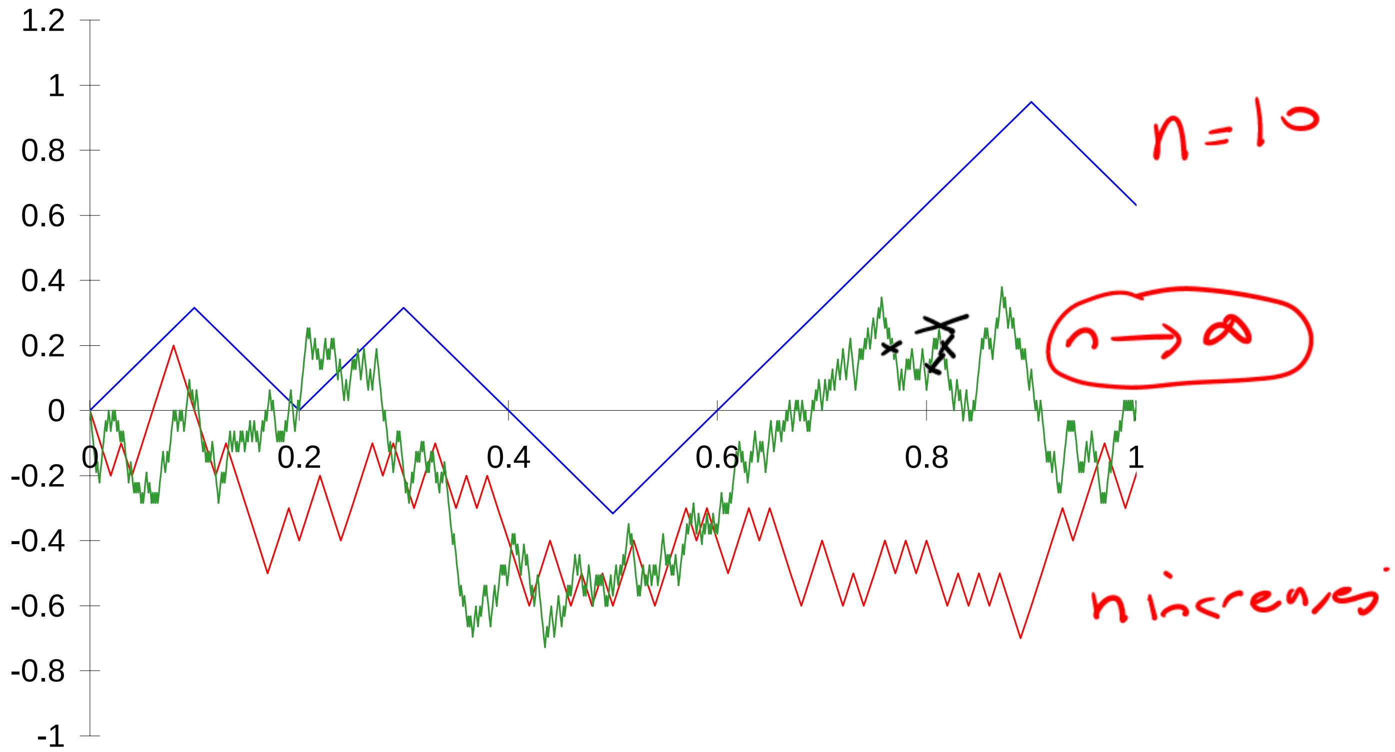
Again, the Markov and martingale properties are retained and the quadratic variation is still

$$Q = \sum_{j=1}^{n} \underbrace{\left(S_j - S_{j-1}\right)^2}_{\left(\pm\sqrt{t/n}\right)} = n \times \left(\sqrt{\frac{t}{n}}\right)^2 = t.$$

finite

$n \to \infty$

Now make n larger and larger. This speeds up the game, decreasing the time between tosses, with a smaller amount for each bet. But the new scalings have been chosen very carefully, the time step is decreasing like n^{-1} but the bet size only decreases by $n^{-1/2}$.



A series of coin tossing experiments.

$\frac{\delta y^2}{\delta t} \sim O(1)$ $\delta y^2 \sim \delta t$ $\delta y^2 \sim O\left(\frac{t}{n}\right)$ $\delta y \sim \sqrt{\delta t}$

As we go to the limit $n = \infty$, the resulting random walk stays finite. It has an expectation, conditional on a starting value of zero, of

$\frac{t}{n} = t \times \frac{1}{n}$ $\frac{1}{n}$

$$\text{mean} = E[S(t)] = 0$$

$\sqrt{t}$ $\frac{1}{\sqrt{n}}$ $\frac{1}{\sqrt{n}}$

and a variance

$$\text{Var} = 2^{nd} \text{ moment} = E[S(t)^2] = t.$$

We use $S(t)$ to denote the amount you have won or the value of the random variable after a time t.

$+\sqrt{\frac{t}{n}}\left(\frac{1}{2}\right) + \left(-\sqrt{\frac{t}{n}}\right)\frac{1}{2} = 0$ $\lim_{n\to\infty}\sum_{i=1}$ $\left(+\sqrt{\frac{t}{n}}\right)^2\frac{1}{2} + \left(-\sqrt{\frac{t}{n}}\right)^2\frac{1}{2}$

- The limiting process for this random walk as the time steps go to zero is called **Brownian motion**, and we will denote it by $X(t)$.

$\int_\infty$

$\boxed{X(t)}$ X_t $\frac{t}{n}$ dt dt W_t B_t

$\pm\sqrt{t/n}$ $\pm\sqrt{dt}$

The important properties of Brownian motion are as follows.

- $X_0 = 0$ a.s.

- *Finiteness*: Any other scaling of the bet size or 'increments' with time step would have resulted in either a random walk going to infinity in a finite time, or a limit in which there was no motion at all. It is important that the increment scales with the square root of the time step.

- $t \mapsto X_t$ is continuous everywhere.

- *Continuity*: The paths are continuous, there are no discontinuities. Brownian motion is the continuous-time limit of our discrete time random walk.

- *Markov*: The conditional distribution of $X(t)$ given information up until $\tau < t$ depends only on $X(\tau)$.

- *Martingale*: Given information up until $\tau < t$ the conditional expectation of $X(t)$ is $X(\tau)$.

- *Quadratic variation*: If we divide up the time 0 to t in a partition with $n+1$ partition points $t_i = i\,t/n$ then

$$\sum_{j=1}^{n} \Big(X(t_j) - X(t_{j-1})\Big)^2 \to t. \quad \text{(Technically `almost surely.')}$$

- *Normality*: Over finite time increments t_{i-1} to t_i, $X(t_i) - X(t_{i-1})$ is Normally distributed with mean zero and variance $t_i - t_{i-1}$.

$$X_{t_i} - X_{t_{i-1}} \sim N(0, |t_i - t_{i-1}|)$$

indep.

Having built up the idea and properties of Brownian motion from a series of experiments, we can discard the experiments, to leave the Brownian motion that is defined by its properties. These properties will be very important for our financial models.

$$\left\{ X_t : t \in \mathbb{R}^+ \right\} \quad X_t \sim \phi \sqrt{dt}$$

$$\phi \sim N(0,1)$$

Very Important Notation

We have seen X as the 'end result' of a random walk, up to some time t.

t

We will often work with the amount by which X changes from moment to moment.

- Think of dX as being an increment in X, i.e. a Normal random variable with mean zero and standard deviation $dt^{1/2}$.

$X_t \sim N(0, t)$ $\qquad X_t - X_0 \sim N(0, |t-0|)$

$t \to t + dt$

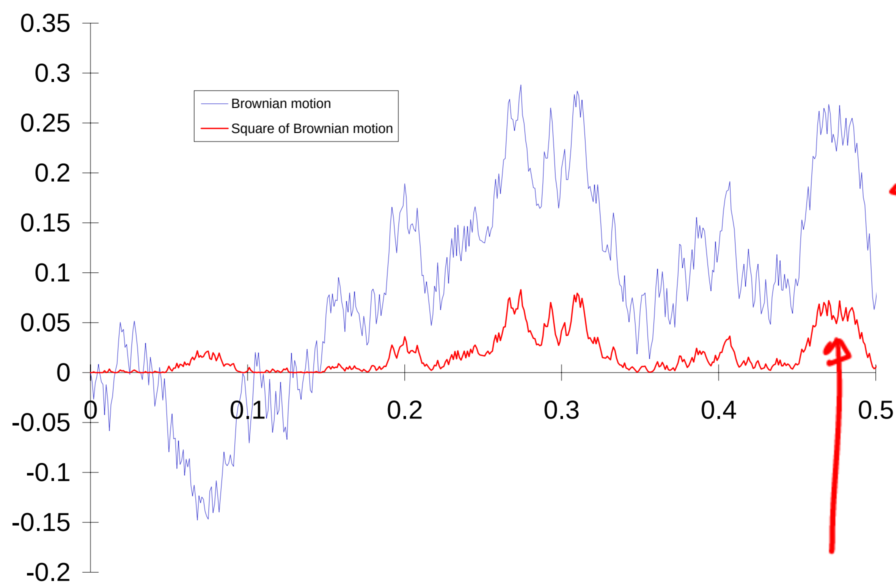
$dX_t = X_{t+dt} - X_t$

Functions of stochastic variables and Itô’s lemma

Now we’ll see the idea of a function of a stochastic variable. Below is shown a realization of a Brownian motion $X(t)$ and the function $F(X) = X^2$.

$F(x)$ fn. of a stoch. variable

X_t



a B.M

B.M squared

Whenever we have functions of a variable it is natural to want to know how to differentiate and manipulate these functions.

What are the rules of calculus when variables are stochastic?

$\therefore \quad dX \sim O(\sqrt{dt}) \qquad (dX)^2$

$dX^2 \sim dt$

$\delta x \to 0$

$\lim_{dt \to 0} \mathbb{E}[dX^2] = \underline{\underline{dt}}$

$\dfrac{f(x+\delta x) - f(x)}{\delta x} = f'(x)$

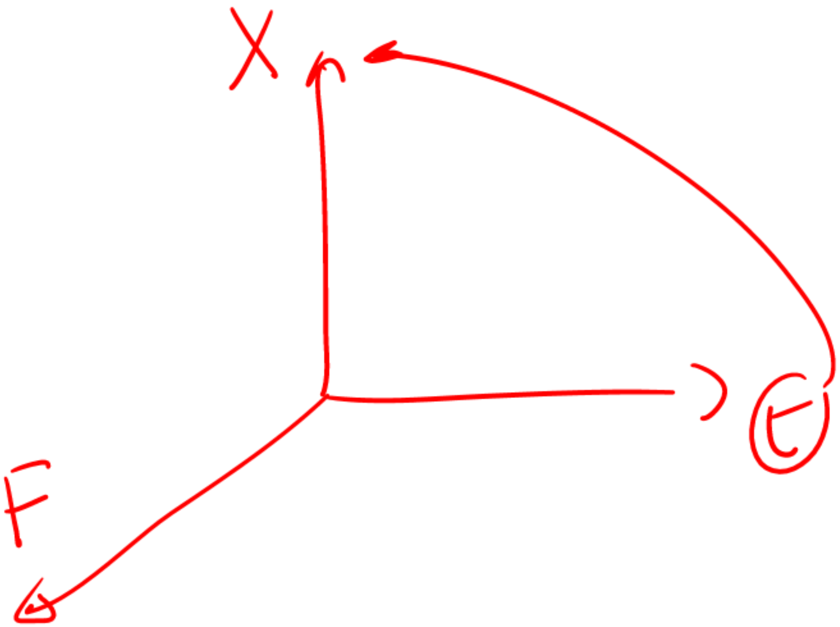
$\therefore \delta x^2$ is small

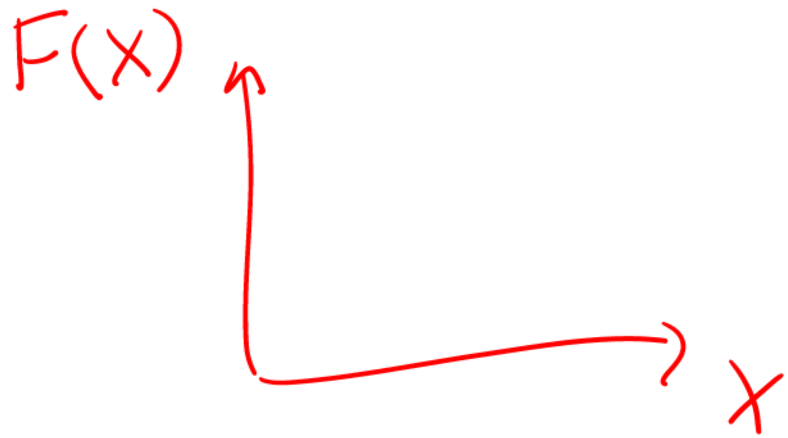
$f(x+\delta x) = f(x) + f'(x)\,\delta x + \cancel{O(\delta x^2)}$

The first point to note is that in the stochastic world we really have two 'variables.'

$X \quad X_t$

These are time t and the Brownian motion X.





X

We are used to writing ordinary and partial differential equations in the form

$$\frac{dF}{d\cdot}$$

or

$$\frac{\partial F}{\partial\cdot}$$

where the quantities on the bottom are the independent variables.

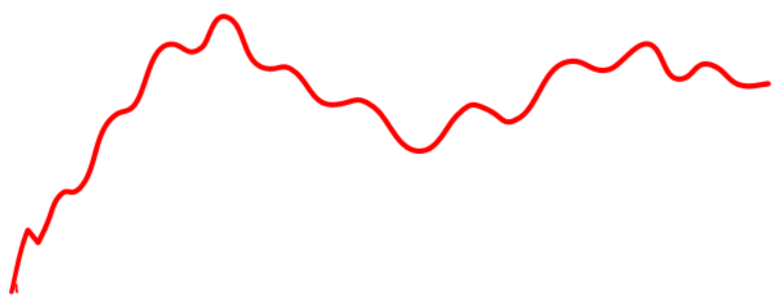
So might expect something similar in the stochastic world.

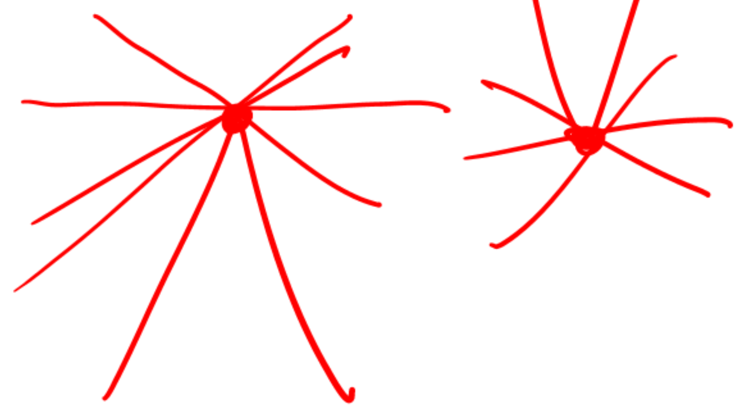
We immediately hit a problem, however.

Because dX is of size $\sqrt{dt}$ it is much bigger than dt.

This means that we have to be careful whenever we think about gradients/slopes/derivatives/sensitivities, since these are limits as dt goes to zero.

$\sqrt{dt} \to 0$ $\qquad$ $dt \to 0$





For this reason, in the stochastic world we instead work with stochastic differential equations.

$dS = \mu S\,dt + \sigma S\,dX$

These take the form

$$\frac{dF}{dt} = A(t, F)$$

O.D.E.

drift

$$dF = \cdots dt + \cdots dX.$$

$dF = A\,dt +$ something random

$O(\sqrt{dt})$

drift $\quad dt$

diffusion

diffusion $\quad \sqrt{dt}$

So, what are the rules of calculus?

$$\boxed{dX^2 = dt}$$

Since X is stochastic, so is F, and we can ask 'what is the stochastic differential equation for F?'

$F = F(X)$ $\quad X \to X + dX$

If $F(X) = X^2$ what is the equation for dF?

$\frac{dF}{dX} = 2X \qquad dF = 2X\,dX$

dt

$$F(X+dX) = F(X) + \frac{dF}{dX}dX + \frac{1}{2}\frac{d^2F}{dX^2}dX^2 + \dots$$

$$dF = F(X+dX) - F(X) = \frac{dF}{dX}dX + \frac{1}{2}\frac{d^2F}{dX^2}dt$$

$$dF = \frac{dF}{dX}dX + \frac{1}{2}\frac{d^2F}{dX^2}dt$$

$\boxed{\text{Itô I}}$

If $F = X^2$ is it true that $dF = 2X\ dX$?

No.

- The ordinary rules of calculus do not generally hold in a stochastic environment.

Then what are the rules of calculus?

We are going to throw caution to the wind, pretend that there are no problems or subtleties, use Taylor series. . . and see what happens!

Taylor Series . . . and Itô

$F = X^2$

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

$\frac{dF}{dx} = 2X$

$$F(X + dX) = F(X) + \frac{dF}{dX}dX + \tfrac{1}{2}\frac{d^2F}{dX^2}dX^2,$$

$\frac{d^2F}{dx^2} = 2$

ignoring higher-order terms.

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

differential

dt

S.D.E. for F

$$dF = \frac{dF}{dX}\underset{=}{dX} + \tfrac{1}{2}\frac{d^2F}{dX^2}dX^2$$

This is *almost* correct.

$dF = 2x\,dx + dt$ SDE

for $f = X^2$

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

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

$$F = e^X = \frac{dF}{dX} = \frac{d^2F}{dX^2}$$

$$dF = e^X dX + \frac{1}{2} e^X dt$$

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 dX^2.

Now consider $F = F(t, X_t)$ e.g $t^2 \sin X_t$

$t^n e^{tX_t}$ $\log tX_t$

$t \to t + dt$

$X \to X + dX$

2D T.S.E.:

$$F(t+dt, X+dX) = \overbrace{F(t,X)}^{dF \leftarrow} + \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial X} dX + \frac{1}{2} \frac{\partial^2 F}{\partial X^2} \underbrace{\cancel{dX^2}}_{= dt}$$

$$dF = \underbrace{\left(\frac{\partial F}{\partial t} + \frac{1}{2} \frac{\partial^2 F}{\partial X^2} \right) dt}_{\text{drift}} + \underbrace{\frac{\partial F}{\partial X} dX}_{\text{diffusion}}$$

Itô II

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

$$\boxed{dX^2 = dt.}$$

  and in practice you will get the right result.

Let's get some intuition now, and then shortly we will do Itô's lemma properly!

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

$$\frac{dF}{dX} = 2X \quad \text{and} \quad \frac{d^2F}{dX^2} = 2.$$

Therefore Itô's lemma tells us that

$$dF = dt + 2X\ dX.$$

This is an example of a **stochastic differential equation**.

Starting with Itô II

$$dF = \left(\frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\right)dt + \underline{\frac{\partial F}{\partial x_t} dx_t} \quad \text{rearrange}$$

$$\int_0^t \frac{\partial F}{\partial x_s} dX_s = \int_0^t dF - \int_0^t \left(\frac{\partial F}{\partial s} + \frac{1}{2}\frac{\partial^2 F}{\partial X_s^2}\right) ds \qquad F = F(t, X_t)$$

$$\underbrace{\int_0^t \frac{\partial F}{\partial x_s} dx_s}_{\text{Itô integral}} = F(\underline{t}, x_t) - F(\underline{0}, x_0) - \underbrace{\int_0^t \left(\frac{\partial F}{\partial s} + \frac{1}{2}\frac{\partial^2 F}{\partial x_s^2}\right) ds}$$

If $F = f(X_t)$ then the above integration formula becomes

$$\int_0^t \frac{dF}{dX_s} dX_s = F(X_t) - f(X_0) - \frac{1}{2}\underbrace{\int_0^t \frac{d^2 F}{dx_s^2} ds}$$

Stochastic differential equations

$$\frac{d}{dX}(X) = 1$$

Stochastic differential equations are used to model random quantities, a stock price for example.

They have two parts to them, a **deterministic** and a **random**.

$$\int_0^T X^4\,dX \equiv \int_0^T \frac{dF}{dX_t}\,dX_t = F(X_T) - F(X_0) - \frac{1}{2}\int_0^T \frac{d^2F}{dX_t^2}\,dt$$

$$\frac{dF}{dX_t} = X_t^4 \rightarrow F = \frac{X_t^5}{5} \quad \therefore \quad \frac{d^2F}{dX_t^2} = 4X_t^3$$

$$dX^2 = dt$$

$$= \underbrace{\frac{1}{5}X_T^5 - \frac{1}{5}X_0^5}\; - \underbrace{2\int_0^T X_t^3\,dt}$$

stop.

Suppose we want to model a stock price as a random quantity. Let's use S to denote that stock price.

A stochastic differential equation for S would look something like this:

$$dS = \mu S\, dt + \sigma S\, dX$$

$$dS = \text{ Deterministic } + \text{ Random } .$$

$$\frac{dS}{S} = \mu\, dt + \sigma \underbrace{dX}_{\phi\sqrt{dt}}$$

In words: "The change in the stock price has a predictable component and a random component."

More precisely

$$dS = \overbrace{\text{Something}}^{\text{drift}}\ dt + \overbrace{\text{Something else}}^{\text{diffusion}}\ dX.$$

The randomness is captured by the dX term.

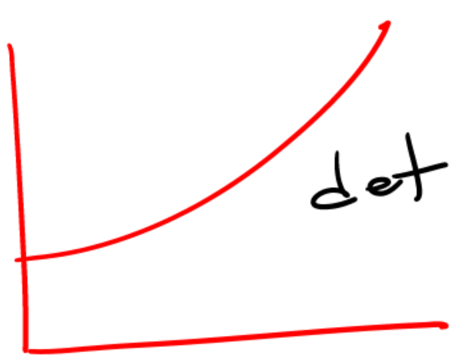
But what are these 'somethings'?

In the standard models they would be functions of S and time, t.

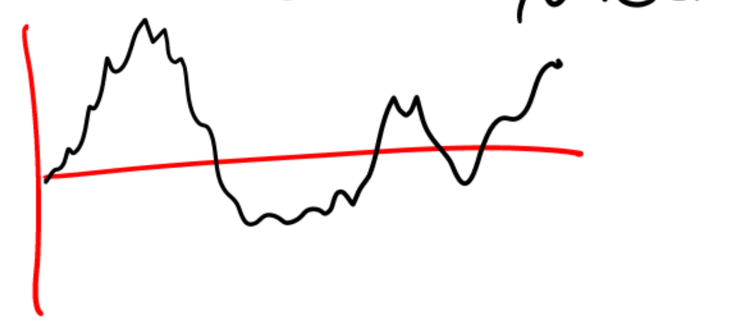
μS σS

$$dS = f(S,t)\,dt + g(S,t)\,dX.$$

The function $f(S,t)$ captures how the predictable bit of the stock model varies with S and t and the $g(S,t)$ function captures the randomness.

① 

\+

② 

③

① + ② → ③

③ → ① & ②

$$dS = f(S,t)\,dt + g(S,t)\,dX.$$

We sometimes call the $f(S,t)$ function the **growth rate** or the **drift**.

The $g(S,t)$ is related to the **volatility** of S.

Some pertinent examples

The first example simple Brownian motion but with a drift:

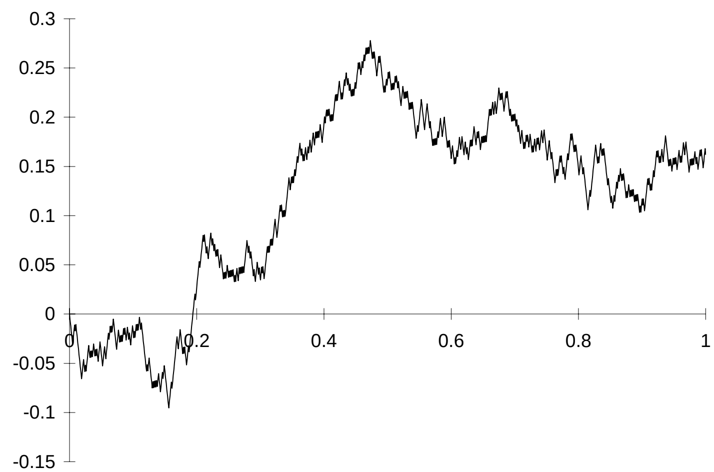
B.M with drift.

$$dS = \underbrace{\mu \, dt}_{\text{drift}} + \sigma \, dX.$$

Generalized Wiener process

$$\boxed{\frac{dS}{S}}$$

$$\frac{S_{i+1} - S_i}{S_i}$$



In this realization S has gone negative.

a $\boxed{G.B.M.}$ lognormal random walk

Our second example is similar to the above but the drift and randomness scale with S:

never hits a negative val?

Switch off randomness $\sigma = 0$:

$$dS = \mu S\, dt + \sigma S\, dX.$$

$\int \frac{dS}{S} = \int \mu\, dt$

$\log S = \mu t + c$

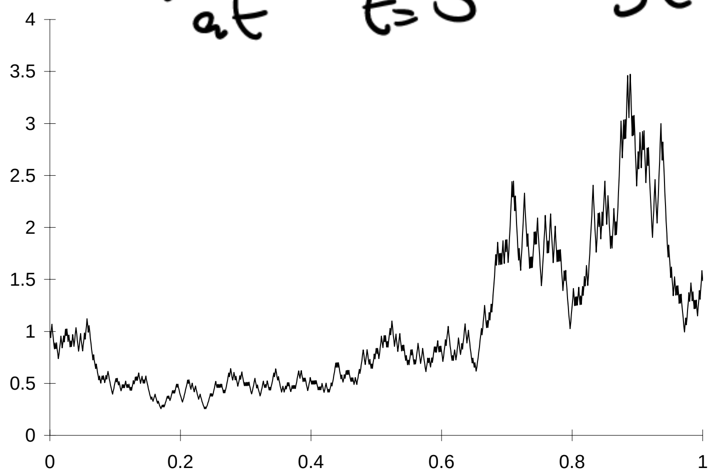
$S(t) = A e^{\mu t}$

$\frac{dS}{S}$

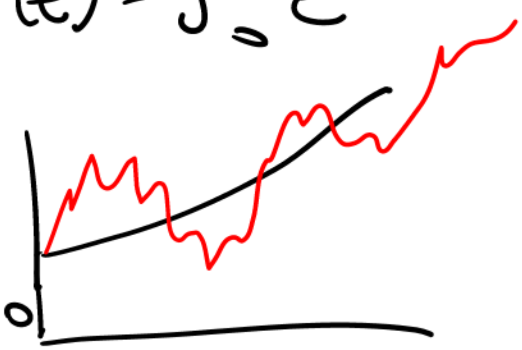
$\mu \to r$

at $t=0$ $S(0) = S_0$ $\therefore A = S_0$

$S(t) = S_0 e^{rt}$



S_0



$\frac{dS}{S} = (\mu - D)dt + \sigma dX$

If S starts out positive it can never go negative; the closer that S gets to zero the smaller the increments dS.

Vasicek

The third example is

ν new

μ mu

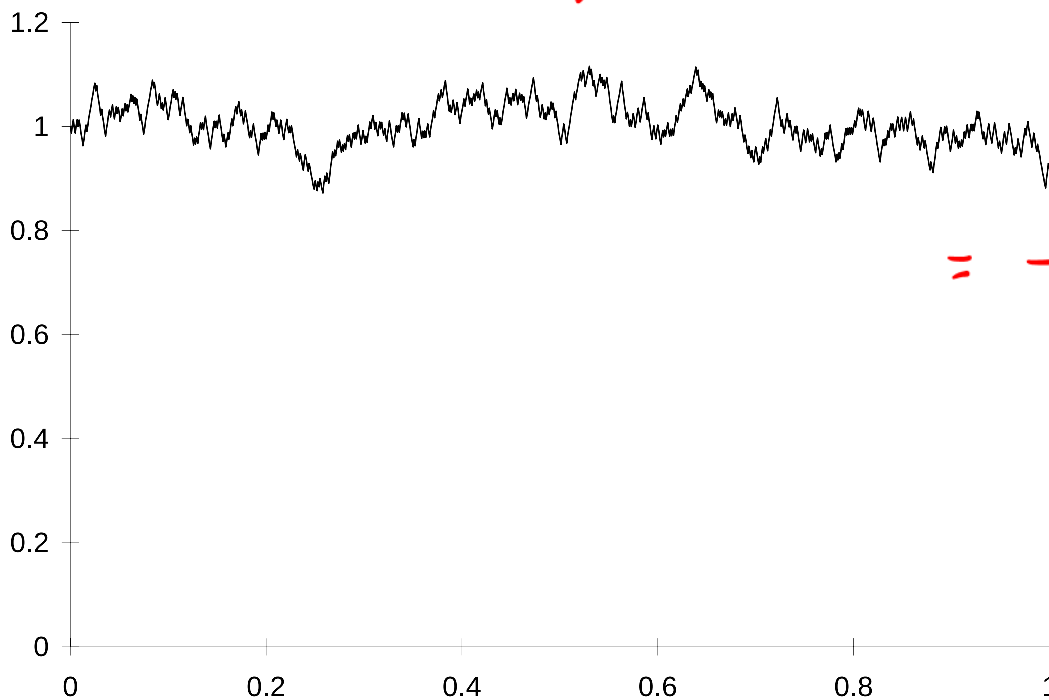
$\frac{\nu}{\mu} = \bar{r}$ mean rate

$$dS = (\nu - \mu S)dt + \sigma\, dX.$$

$dr = (\nu - \mu r)dt + \sigma dX = \mu\left(\frac{\nu}{\mu} - r\right)dt + \sigma dX$

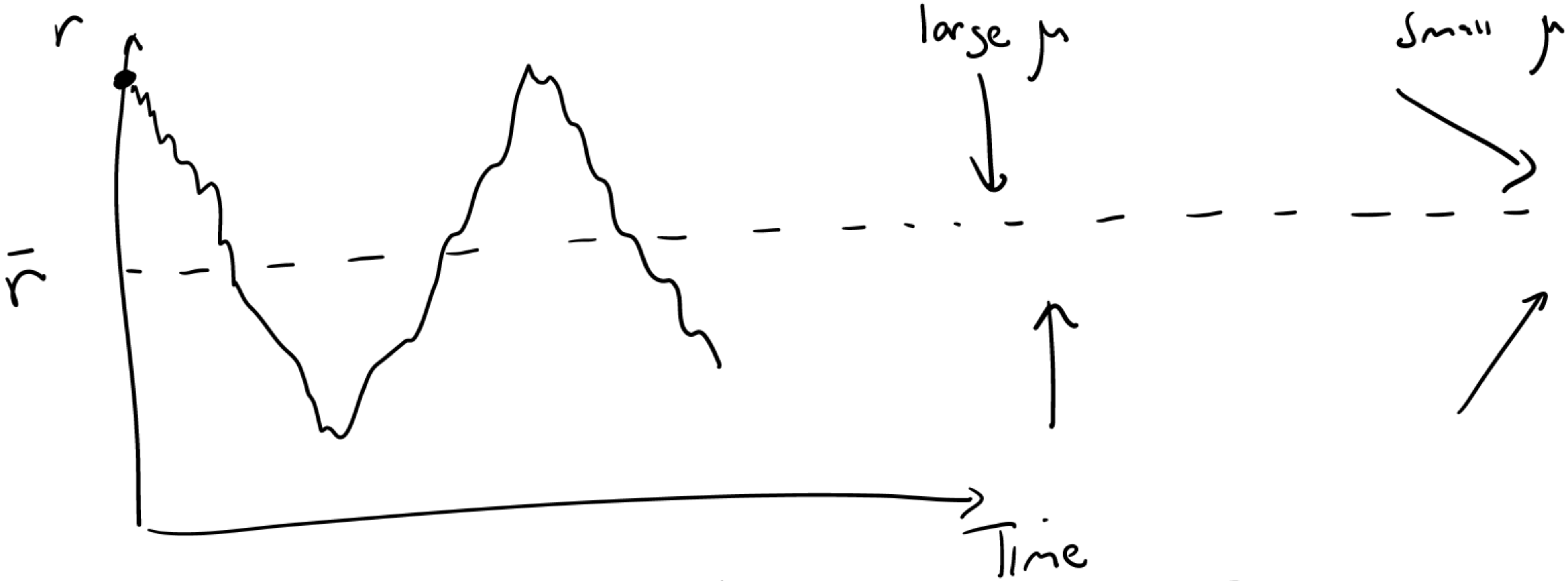
μ - speed

$= \underbrace{-\mu(r - \bar{r})}_{\text{mean reverting}}dt + \sigma dX$



$r > \bar{r}$ -ve (+ve)

$r < \bar{r}$ -ve(-ve)

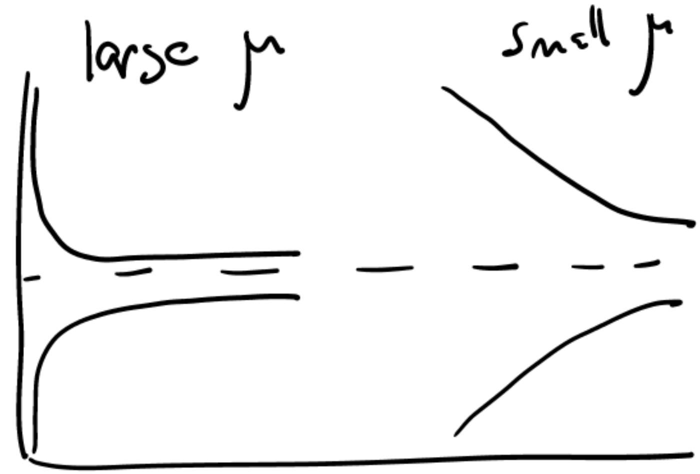


$dr = -\mu(r-\bar{r})dt + \sigma dX$ switch off randomness $\sigma = 0$

$dr = -\mu(r-\bar{r})dt$ $\quad \int \frac{dr}{r-\bar{r}} = -\mu \int dt \quad \log(r-\bar{r}) = -\mu t + c$

$r = \bar{r} + Ae^{-\mu t}$

exponential decay



The random walk

$$dS = (\nu - \mu S)dt + \sigma\, dX$$

is an example of a **mean-reverting** random walk.

If S is large, the negative coefficient in front of dt means that S will move down on average, if S is small it rises on average. There is still no incentive for S to stay positive in this random walk.

With r instead of S this random walk is the Vasicek model for the short-term interest rate.

The final example is similar to the third but we are going to adjust the random term slightly:

Cox-Ingersoll-Ross

As $r \to 0$ randomness switches off and left with $-\mu(-\bar{r})$: $-ve \times -ve$ $+ve$

$$dS = (\nu - \mu S)dt + \sigma S^{1/2}\, dX.$$

$$= -\mu(r - \bar{r})\,dt + \sigma\sqrt{r}\,dX$$

Now if S ever gets close to zero the randomness decreases, perhaps this will stop S from going negative?

This particular stochastic differential equation for S will be important later on, it is the Cox, Ingersoll & Ross model for the short-term interest rate.

Pursuing this idea further, imagine what might be meant by

$\bar{r} = \frac{\nu}{\mu}$

$\boxed{\nu = \bar{r}\mu}$

$$dW = g(t)\,dt + f(t)\,dX.$$

$\nu - \mu\bar{r}$

$\mu(\nu/\mu)$

- Equations like this are called **stochastic differential equations**. Their precise meaning comes, however, from the technically more accurate equivalent stochastic integral.

$G_0 = g$ I.C.

$(*)\quad dG_t = A(t, G_t)\,dt + B(t, G_t)\,dX_t$

This equation above is shorthand for

$\int_0^t dG_s = \int_0^t A(s, G_s)\,ds + \int_0^t B(s, G_s)\,dX_s$

SDE for G_t or RW for dG_t

$$W(t) = \int_0^t g(\tau)\,d\tau + \int_0^t f(\tau)\,dX(\tau).$$

$\{G_t : t \in \mathbb{R}^+\}$

$(**)\quad G_t = g + \int_0^t A\,ds + \int_0^t B\,dX_s$

$(*) \equiv (**)$

The mean square limit

This is useful in the precise definition of stochastic integration.

Examine the quantity

Sequence of random variables

$$\lim_{n\to\infty} E\left[\left(\underbrace{\overbrace{\sum_{j=1}^{n}(X(t_j)-X(t_{j-1}))^2}^{\text{Sequence of random variables}}}_{}-t\right)^2\right]=0$$

$\boxed{dX^2 \sim dt}$

RV limit

where

$$t_j = \frac{jt}{n}.$$

$\mathbb{E}[\ldots]$ sequence of real numbers

$\sum dX^2$

$X^2 \to t$

$\lim_{n\to\infty}$ ordinary limit of real sequence

$dX^2 \to dt$

Let $Y(t_j) = X(t_j) - X(t_{j-1})$

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

$$\underbrace{(Y_1 + \dots + Y_n - t)(Y_1 + \dots + Y_n - t)}$$

This can be expanded as

$$E\left[\sum_{j=1}^{n} (X(t_j) - X(t_{j-1}))^4\right.$$

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

$$\left.-2t\sum_{j=1}^{n}(X(t_j) - X(t_{j-1}))^2 + t^2\right].$$

Since $X(t_j) - X(t_{j-1})$ is Normally distributed with mean zero and variance t/n we have

$$E\left[(X(t_j) - X(t_{j-1}))^2\right] = \frac{t}{n} \qquad dt$$

and

Kurtosis

$$E\left[(X(t_j) - X(t_{j-1}))^4\right] = \frac{3t^2}{n^2}. \qquad 3 \times dt^2$$

Thus the required expectation becomes

$$n\frac{3t^2}{n^2} + n(n-1)\frac{t^2}{n^2} - 2tn\frac{t}{n} + t^2 = O\left(\frac{1}{n}\right).$$

$$\frac{3t^2}{n} + n^2\frac{t^2}{n^2} - n\frac{t^2}{n^2} - 2t^2 + t^2 = \frac{3}{n}t^2 + t^2 - \frac{1}{n}t^2 - t^2 = 2t^2 \cdot \frac{1}{n}$$

As $n \to \infty$ this tends to zero. We therefore say that

$$\boxed{\sum_{j=1}^{n}\left(X(t_j) - X(t_{j-1})\right)^2 = t}$$

in the 'mean square limit.' This is often written, for obvious reasons, as

$$\boxed{\int_0^t (dX)^2 = t.} \longrightarrow \int_0^t dt = t$$

Whenever we talk about 'equality' in the following 'proof' we mean equality in the mean square sense.

Summary

Please take away the following important ideas

- B. M + properties

- Functions of random variables can't be differentiated in quite the same way as functions of deterministic variables.

$$F = F(X) \qquad F = F(t, X)$$

- Instead of using Taylor series you must use Itô's lemma. However, they are very similar and a simple rule of thumb can usually be used to get from Taylor to Itô.

$$\text{Taylor} \xrightarrow{dX^2 = dt} \text{Itô}$$