Fitch Group

CQF Module 1 Lecture 4

Introduction to Stochastic Calculus

Stochastic Calculus and Itô's Lemma

By the end of this lecture you will be able to

- understand where Brownian motion and diffusion processes come from

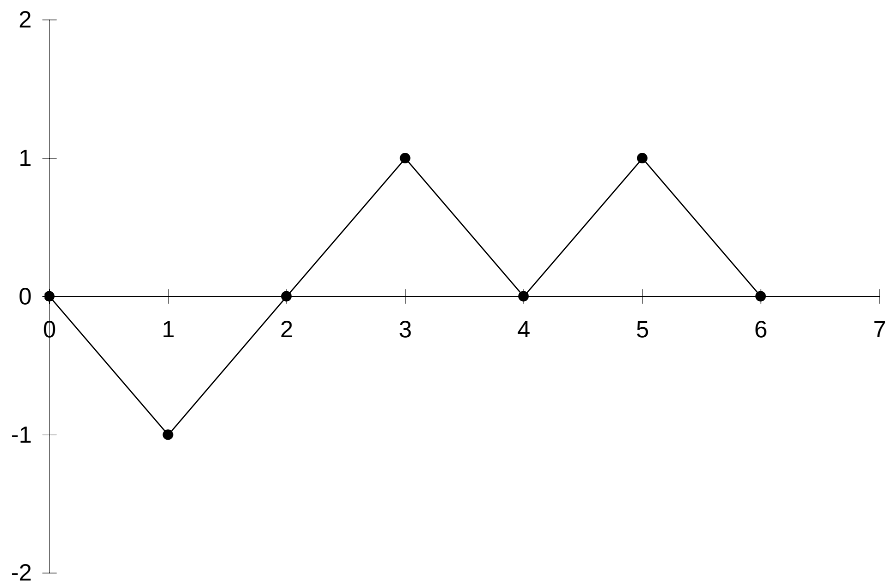
- manipulate functions of random variables

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.



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

$$E[R_i] = 0, \quad E[R_i^2] = 1 \quad \text{and} \quad E[R_i R_j] = 0.$$

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

$$S_i = \sum_{j=1}^{i} R_j.$$

  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?

- This is the **conditional expectation**.

Quadratic variation

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

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

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:

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

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

Change the rules of the coin-tossing experiment slightly.

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 $\sqrt{t/6}$.

The quadratic variation measured over the whole experiment is

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

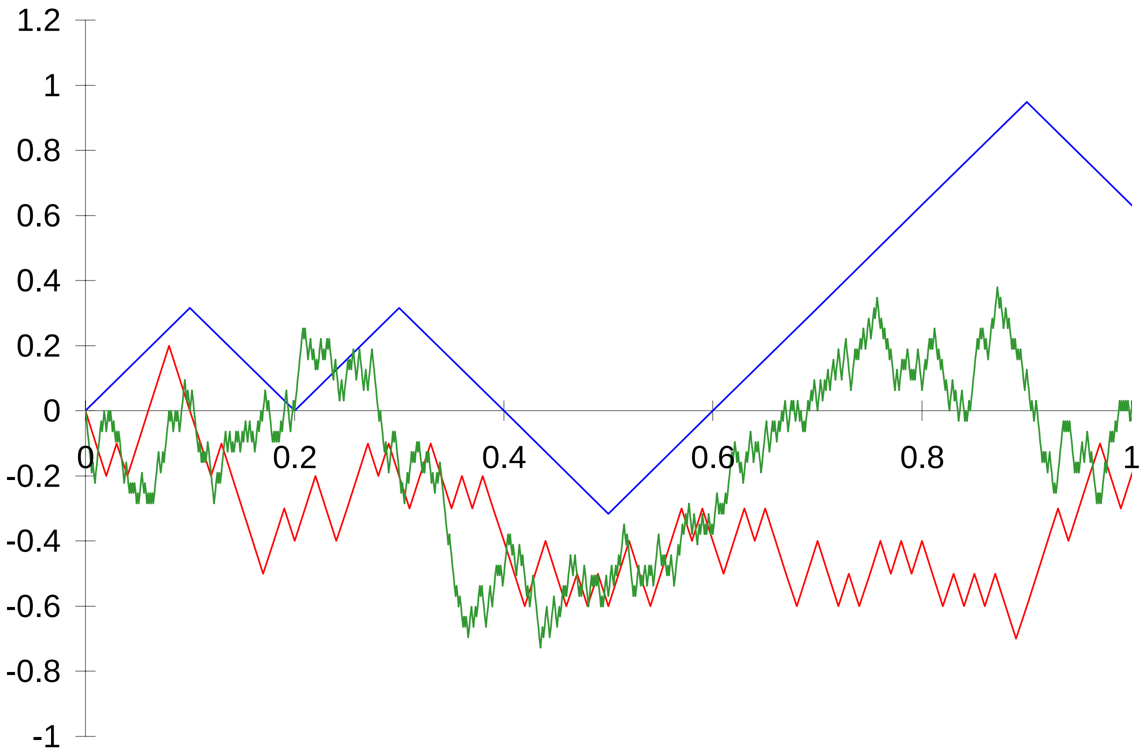
Change the rules again, to speed up the game.

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

The quadratic variation is still

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

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.

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

$$E[S(t)] = 0$$

and a variance

$$E[S(t)^2] = t.$$

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

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

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.

Very Important Notation

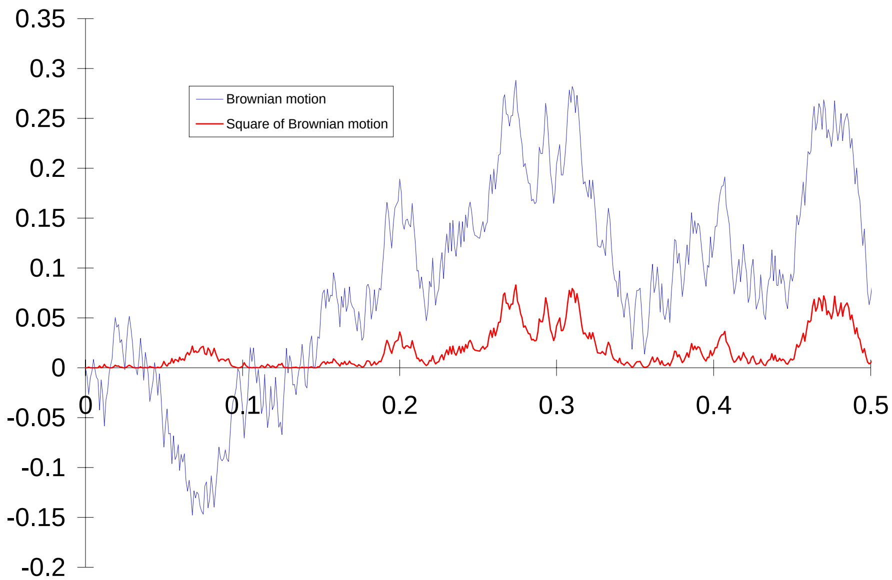
We have seen X as the 'end result' of a random walk, up to some time t.

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 $dt^{1/2}$.

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 $F(X) = X^2$.



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?

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

These are time t and the Brownian motion 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.

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

These take the form

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

So, what are the rules of calculus?

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

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

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

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

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

ignoring higher-order terms.

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

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

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 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.

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.

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

$$dX^2 = dt.$$

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**.

Stochastic differential equations

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**.

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:

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

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

More precisely

$$dS = \text{ Something } \; dt + \text{ Something else } \; dX.$$

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.

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

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.

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

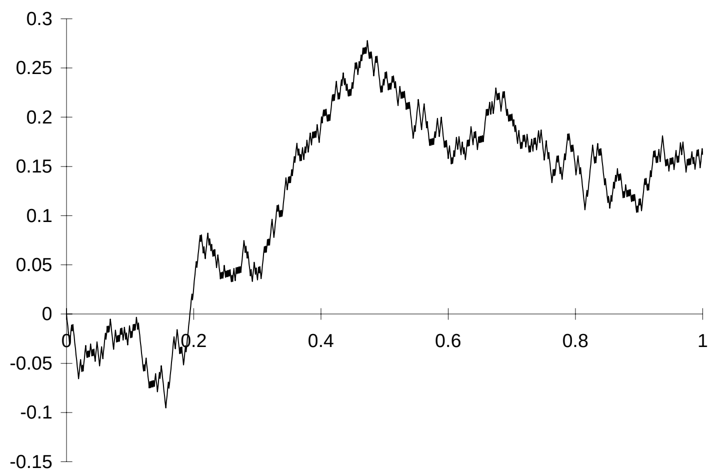
We sometimes call the $f(S,t)$ function the **growth rate** or the **drift**.

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

Some pertinent examples

The first example simple Brownian motion but with a drift:

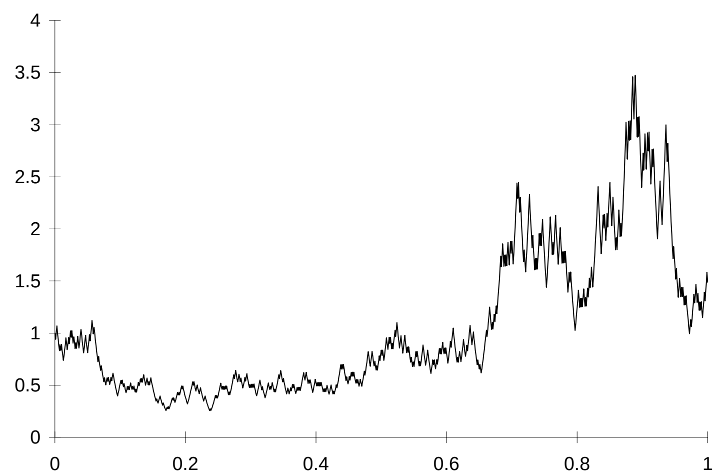
$$dS = \mu \, dt + \sigma \, dX.$$



In this realization S has gone negative.

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

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



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

The mean square limit

This is useful in the precise definition of stochastic integration.

Examine the quantity

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

where

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

This can be expanded as

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

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

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

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

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

and

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

Thus the required expectation becomes

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

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

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

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

$$\int_0^t (dX)^2 = t.$$

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

Slow but accurate

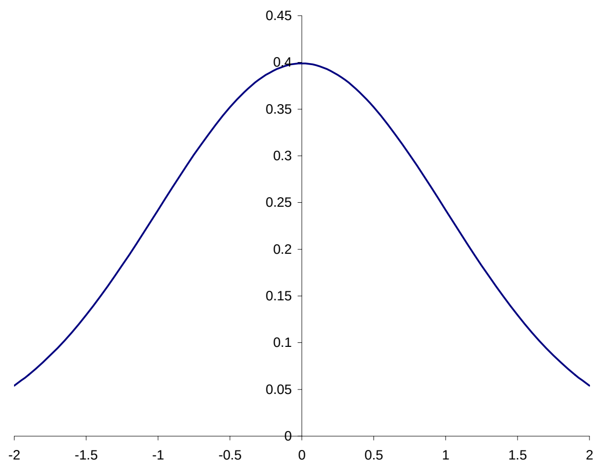
The Excel spreadsheet function `RAND()` gives a uniformly-distributed random variable.

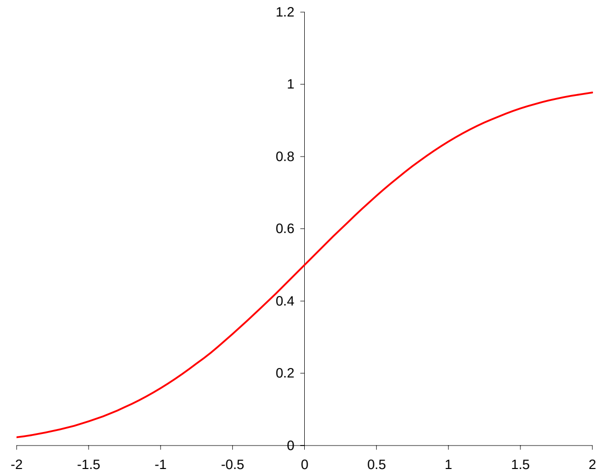
This can be used, together with the inverse cumulative distribution function `NORMSINV` to give a genuinely Normally distributed number:

- `NORMSINV(RAND())`.

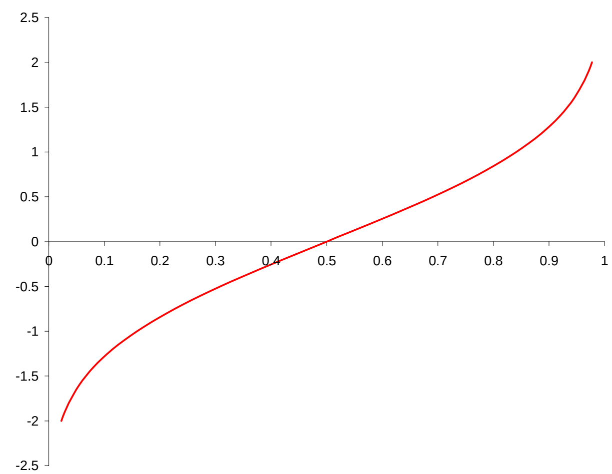
Why does this work?

The pdf and cdf for the Normal distribution





The inverse cumulative distribution function



Fast but inaccurate

An approximation to a Normal variable that is fast in a spreadsheet, and quite accurate, is simply to add up twelve random variables drawn from a uniform distribution over zero to one, and subtract six:

- $$\left(\sum_{i=1}^{12} \mathtt{RAND()}\right) - 6.$$

Why 12?

Any ‘large’ number will do. The larger the number, the closer the end result will be to being normal, but the slower it is.

Why subtract off 6?

The random number must have a mean of zero.

And the standard deviation?

Must be 1.

Summary

Please take away the following important ideas

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

- 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.