# A Brief Intro to Stochastic Processes

Mike Zwecher

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

# Ito Processes Intro

- Ito processes are named for Kiyosi Itô – Wikipedia. They provide a way to extend the toolset of calculus to processes with a random component – stochastic processes.

- All Ito processes have sample paths moving through time that are continuous (roughly meaning if you are drawing them, your pen stays on the paper), but because of the random blips, being continuous doesn't mean they are differentiable.

- Working with processes that are continuous but not differentiable is tricky (to me anyway) – thankfully Ito paved the way.

# How Do We Get to Ito's Lemma?

- In the regular old chain rule, when taking the derivative of a function of a function, e.g., y = f(g(x)), we take the derivative of the outside function, and then the inside function $\frac{df\left(g(x)\right)}{dx} = f'\left(g(x)\right) \cdot g'(x)$ write as $df\left(g(x)\right) = f'\left(g(x)\right)dg(x)$, i.e. define, $g'(x)dx \equiv dg(x)$

- The Mean Value Theorem is also very handy for understanding Ito.

$$f(b) = f(a) + (b-a)f'(c) \; where \; c \in (a, b)$$

where the exact Taylor expansion becomes

$$f(b) = f(a) + (b-a)f'(a) + \frac{1}{2}(b-a)^2 \, f''(c)$$

g

# One more Backgrounder

- Sticking with our MVT / Taylor expansion we have

$$f(b) = f(a) + (b-a)f'(a) + \tfrac{1}{2}(b-a)^2\, f''(c)$$

- Which we re-write as

$$f(b) - f(a) = f'(a)(b-a) + \tfrac{1}{2}f''(c)\,(b-a)^2$$

$$df = \frac{\partial f}{\partial x}dx + \frac{1}{2}\frac{\partial^2 \tilde{f}}{\partial x^2}dx^2$$

The second term is important for Ito, (in deterministic processes, “c” is some intermediate point; where “c” is evaluated in stochastic processes matters a bit) and it requires some hand waving in what follows, but proofs of Ito abound.

# Weiner & Ito Processes

- A Weiner process (Norbert Wiener – Wikipedia) is the fundamental process of Brownian motion

$$\Delta z = \tilde{u}\sqrt{\Delta t} \text{ where } \tilde{u} \sim \mathcal{N}(0,1)$$

And for any two different intervals of time the $\varepsilon$ are independent.

In the limiting continuous case, we write $\mathrm{d}z = \tilde{u}\sqrt{dt}$

Note: $E[dz] = 0, E[dz^2] = dt$

An Ito process is the generalization of the Weiner process (often people interchange the terms when speaking)

$$dx(t) = \mu(x,t)dt + \sigma(x,t)dz(t)$$

# Ito Process Representation

- A univariate Ito process is written as

$$dx(t) = \mu(x,t)dt + \sigma(x,t)dz(t)$$

- Read this loosely as the change in x has a mean change and a random component to the change
- Where

$$dz(t) \equiv = \lim_{\Delta t \to 0} \tilde{u}\sqrt{\Delta t}$$

  - Where, $\tilde{u} \sim \mathcal{N}(0,1)$, $\tilde{u}$ is standard normal random variable, i.e.,

$$\mathrm{E}\big(dz(t)\big) = 0 \; and \; E\left[\big(dz(t)\big)^2\right] = 1 \cdot dt$$

# Rules about Ito Processes

For our Ito process,

$$dx(t) = \mu(x,t)dt + \sigma(x,t)dz(t)$$

There are a few rules:

$$dt^2 = dz(t)dt = 0$$

and

$$dz^2 = dt$$

meaning

$$V[(dx(t)|X_t)] = \sigma(x,t)^2 dt$$

# Ito Processes Imagined

- Imagine watching an asset price through time. As it begins its journey, you watch it move forward. Each day's movement can be decomposed into an average move (often not clear until after some period of observation), and random movement.

- We'll have more to say about this later, but for internal visualization purposes you may think of it in terms of the fundamental driver

$$dP(t) = \mu(p,t)dt + \sigma(p,t)dz(t)$$

or

$$dfwd\_rate(t) = \hat{\mu}(fwd,t)dt + \hat{\sigma}(fwd,t)dz(t)$$

# Ito's Lemma

- Ito's Lemma is the Stochastic Calculus equivalent of the chain rule coupled with the mean value theorem:

- Suppose F(x,t) is at least twice differentiable in x and once differentiable in t.

$$dF(x,t) = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial t}dt + \tfrac{1}{2}\frac{\partial^2 F}{\partial x^2}dx^2$$

$$= \frac{\partial F}{\partial x}dx + \left[\frac{\partial F}{\partial t} + \tfrac{1}{2}\sigma^2(x,t)\frac{\partial^2 F}{\partial x^2}\right]dt$$

Note how familiar this becomes if $\sigma = 0$

# How to Work with Ito’s Lemma

- Working with Ito’s Lemma from here on out involves the process of chug and plug.

- For any function F(x,t), you’ll want to first take the derivatives with respect to x and t, then substitute in the Ito process for dX before simplifying the expression.

- It’s pretty easy to do, plus you get the benefit of being able to lean on a bar and tell people that you work with stochastic differential equations – before watching their eyes glaze over.

# Common Functions E - 1

- Suppose $dX(t) = \mu X dt + \sigma X dz$; Start with the simple $F = aX^2$

Then $\mathrm{dF} = \frac{\partial F}{\partial x} dx + \left[\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 F}{\partial x^2}\right] dt$

Becomes $\mathrm{d}(aX^2) = 2aXdx + \left[\frac{1}{2}\sigma^2 2a\right]dt$

$$= 2aX(\mu X dt + \sigma X dz) + a\sigma^2 dt$$

$$= [2aX^2 + a\sigma^2]dt + 2aX^2\sigma dz$$

# Common Functions E - 2

- Suppose $dX(t) = \mu X dt + \sigma X dz$ And we consider $F = log(X)$

Then $\mathrm{dF} = \frac{\partial F}{\partial x} dx + \left[\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 F}{\partial x^2}\right] dt$

Becomes $\mathrm{dLog}(X) = \frac{1}{X} dx + \left[\frac{\partial F}{\partial t} - \frac{1}{2}\sigma^2 X^2 \frac{1}{X^2}\right] dt$

$$= \frac{1}{X}(\mu X dt + \sigma X dz) + \left[0 - \frac{1}{2}\sigma^2 X^2 \frac{1}{X^2}\right] dt$$

$$= \left[\mu - \frac{1}{2}\sigma^2\right] dt + \sigma dz$$

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

# Common Functions E – 2a (Merton Model)

- Suppose $dX(t) = (\mu X - C)dt + \sigma X dz$ And we consider $F = (X, t)$

Then dF $= \frac{\partial F}{\partial x} dx + \left[\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 F}{\partial x^2}\right] dt$

Becomes dF $= F_x dx + \left[\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 X^2 F_{xx}\right] dt$

$$= F_x\big((\mu X - C)dt + \sigma X dz\big) + \left[F_t + \frac{1}{2}\sigma^2 X^2 F_{xx}\right] dt$$

$$= \left(\frac{1}{2}\sigma^2 X^2 F_{xx} + (\mu X - C)F_x + F_t\right) dt + \sigma X dz$$

# Common Functions E - 3

- Suppose $dx(t) = k(\bar{x} - x)dt + \sigma dz$ (Ornstein-Uhlenbeck)
- And we consider $F = (x - \bar{x})e^{k(t-t_0)}$

Then $\mathrm{dF} = \frac{\partial F}{\partial x}dx + \left[\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 F}{\partial x^2}\right]dt$ becomes

$$\mathrm{d}\left((x - \bar{x})e^{k(t-t_0)}\right) = e^{k(t-t_0)}\,[k(\bar{x} - x)dt + \sigma dz + k(x - \bar{x})dt]$$

$$= \sigma e^{k(t-t_0)}dz$$

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

# Black Scholes in Differential Equation Form part 1/3

- Start with $dS(t) = \mu S dt + \sigma S dz$, a GBM process
- let $C = C(S,t)$ denote a call such that $C_T = max(S_T - K, 0)$
- We create a portfolio $P = C(S,t) - \delta S$, (a Call minus some number of shares we get to choose, i.e., ‘delta’ shares) so that

$$dP = dC(S,t) - \delta dS \equiv dC - \delta dS$$

We’ll use Ito for $dC(S,t)$, and create a hedged portfolio.

$$dP = \frac{\partial C}{\partial S} dS + \left[\frac{\partial C}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right] dt - \delta dS$$

# Black Scholes in Differential Equation Form part 2/3

Choosing $\delta^*$ $such\ that \delta^* dS = \frac{\partial C}{\partial S} dS, i.e., \delta^* = \frac{\partial C}{\partial S},$

The $\frac{\partial C}{\partial S} dS \ - \delta^* \frac{\partial C}{\partial S}$ term equals zero, and $dP$ becomes

$$dP = \left[\frac{\partial C}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right] dt$$

With $dP = \left[\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right] dt$ we have something that depends on the level of S, but not on the movement of $S, dS$ (either component), i.e., $\mu$ has dropped out, and there is no $dz$ risk here either. So now $dP$ represents movement in a risk-free portfolio $P$.

# Black Scholes in Differential Equation Form part 3/3

- As a risk-free portfolio, it can grow at the risk-free rate, and we have

$$dP = rPdt$$

Now let's substitute the terms back to their definitions and finish

$$\left[\frac{\partial C}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right] dt = r\left(C - \frac{\partial C}{\partial S} S\right)$$

This (parabolic differential equation) shows the call option represented by its greeks: theta, gamma, delta; $r$, $\sigma$, and $S$. (note, no $\mu$)

Solving this differential equation is tough, and requires remembering that $C_T = max(S_T - K, 0)$, $C(0, t) = 0$, $C(S, t) \to S - K \; as \; S \to \infty$

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

# Some Popular Processes

- Simple (Arithmetic) Brownian Motion –

$$dX(t) = \mu dt + \sigma dz$$

- Ho-Lee –

$$dX(t) = \mu(t)dt + \sigma dz$$

- Geometric Brownian Motion (Stock prices) –

$$dX(t) = \mu X dt + \sigma X dz$$

- Mean Reverting (Ornstein-Uhlenbeck, Vasicek) –

$$dX(t) = a\big(b - X(t)\big)dt + \sigma dz$$

- Cox Ingersoll Ross –

$$dX(t) = \big(b - aX(t)\big)dt + \sigma\sqrt{X(t)}dz$$

# Multivariate Version of Ito's Lemma

- We start with the one dimensional process:

$$dx(t) = \mu(x,t)dt + \sigma(x,t)dz(t)$$

But we know that many formulas of financial interest involve products, ratios, and other functions of multiple random variables.

So the above expression becomes:

$$dx_i(t) = \mu_i(x_i,t)dt + \sigma_i(x_i,t)dz_i(t)$$

With $cov\left(dz_i, dz_j\right) \equiv \sigma_{i,j} \equiv \rho_{i,j}\sigma_i\sigma_j$

# Multivariate Ito Formula

- With our vector of Ito processes, Ito's Lemma becomes

$$dF(\boldsymbol{x}, t) = \sum_i \frac{\partial F}{\partial x_i} dx_i + \frac{\partial F}{\partial t} dt + \frac{1}{2} \sum_i \sum_j \frac{\partial^2 F}{\partial x_i \partial x_j} \sigma_{ij} dt$$

As an exercise, consider an option $C = C(E, t)$, where $E = SX$ where $S$ is the stock price in its native currency, and $X$ is the exchange rate.

Follow the Black Scholes derivation format, and watch the magic.

# Multivariate Ito – Stochastic Vol

Stochastic Vol Models are an application within the multivariate Ito framework. Combined with dynamics for the underlying, a second process is added expressing the dynamics of volatility e.g.:

Cox Ross Rubinstein -

$$dS(t) = \mu S dt + \sqrt{\sigma} S dz$$
$$d\sigma(t) = \alpha(\sigma, t)dt + \beta(\sigma, t)dw$$
$$dz \cdot dw = \rho dt$$

GARCH -

$$dS(t) = \mu S dt + \sqrt{\sigma} S dz$$
$$d\sigma(t) = a\big(b - \sigma(t)\big)dt + \varepsilon\sigma(t)dz$$
$$dz \cdot dw = \rho dt$$

# Stochastic Alpha, Beta, Rho - SABR

SABR -

$$dF(t) = \sigma(t)\big(F(t)\big)^{\beta} dz$$
$$d\sigma(t) = \alpha\sigma(t)dw$$
$$dz \cdot dw = \rho dt$$

Note: with ethe SABR model, if $\beta$=1, then we're working with lognormal forwards (think of the equity world), and if $\beta$=0, then we're working with a normally distributed process (the HJM-ish rates world). Sometimes $\alpha$ is referred to as the vol of vol.

# Approximating SDEs - Finite Difference

Starting with

$$0 = \left[\frac{\partial C}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right] dt - r\left(C - \frac{\partial C}{\partial S} S\right)$$

We can discretize this, and work backwards remembering our boundary conditions

$C_T = max(S_T - K, 0)$, $C(0, t) = 0$

Now let's discretize the derivatives $\frac{\partial C}{\partial t}$, $\frac{\partial C}{\partial S}$, and $\frac{\partial^2 C}{\partial S^2}$ This will give us most of the pieces to a puzzle we can solve based on the boundary values we know, and the node values we need to figure out.

# Explicit Finite Difference (1)

*Taking some liberty with notation for familiarity with a generic function we are trying to approximate.*

$$\frac{\partial C}{\partial t} \approx \frac{\big(f(time\ t+1, stock\ node\ m) - f(t,m)\big)}{\Delta T}$$

$$\frac{\partial C}{\partial S} \approx \frac{\big(f(t+1, m+1) - f(t+1, m-1)\big)}{2\Delta S}$$

$$\frac{\partial^2 C}{\partial S^2} \approx \frac{\big(f(t+1, m+1) + f(t+1, m-1) - 2f(t+1, m)\big)}{(\Delta S)^2}$$

# Explicit FD 2

*Now we plug the mess on the previous page into our approximation for the fundamental SDE, but note that thankfully some of the "f" terms can be collected.*

This can all be re-written as

$$f(t,m) = \mathrm{a}(m) \times f(t+1,m) + \mathrm{b}(m) \times f(t+1,m+1) + \mathrm{c}(m) \times f(t+1,m-1)$$

Where, some messy arithmetic later, we get

$$\mathrm{a}(m) = \frac{.5(-rm + \sigma^2 m^2)\Delta T}{1 + r\Delta T}$$

$$\mathrm{b}(m) = \frac{(1 - .5\sigma^2 m^2)\Delta T}{1 + r\Delta T}$$

$$\mathrm{c}(m) = \frac{.5(rm + \sigma^2 m^2)\Delta T}{1 + r\Delta T}$$

# Implicit Method

$$\frac{\partial C}{\partial t} \approx \frac{\left(f(time\ t+1, stock\ node\ m) - f(t,m)\right)}{\Delta T}$$

*Same as before*

$$\frac{\partial C}{\partial S} \approx \frac{\left(f(t,m+1) - f(t,m-1)\right)}{2\Delta S}$$

*Note the change from t+1 moves to the time t response in f to m up vs down*

$$\frac{\partial^2 C}{\partial S^2} \approx \frac{\left(f(t,m+1) + f(t,m-1) - 2f(t,m)\right)}{(\Delta S)^2}$$

*Note this change too*

# Implicit Finite Difference Stage 1

*As before, we start with*

$$0 = \left[\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right] dt - r\left(C - \frac{\partial C}{\partial S} S\right)$$

Becomes

$$0 = \left[\frac{\left(f(t+1,m) - f(t,m)\right)}{\Delta T} + \frac{1}{2}\sigma^2 S^2 \frac{\left(f(t,m+1) + f(t,m-1) - 2f(t,m)\right)}{(\Delta S)^2}\right] dt - r\left(C - \frac{\left(f(t,m+1) - f(t,m-1)\right)}{2\Delta S} S\right)$$

Here we have to find $f(t,m+1), f(t,m), and\ f(t,m-1)$.

# IFD Stage 2

As usual, we'll put our unknowns on the left side of the equation, and our knowns on the right.

$$\mathrm{a}(m) \cdot f(t, m-1) + \mathrm{b}(m) \cdot f(t, m) + \mathrm{c}(m) \cdot f(t, m+1) = f(t+1, m)$$

Where

$$\mathrm{a}(m) = .5(rm - \sigma^2 m^2)\Delta T$$

$$\mathrm{b}(m) = (1 + r + .5\sigma^2 m^2)\Delta T$$

$$\mathrm{c}(m) = .5(rm - \sigma^2 m^2)\Delta T$$

This holds for all values, and requires a matrix solution, but remember we know the boundary values at $f(t, 0), f(t, M)$, and all $f(T, m)\ m = 0, \ldots, M$

# Implicit Finite Diff in Matrix Form

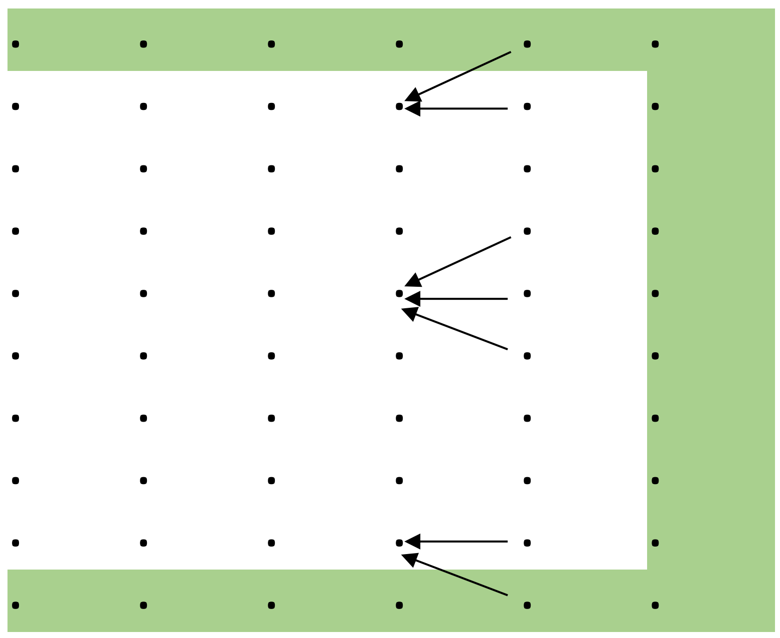
- Since the boundary values $f(t,0), f(t,M)$, and all $f(T,m)\ m = 0, \ldots, M$ are either known from the get go (end-point boundary values) or directly calculable from info (upper and lower boundary), we only need to solve the unknowns for the guts of the grid.

$$AF_t = F_{t+1} + B_t$$

- Where $A = \begin{bmatrix} b_1 & c_1 & 0 & & 0 & 0 \\ a_2 & b_2 & c_2 & \cdots & 0 & 0 \\ 0 & a_3 & b_3 & & 0 & 0 \\ & \vdots & & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & a_{M-1} & b_{M-1} \end{bmatrix}, F_t = \begin{bmatrix} F_{t,1} \\ F_{t,2} \\ F_{t,3} \\ \vdots \\ F_{t,M-1} \end{bmatrix}, B_t = \begin{bmatrix} a_1 f_{t,0} \\ 0 \\ \vdots \\ \vdots \\ 0 \\ c_{M-1} f_{t,M} \end{bmatrix}$

- Noe A is M-1 X M-1 reflecting the fact that we only need to solve the guts of the grid as a simultaneous system. $F_T$ is the known end boundary, and B reflects that top (bottom) boundary values affect the evaluation of the 1st and M-1st F values.

# An Attempt at an Illustration



Meant to show how the arithmetic of the top and bottom interior rows differs slightly from the rest of the interior, necessitating $B_t$