From discrete to continuous time model

Math 485

October 13, 2015

1 Introduction

We have seen that the discrete time binomial model provides some rich features to our modeling of a financial asset. In particular, it allows us to discuss the pricing path dependent derivatives and American options, which would be trivial in the 1 period model. However, the discrete time model is still limiting in a sense: we're only allowed to take action; and the underlying asset can only change value at certain discrete points in time. Between these points, no change and no action can happen.

Of course the situation can be improved by adding more discrete time points. The finer the decision making process (making decisions more often), or the richer the movement of the underlying becomes, the more points we will have to add to the model. Eventually you can see that our model approaches some kind of continuous time process. The question is how can we do it in a tractable way? Let's call our process S_t and suppose it's defined for every t in $[0, T]$ (by approximation). Tractable here means, for one thing, that our approximation should converge. (Not every refining procedure will converge, as you may have learned with a non-Riemann integrable function - the refining procedure there for the integral of such a function does not converge). For another, we want to be able to compute the distribution of S_t at any time t, as well as the joint distribution of $S_{t_1}, S_{t_2}, \cdots, S_{t_n}$ for any time points $t_1, t_2, \cdots, t_n$. Beyond this we would like to know as much about S_t as possible. The key tool for us is the CLT.

2 The approximating discrete time binomial model

As the introduction section alluded to, we need to do our approximation in the right way. For example, observe that if $S_N = S_0 X_1 X_2 \cdots X_N$ then

$$\log(S_n) = \log(S_0) + \sum_{k=1}^{N} \log(X_k).$$

Thus

$$\begin{aligned} E(\log(S_N)) &= \log(S_0) + NE(\log(X_1)); \\ Var(\log(S_N)) &= NVar(\log(X_1)). \end{aligned}$$

As we add more time points (increasing N) we would want to keep the $E(\log(S_N))$ and $Var(\log(S_N))$ constant. After all, it is the distribution of the asset at the terminal time $T = N\delta T$. Also, the spirit of CLT is that if we have a sequence of partial sum of i.i.d random variables, *scaled in such a way that ther expectation and variance remains constant*:

$$\begin{aligned} E\left(\frac{\sum_{i=1}^{n} X_i - \mu}{\sigma\sqrt{n}}\right) &= 0; \\ Var\left(\frac{\sum_{i=1}^{n} X_i - \mu}{\sigma\sqrt{n}}\right) &= 1. \end{aligned}$$

then the sequence of partial sum will converge.

This motivates us to model X_i as followed: for a fixed μ, σ

$$X_i = e^{\mu\Delta T + \sigma\sqrt{\Delta T}\xi_i},$$

where $P(\xi_i = 1) = P(\xi_i = -1) = 1/2$, and the ξ_i are independent.

Note that this is just another way to express our previous discrete time model, with

$$\begin{aligned} u &= e^{\mu\Delta T + \sigma\sqrt{\Delta T}}; \\ d &= e^{\mu\Delta T - \sigma\sqrt{\Delta T}}. \end{aligned}$$

Given u, d, we can solve for μ, σ such that the above system is satisfied. Note also that the no arbitrage condition requires:

$$\mu\Delta T - \sigma\sqrt{\Delta T} < r\Delta T < \mu\Delta T + \sigma\sqrt{\Delta T}.$$

Also note that the probability given above is *the real world probability, not the risk neutral probability.* The choice of 1/2 may seem a bit arbitrary, but keep in mind that the real world probability is irrelevant for pricing. We only need to build a model with rich enough structure to capture the up and down movement of the asset. The choice of 1/2 helps us to prove the convergence of the model in an easier way. Also note that we can even make μ and σ be random variables, if you feel adventurous. For now we'll just keep them constants.

Lastly we want to verify that the expectation and variance of $\log(S_N)$ is constant:

$$\begin{aligned} E(\log(S_N)) &= \log(S_0) + NE(\log(X_1)) = \log(S_0) + \mu N \Delta T = \log(S_0) + \mu T; \\ Var(\log(S_N)) &= NVar(\log(X_1)) = \sigma^2 N \Delta T = \sigma^2 T. \end{aligned}$$

3 Converging to the continuous model

We will now denote S_N as S_T (since $T = N\Delta T$). From the previous section, we have

$$\begin{aligned} \log(S_T) &= \log(S_0) + \mu T + \sum_{i=1}^{N} \sigma \sqrt{\Delta T} \xi_i \\ &= \log(S_0) + \mu T + \sigma \sqrt{T} \frac{\sum_{i=1}^{N} \xi_i}{\sqrt{N}}. \end{aligned}$$

Let $N \to \infty$, by the CLT, we see that S_T has distribution

$$\log(S_T) = \log(S_0) + \mu T + \sigma W_T,$$

where W_T has distribution $N(0,T)$ (one can represent $W_T = \sqrt{T} N(0,1)$ as well).

By a similar argument, partitioning the interval $[t,T]$ into N sub-interals, and let $N \to \infty$ we also have

$$\log(S_T) = \log(S_t) + \mu(T-t) + \sigma W_{T-t},$$

where W_{T-t} has $N(0, T-t)$ distribution.

Moreover, since the increments ξ_i are independent, it can be shown that W_{T-t} is indendent of W_t (where W_t is the RV we get for partitioning the interval $[0,t]$ and let $N \to \infty$). In fact, by the same argument, you can see that W_{T-t} is indendent of $W_r, 0 \le r \le t$. This is the so called independent increment propety we'll discuss later.

Thus, in summary, one can say that our continuous model, derived from the limit of the discrete binomial model, satisfies the following properties in distribution:

a. For any $s < t$

$$S_t = S_s e^{\mu(t-s)+\sigma W_{t-s}},$$

W_{t-s} has $N(0, t-s)$ distribution, W_{t-s} is indendent of $W_r, 0 \le r \le s$.

b. In particular,

$$S_t = S_0 e^{\mu t+\sigma W_t}.$$

We say S_t has log normal distribution (the log of S_t has Normal distribution).

4 The dynamics of S_t

4.1 Evolution in discrete time

In the discrete time model, the evolution of S_n is clear. At the next time point $n+1$, to get the value of S_{n+1}, we just have to multiply S_n with X_{n+1}, where $X_{n+1} = e^{\mu\Delta T+\sigma\sqrt{\Delta T}\xi_{n+1}}$ as described above. In other words

$$S_{n+1} = S_n e^{\mu\Delta T+\sigma\sqrt{\Delta T}\xi_{n+1}}.$$

Equivalently, at the next time point $n+1$, to get the value of $\log(S_{n+1})$, we just have to add $\log(S_n)$ with $\mu\Delta T + \sigma\sqrt{\Delta T}\xi_{n+1}$. We say

$$\Delta \log(S_n) := \log(S_{n+1}) - \log(S_n) = \mu\Delta T + \sigma\sqrt{\Delta T}\xi_{n+1}.$$

This is a discrete recursion equation that specifies the dynamics of $\log(S_n)$. We would like to get an equation for the dynamics of S_n:

$$\Delta S_n := S_{n+1} - S_n = ?$$

But in our discrete time model, this won't be anything nice. You'll see that the situation is different in the continuos time.

4.2 Evolution in continuous time

In the continuous time context, ΔT becomes dt (dt is not a real physical quantity, it's a differential and only makes sense within an integral sign. Nevertheless, it can be used to specify the dynamics of a process in time, as long as we are clear on what we mean by using dt).

Formally, we have

$$S_{t+dt} = S_t e^{\mu dt + \sigma\sqrt{dt}\xi_t},$$

where ξ_t takes values ± 1 with probability $1/2$ each. Apply Taylor's expansion on the exponential term, we get

$$\begin{aligned} S_{t+dt} &= S_t(1 + \mu dt + \sigma\xi_t\sqrt{dt} + \frac{1}{2}\sigma^2\xi_t^2 dt + \text{higher order terms}) \\ &= S_t(1 + [\mu + \frac{1}{2}\sigma^2]dt + \sigma\xi_t\sqrt{dt} + \text{higher order terms}) \\ &= S_t + S_t[\mu + \frac{1}{2}\sigma^2]dt + \sigma S_t\xi_t\sqrt{dt} + S_t\text{higher order terms}). \end{aligned}$$

since $\xi_t^2 = 1$. What we mean by the above is, provide we can make sense of the intergration, we have for $s < t$

$$S_t = S_s + \int_s^t [\mu + \frac{1}{2}\sigma^2]S_u du + \int_s^t \sigma S_u \xi_u \sqrt{du} + \int_s^t \text{higher order terms}.$$

By higher order terms, we mean terms of order higher than $dt^{3/2}$. As we will explain later, the intergral of the form

$$\int_s^t \text{higher order terms} = \int_s^t O(dt^{3/2}) = 0.$$

Also the integral needs to be explained, because of the term

$$\int_s^t \sigma S_u \xi_u \sqrt{du}.$$

As we'll also explain, $\int O(\sqrt{dt}) = \infty$. Thus the term $\int_s^t \sigma S_u\xi_u\sqrt{du}$ is undefined. However, S_t is defined (we got it as a limit of convergence of the discrete model, and we have a distribution for S_t). Thus there must be a way to define $\int_s^t \sigma S_u\xi_u\sqrt{du}$. As you'll see, this will lead to the definition of Brownian motion and Ito Calculus.

The bottomline is, provided we can make sense of these technical details, we have arrived at the dynamics of S_t as we wished for in the continuous time, using the Taylor's expansion:

$$dS_t := S_{t+dt} - S_t = S_t[\mu + \frac{1}{2}\sigma^2]dt + \sigma S_t \xi_t \sqrt{dt},$$

where we thow away the higher order terms since it disappears in the integral, which is the rigorous sense we want to give to the above dynamical equation anyway.

5 Prelude to Brownian motion and Ito Calculus

5.1 Brownian motion

As you can observe from the previous sections, the term that makes S_t a random variable is W_t in

$$S_t = S_0 e^{\mu t + \sigma W_t},$$

or W_{t-s} is

$$S_t = S_s e^{\mu(t-s) + \sigma W_{t-s}},$$

or ξ_u in

$$S_t = S_s + \int_s^t [\mu + \frac{1}{2}\sigma^2] S_u du + \int_s^t \sigma S_u \xi_u \sqrt{du}.$$

Indeed, there is reason to believe these are all different forms of one single process, let's call it W_t where corresponding to the above we have

$$\begin{aligned} W_t &= W_t \\ W_{t-s} &= W_t - W_s \\ \sqrt{du}\xi_u &= dW_u. \end{aligned}$$

Let's see what we have learned about this process W_t so far:

a. $W_t - W_s$ is independent of $W_r, 0 \leq r \leq s$.

b. $W_t - W_s$ has $N(0, t-s)$ distribution.

Surprisingly, these two characteristics are enough to specify a unique stochastic (random) process called the Brownian motion. As typical in mathematics, now that we have the intuition, we'll take the reverse approach and *define* Brownian motion as a process satisfying properties a and b. We then *build* a model for the underlying S_t out of this Brownian motion W_t. The question is: how do we build S_t?

5.2 Ito Calculus

The short answer to the question how to build S_t is to specify its dynamics:

$$\begin{aligned} dS_t &= S_t \mu dt + \sigma S_t dW_t, \\ S_0 &= x. \end{aligned}$$

Then we will find, if possible, a process S_t that has the above dynamics. Yet this involves several significant difficulties. First of all, the above is a *Stochastic Differential Equation* (SDE). Equation because S_t appears on both sides of the equality (recall Ordinary Differental Equation (ODE) and what makes it a differntial equation). Stochastic because the equation relates random quantities on both sides. Differential because it involves dt and dWt.

Even ignoring the issue that we are dealing with a SDE, there is even a more fundamental issue: what do we mean by

$$\int_0^t f(u)dW_u,$$

for a (possibly random) process $f(u)$. The reason is we have seen in some sense

$$\int_0^t f(u)dW_u = \int_s^t f(u)\xi_u\sqrt{du},$$

and integrating with respect to the term $\sqrt{du}$ has to be interpreted in a special way. Being able to give a sense to integrating with dWu is one fundamental result in the Ito Calculs: the so-called Ito's integral.

Similar to classical Calculus (but in some sense developing things in the reverse order), after we have a notion of the integral, we want to have notion of differentiation. That is, what is the "derivative" with respect to t (to time) of a term like

$$S_t = \int_0^t \alpha(u)d_u + \int_0^t \sigma(u)dW_u?$$

In some sense, the answer has been given above, the "derivative" is given in terms of the differential:

$$dS_t = \alpha(t)dt + \sigma(t)dW_t.$$

The reason we don't have a proper derivative with respect to t is because W_t is NOT differentiable in t, another characteristic property of Brownian motion.

Then what about the "derivative" of a function os S_t, say S_t^2? This is to ask for the chain rule for the stochastic calculus. The chain rule in this case is referred to as Ito's formula.

Lastly, how can we solve a SDE? After having the chain rule, we can develop some basic techniques to find the explicit solution for some basic form of SDEs, in which the equation

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

is included. S_t in this case will be referred to as Geometric Brownian motion.

6 The Lebesgue-Stiltjes integral

6.1 The Riemann integral

Suppose we have a partition of the interval $[0, T]$: $0 = t_0 < t_1 < \cdots < t_n = T$ and a function $f(t)$ on $[0, T]$ that has nice properties. Define $\|\Delta\| = \max_i (t_{i+1} - t_i)$ as the mesh of this particular partition. Then

$$\sum_{i=1}^{n} f(t_i)(t_{i+1} - t_i) \to \int_0^t f(s) ds,$$

as $\|\Delta\| \to 0$.

Additionally, if we try the following:

$$\sum_{i=1}^{n} f(t_i)(t_{i+1} - t_i)^{1+\varepsilon}$$

or

$$\sum_{i=1}^{n} f(t_i)(t_{i+1} - t_i)^{1-\varepsilon}$$

for some $\varepsilon > 0$ you'll see that

$$\sum_{i=1}^{n} f(t_i)(t_{i+1} - t_i)^{1+\varepsilon} \to 0$$

and

$$\sum_{i=1}^{n} f(t_i)(t_{i+1} - t_i)^{1-\varepsilon} \to \infty$$

as $\|\Delta\| \to 0$. The intuitive reason is because

$$\sum_{i=1}^{n} f(t_i)(t_{i+1} - t_i)^{1+\varepsilon} = \sum_{i=1}^{n} f(t_i)(t_{i+1} - t_i)(t_{i+1} - t_i)^{\varepsilon}$$

and since

$$\sum_{i=1}^{n} f(t_i)(t_{i+1}-t_i) \to \int_0^t f(s)ds,$$

$$(t_{i+1}-t_i)^{\varepsilon} \to 0,$$

the product converges to 0. Similarly,

$$\sum_{i=1}^{n} f(t_i)(t_{i+1}-t_i)^{1-\varepsilon} = \sum_{i=1}^{n} \frac{f(t_i)(t_{i+1}-t_i)}{(t_{i+1}-t_i)^{\varepsilon}}$$

which goes to infinity.

Formally we write

$$\begin{aligned} \int_0^t f(t)(dt)^{1+\varepsilon} &= 0 \\ \int_0^t f(t)(dt)^{1-\varepsilon} &= \infty, \end{aligned}$$

where

$$\begin{aligned} \int_0^t f(t)(dt)^{1+\varepsilon} &= \lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(t_i)(t_{i+1}-t_i)^{1+\varepsilon} \\ \int_0^t f(t)(dt)^{1-\varepsilon} &= \lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(t_i)(t_{i+1}-t_i)^{1-\varepsilon} \end{aligned}$$

In particular,

$$\begin{aligned} \int_0^t f(t)(dt)^2 &= 0 \\ \int_0^t f(t)\sqrt{dt} &= \infty, \end{aligned}$$

6.2 Some examples

Example 6.1.

$$\lim_{n\to\infty} \sum_{i=1}^{n} \frac{i}{n^2} = \int_0^1 tdt = \frac{t^2}{2}\Big|_0^1 = \frac{1}{2},$$

where we see the sum on the LHS as the function $f(t) = t$ evaluated at the grid point $\frac{i}{n}$ and $t_{i+1} - t_i = \frac{i+1}{n} - \frac{i}{n} = \frac{1}{n}$.

Note that the sum above can be checked straightforwardly by using the identity

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}.$$

The point is this simple example allows us to verify that

$$\begin{aligned} \int_0^1 t\sqrt{dt} &= \infty \\ \int_0^1 t(dt)^2 &= 0 \end{aligned}$$

since

$$\begin{aligned} \lim_{n\to\infty} \sum_{i=1}^{n} \frac{i}{n}\frac{1}{n^2} &= \lim_{n\to\infty} \frac{n(n+1)}{2n^3} = 0 \\ \lim_{n\to\infty} \sum_{i=1}^{n} \frac{i}{n}\frac{1}{\sqrt{n}} &= \lim_{n\to\infty} \frac{n(n+1)}{2n\sqrt{n}} = \infty. \end{aligned}$$

6.3 The Lebesgue-Stiltjes integral

You may ask why do we even care about expression like $\int_0^t f(s)\sqrt{ds}$ or $\int_0^t f(s)(ds)^2$, since we never see them in calculus. This is true, because all we've dealt with there were Riemann integral. However, we can generalize the notion of integration in the following way: let $g(x)$ be a function defined on $[0, T]$ with nice property. Then we can define

$$\int_0^T f(s)dg(s) := \lim_{\|\Delta\|\to 0} \sum_{i=1}^{n} f(t_i)(g(t_{i+1}) - g(t_i)),$$

if the limit on the RHS exists. This is called the Lebesgue-Stiltjes integral of f against g.

And since now we deal with general function g, you can see that there are such functions that when $t_{i+1} - t_i$ is small

$$g(t_{i+1}) - g(t_i) \approx \sqrt{t_{i+1} - t_i}.$$

(Such function g is not one of your classical calculus examples. At least if you look among the differentiable functions you won't find one. The reason is if g is differentiable, then first order approximation tells us that

$$g(t_{i+1}) - g(t_i) \approx g'(t_i)(t_{i+1} - t_i),$$

so it's of order $O(dt)$, not $O(\sqrt{dt})$.)

Our discussion above shows us that for such function g the Lebesgue-Stiltjes integral for f against g does not exist. As you will see, the Brownian motion paths give us such an example of a function g, which makes it necessary to define the Ito's integral.

6.4 An example

Actually if you have done u-substitution in Calculus, you have performed the Lebesgue-Stiltjes integral (possibly without realizing it). Consider the following example:

$$\int_0^1 2xe^{x^2}dx = \int_0^1 e^u du,$$

where you made the substitution

$$\begin{aligned} u &= x^2 \\ du &= 2xdx. \end{aligned}$$

More explicitly you consider u as a function of x:

$$\begin{aligned} u(x) &= x^2 \\ du/dx &= 2x. \end{aligned}$$

But another way to write this is you're evaluating the integral

$$\int_0^1 e^{x^2}du(x) = \int_0^1 e^{x^2}u'(x)dx = \int_0^1 e^{x^2}2xdx.$$

That is you integrate e^{x^2} with respect to $u(x) = x^2$ over the interval $[0, 1]$.

6.5 The Lebesgue-Stiltjes integral as an example of portfolio value

You may ask when we need to use the Lebesgue-Stiltjes integral. Consider the following example.

Let $0 = t_0 < t_1 < ... < t_n = T$ be a partition of $[0, T]$

Consider an investor who invests in an underlying asset S and the saving account such that the portfolio is self-financing. Let $\pi_k = \pi_{t_k}$ be the value of the portfolio and Δ_k be the number of shares of S he holds at time k. Then

$$\pi_{k+1} = \Delta_k S_{k+1} + e^{r\Delta T}(\pi_k - \Delta_k S_k).$$

We will replace $e^{r\Delta T}$ with $1 + r\Delta T$, i.e. continuous compounding with discrete compouding. They should be very close, if ΔT is small. Then the self-financing equation reads

$$\pi_{k+1} = \Delta_k S_{k+1} + (1 + r\Delta T)(\pi_k - \Delta_k S_k),$$

or

$$\begin{aligned}
\pi_{k+1} &= \pi_k + \Delta_k(S_{k+1} - S_k) + r\Delta_T(\pi_k - \Delta_k S_k) \\
&= \pi_k + \Delta_k(S_{k+1} - S_k) + y_k r(t_{k+1} - t_k) \\
&= \sum_{i=1}^{k} \Delta_i(S_{i+1} - S_i) + y_i r(t_{i+1} - t_i).
\end{aligned}$$

where y_k is the amount of cash we holds at time k. Thus you see that if we consider $\Delta(t), y(t)$ as a function of t, self-financing requiring that $\Delta(t) + y(t) = \pi(t)$, letting $\|\Delta\| \to 0$ we get

$$\pi_t = \int_0^t \Delta_u dS_u + \int_0^t y_u r du.$$

Thus the amount of money we get from investing in the stock in the continuous time is a Lebesge-Stiltjes integral.

Note: if you replace $y_u = \pi_u - rS_u$, then the above equation reads

$$\pi_t = \int_0^t \Delta_u(dS_u - rS_u du) + \int_0^t \pi_u r du,$$

which has the following interpretation:

When you invest in a risky asset (S) in such a way that your portfolio is self-financing, your gain can be decomposed in two components: the deterministic component, which is just the saving account: $\int_0^t \pi_u r du$. The other component is how the underlying asset performs versus the saving account: if it performs better: $dS_u > rS_u du$ then your portfolio will perform better than the traditional saving. If it performs worse: $dS_u < rS_u du$ then your portfolio will perform worse than the traditional saving.