Introduction to Simulation

Goals:

- Review where we are in the course
- Introduce the final core mathematical technique we will learn: simulation
- Discuss the steps we need to take in order to solve problems via simulation

Relevant literature:

- Hirsa, Chapter 6; Glasserman, Chapter 1

Review: Comparison of Greeks in Heston vs. Black-Scholes model

- In your last homework, you saw that Greeks in the Heston model can diverge from the traditional Black-Scholes Greeks.
  - Why is this, and which should we trust?
- In reality, the comparison is challenging conceptually as we should not generally expect parameters to have a one to one mapping from one model to another.
- Adding another level of complexity is the fact that the asset price and its volatility depend on each other by definition in a stochastic volatility model.
- This means that if we shift one, we might be implicitly shifting the other. How do we make sure our Greeks incorporate this?

Comparison of Greeks in Heston vs. Black-Scholes model: Delta

- Δ is an exception as the definition of a shift in the underlying is unambiguous and comparable regardless of model. Yet even in this case, the model Δ's were different.
- The fundamental reason is the interaction between an asset's price and its volatility in stochastic volatility model.
- To be precise, let's define Δ as the sensitivity to changes in the underlying asset leaving all other parameters, including volatility, constant.
- That is, we perform the following transformation:

$$S_0 = S_0 + \delta S \quad (1)$$

$$\sigma_{\text{impl}} = \sigma_{\text{impl}} \quad (2)$$

where δ is the shift amount.

Comparison of Greeks in Heston vs. Black-Scholes model: Delta

- If we are calculating the Δ of a call option, $C(S_0, \sigma_{\text{impl}})$, then the chain rule tells us that the sensitivity will be:

→ $\rho = 0$ or BS model

terms $= 0$

$$\Delta = \frac{\partial C}{\partial S} + \underbrace{\frac{\partial C}{\partial \sigma_{\text{impl}}} \frac{\partial \sigma_{\text{impl}}}{\partial S}} \tag{3}$$

- In the Black-Scholes model, the second term is zero and can be ignored since the implied volatiliy does not depend on the asset price. This will not be true in other models, however, in particular those with stochastic volatility if $\rho \neq 0$

Comparison of Greeks in Heston vs. Black-Scholes model: Delta

- This leads to the concept of multiple Δs, those that are standard and those that are smile-adjusted. By smile-adjusted, we mean those that account for the correlation between spot and volatility when adjusting the spot and volatility parameters.

- In particular, a smile-adjusted delta would create a scenario consistent with the stochastic volatility

$$S_0 = S_0 + \delta S \quad (4)$$

$$\sigma_{\text{impl}} = \sigma_{\text{impl}} + \delta_S \sigma_{\text{impl}} \quad (5)$$

where δ_S is defined by ρ and the particular SDE we are working with.

Comparison of Greeks in Heston vs. Black-Scholes model: Vega

- For vega, we need to agree on what a shift in volatility means.
  - In Black-Scholes, it is clearly defined as the sensitivity to the option's implied volatility.
  - In the case of Heston, there is no single volatility parameter:
    * θ can be thought of as a long-run equilibrium volatility.
    * ν_0 can be thought of as a short-term volatility.
    * Consequently, we can think of moving both together as a **parallel shift** of the volatility surface.
    * The alternative would be to bump all the input volatilities and re-calibrate our Heston model parameters. This would also measure change to a parallel vol. shift.
- Even once we agree upon the definition of a shift in volatility, the same argument as we used for delta applies.

Review: What techniques have we covered so far?

- Quadrature
- Fourier Transform Techniques
- Numerical Solutions to PDE's
- Optimization and Calibration

Review: What can we do with these techniques

- Quadrature
  - Price European options with known density functions.
  - Price multi-dimensional exotic options under certain model assumptions.
- Fourier Transform Techniques
  - Price European options with more complex stochastic processes.
- Numerical Solutions to PDE's
  - Price path-dependent & American options
- Optimization and Calibration
  - Enables us to find the optimal set of model parameters for a given set of market data.

Review: What's left?

- So far, the only technique that we've studied that deals with path dependent options is numerical methods for PDE's.
- Another important technique for solving these path dependent problems is **simulation**
- For many path dependent securities, simulation and PDE's will both produce reliable solutions and it is a matter of preference which technique we apply.
- In some cases, PDE solutions will become intractable and we will be forced to use simulation.
  - Multi-Asset Path Dependent Options
- In other cases, such as exercise decisions for American options, simulation will prove to be problematic.

Review: Solving Path Dependent Payoff Problems

Before diving into simulation, let's have a quick look at the process we generally follow in order to solve a path dependent pricing problem:

- We almost always start with market data for European options. Market data is usually unavailable for the path dependent payoff we are working with.
- We then specify a model and European options pricing technique and run a calibration algorithm in order to find the optimal parameters for the given market data.
  - To do this, we would choose a technique that handles European options robustly and is highly efficient (e.g. FFT).
  - Generally speaking, the computational cost of simulation and PDE techniques prevents us from using them in a calibration routine.

- We then take the model parameters from our calibration and apply these parameters to our path dependent pricing problem.
  - Once we know the model parameters, we need to choose a technique for our solving the path dependent pricing problem, which is generally either simulation or solving a PDE numerically.

Review: American Digital Options

- Finally, let's consider an American digital option. This option can be modeled using either a PDE approach or a simulation approach.
- In the next few lectures we will walk you through how to solve this problem via simulation for an arbitrary SDE.
- Before we do that, we are going review how to model this using the PDE approach.
- But prior to that, let's see if there are any ways to simplify the problem that avoid using these complicated methods...

Review: American Digital Options

- An American Digital option is an option that pays $1 if an asset price touches a certain barrier level *at any point* in its path.
- This is similar to a European Digital option which pays $1 if an asset price touches a barrier at expiry.
- In FX, these are referred to as **one-touch** and **no-touch** options.
  - There are also **double no-touch** options which pay if the exchange rate never leaves a given range.
  - European touch options only observe the barrier at expiry.
  - American touch options observe the barrier at every tick.

Review: American Digital Options

- Let's consider a simple up-and-in one-touch option with barrier K.
- For simplicity let's assume that the payout is at maturity regardless of when the barrier is triggered. This will be an easy assumption to relax as we move forward.
- The price of this option can be computed via the following expectation:

what if asset follow geo BM? what's max?

$$D_T(K) = \tilde{\mathbb{E}}\left[e^{-\int_0^T r_u\,du} 1_{\{M_T > K\}}\right] \quad (6)$$

reflecton princple

where M_T is the maximum value S takes over the interval $[0, T]$

- The European equivalent would be:

$$D_T(K) = \tilde{\mathbb{E}}\left[e^{-\int_0^T r_u\,du} 1_{\{S_T > K\}}\right] \quad (7)$$

Review: Relationship between American & European Digital Options

- Clearly the American versions of these options are path dependent and quadrature & FFT methods will not help us.
- However, if we know the distribution of the maximum (or minimum) of the asset value, then the problem reduces to a 1D integral which we can compute analyitcally or via quadrature.
- We can also use the **reflection principle** in order to give us a rough estimate for the price of the American digital option using a European Digital Option. (Why?)

Review: Relationship between American & European Digital Options

- It is clear that the American Digital pays out on all paths that the European Digital pays out by construction, as the expiry time is part of the monitoring period of the American.
- But let's consider a path where the American Digital is triggered prior to expiry. What is the probability that the European will also trigger?
  - The reflection principle tells us that for each realization that finishes above the barrier there is an inverted path with equal probability that finishes below the barrier.
  - As a result, the probability that the European will trigger is $\approx 50\%$. twice expensive.

Review: Relationship between American & European Digital Options

- So, if the conditional distribution is symmetric at each point where the American Digital is triggered, then the European should be half the cost of the American.
- What determines the accuracy of this rule of thumb?
  - Drift?
  - Vol?
  - Skew?

Review: Relationship between American & European Digital Options

- While the reflection principle technique is fairly elegant and an objectively interesting trick, it only works under certain assumptions. (Which ones?)
- So in reality using the reflection principle approach gives us little more than a sanity check for our code.
- Alas, it was worth a try, but we have no choice but to use simulation or PDE's...

Review: American Digital Options Pricing via PDE Methods

- In the last few lectures, we learned how to solve these types of problems by solving PDE's numerically.
- Before moving on to simulation, let's use this as an example and see how it fits into our PDE technique.
- What do we need to do to solve this problem numerically via PDE?
  - SDE/PDE
  - Grid in Time & Asset Price
  - Boundary Conditions
  - Discretization Scheme

Review: American Digital Options Pricing via PDE Methods

- Let's assume that we are working with the Black-Scholes PDE:

$$\frac{\partial D}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 D}{\partial s^2} + rs\frac{\partial D}{\partial s} - rD = 0 \tag{8}$$

- Recall that the PDE does not depend on the payoff. The only thing that changes with the payoff are the boundary conditions.
- Also note that while we have been working exclusively with the Black-Scholes PDE, we can write the PDE for an arbitrary diffusion process.
  - If we use a stochastic volatility model, for example Heston, this will introduce another dimension into our PDE.
  - If we use a jump process, this will introduce an integral into our PDE, making it a PIDE. PIDE's are beyond the scope of this course, but are covered very well in Hirsa Chapter 5.

Review: American Digital Options Pricing via PDE Methods

- We also need to define a mesh in time & asset price.
  - Let's break up our grid into N time intervals and M asset price intervals. This leads to:
    $h_t = \frac{T}{N}$
  - 0 is a natural lower bound for most assets. We need to define some $S_{\max}$ as the upper bound. That gives us:
    $h_s = \frac{S_{\max}}{M}$
- We will define $D(S_i, t_j)$ as the price of the American digital at the i^{th} price on our grid and the j^{th} time on our grid.

$$
\begin{aligned}
i &= \{0, \ldots, M\} \\
j &= \{0, \ldots, N\}
\end{aligned}
$$

Review: American Digital Options Pricing via PDE Methods

- We also need boundary conditions.
  - The "bottom" boundary condition is: $D(0, t_j) = 0, \quad \forall j$.
  - The "top" boundary condition is: $D(S_M, t_j) = 1, \quad \forall j$.
  - The terminal value conditions (at maturity) are:
    $D(S_i, t_N) = 1_{\{S_i > K\}}, \quad \forall i$

Review: American Digital Options Pricing via PDE Methods

- Lastly, we need to define our approximations for $\frac{\partial D}{\partial t}$, $\frac{\partial^2 D}{\partial s^2}$ and $\frac{\partial D}{\partial s}$ respectively. That is, we need to choose a discretization method.
- We want to use a first order central difference in order to approximate $\frac{\partial D}{\partial s}$. (Why?) Accuracy (Taylors, terms cancel)
- We will use a second order central difference in order to approximate $\frac{\partial^2 D}{\partial s^2}$.
- Remember that if use a backward difference for $\frac{\partial D}{\partial t}$ then we are using the explicit discretization scheme, and if we use a forward difference then we are using the implicit discretization scheme.
  - In this example let's use the explicit scheme.

Review: American Digital Options Pricing via PDE Methods

- Using these **finite difference approximations** in (8), we have:

$$
\begin{aligned}
&\frac{D(S_i,t_j)-D(S_i,t_{j-1})}{h_t} \\
+\quad &\frac{1}{2}\sigma^2 S_i^2 \frac{D(s_{i+1},t_j)-2D(s_i,t_j)+D(s_{i-1},t_j)}{h_s^2} \\
+\quad &rs_i\left(\frac{D(S_{i+1},t_j)-D(S_{i-1},t_j)}{2h_s}\right)-rD(S_i,t_j)=0 \quad (9)
\end{aligned}
$$

- The only unknown in the equation is $D(S_i,t_{j-1})$. Notice that this term only occurs once and is a function of the model parameters, r and σ as well as three values at the next time step: $D(S_{i-1},t_j)$, $D(S_i,t_j)$, $D(S_{i+1},t_j)$

- We can then solve the PDE by putting it in matrix form and iterating backward, starting at $j=N$ until we reach $j=0$.

Simulation: Overview

- Simulation is the process of using randomly generated numbers in order to approximate an expectation.
- In contrast to the previous methods that we have studied, which use deterministic estimators, simulation methods rely on random estimators.
- The idea behind simulation is that we **randomly sample** from some distribution **repeatedly** in order to get an estimate of some function of that distribution.
- There are some theoretical results that tells us that if this is done properly, the distribution of our sample will converge to the true distribution as we increase the number of draws.

Uses of Simulation

- Simulation is the most popular technique for dealing with problems with high dimensionality as it does not suffer from the **curse of dimensionality**.
- Where are simulation methods most commonly applied in finance?
  - Option Pricing: Calculating multi-dimensional path dependent option prices (& greeks)
  - Risk Mgmt: Computing risk metrics of complex portfolios
  - Portfolio Mgmt: Simulating returns as inputs to optimal portfolios (e.g. re-sampling)
  - Credit: Simulating default events of underlying entities.
  - Fixed Income: Modeling bond prices and interest rate curves.
  - Statistics: Generating synthetic data from an empirical distribution (e.g. Bootstrapping)

Option Pricing Example

- Let's start with the simplest possible example: a European Call Option.
- Clearly, simulation methods are unnecessary for solving these types of problems, but it will help give us some intuition around simulation before moving to more complex problems.
- Recall that the risk-neutral valuation formula for a European call is:

$$c_0 = \tilde{\mathbb{E}}\left[e^{-rT}\left(S_T - K\right)^+\right] \tag{10}$$

$$c_0 = e^{-\int_0^T r_u\,du} \int_{-\infty}^{+\infty} (S_T - K)^+ \,\phi(S_T)\, dS_T \tag{11}$$

- So far in this course, we have used quadrature and FFT methods to approximate these integrals.

Option Pricing via Exact Simulation: Overview

- Assuming $\phi(S_T)$ is known and we can generate samples from it, another approach to solving this problem is **exact simulation**.
- To see this, let's define Z_i as a set of i.i.d. random variables that are sampled from $\phi(S_T)$.
- The fact that they are i.i.d. means that each draw from $\phi(S_T)$ is *independent*.
- Additionally, this means that the draws are *identically distributed*, or come from the same distribution, namely $\phi(S_T)$.
- As we will see later when we discuss variance reduction techniques, we don't necesarily need the samples to be independent, but we clearly will always want them to be identically distributed.

Option Pricing via Exact Simulation: Central Limit Theorem & Law of Large Numbers

- The **law of large numbers** tells us that the average of the Z_i's converges in expectation (as $N \to \infty$) to the mean of $\phi(S_T)$, that is:

$$\bar{Z}_N = \frac{1}{N} \sum_{i=1}^{N} Z_i \to \tilde{\mathbb{E}}\left[S_T\right] = \mu \tag{12}$$

- The **central limit theorem** tells us not only that Z_N converges in expectation, but also how the standard deviation changes as we increase N.

  In particular, if $\phi(S_T)$ has mean μ and variance σ^2, the central limit theorem says:

Converges

$$\bar{Z}_N \to N\left(\mu, \frac{\sigma^2}{N}\right) \tag{13}$$

Option Pricing via Exact Simulation: Central limit theorem

- To see this, let's look at:

$$\epsilon_N := \frac{1}{N}\sum_{i=1}^{N} Z_i - \mu$$

- The law of large numbers tells us that $\mathbb{E}[\epsilon_N] = 0$ as $N \to \infty$.

- Let's have a look at $\hat{\sigma}_N$, the standard deviation of ϵ_N:

$$\hat{\sigma}_N = \sqrt{Var(\epsilon_N)} \tag{14}$$

$$= \sqrt{Var(\bar{Z}_N - \mu)} = \sqrt{Var(\bar{Z}_N)} \tag{15}$$

$$= \frac{1}{N}\sqrt{Var\left(\sum_{i=1}^{N} Z_i\right)} \tag{16}$$

Option Pricing via Exact Simulation: Central limit theorem

- Recall that we have assumed the Z_i's are independent. Therefore, by the properties of the variance we have:

$$
\begin{aligned}
\hat{\sigma}_N &= \frac{1}{N}\sqrt{\mathbb{E}\left[\sum_{i=1}^{N}(Z_i-\mu)^2\right]} && (17)\\
&= \frac{1}{N}\sqrt{\sum_{i=1}^{N}\mathbb{E}\left[(Z_i-\mu)^2\right]} && (18)\\
&= \frac{\sigma}{\sqrt{N}} && (19)
\end{aligned}
$$

- In the last step we used the fact that the Z_i's are identically distributed.

Option Pricing via Exact Simulation: Rate of Convergence

- In other words, as we increase N, the standard deviation decreases at a rate of $\frac{1}{\sqrt{N}}$.
- This is equivalent to saying that the rate of convergence is $\frac{1}{\sqrt{N}}$, or that it is $\mathcal{O}(N^{-\frac{1}{2}})$.
- Although this rate of convergence is slower than the other techniques we have considered so far, the rate of convergence does not depend on the dimensionality of the problem making it more efficient relative to other techniques as we add dimensions.

Option Pricing via Exact Simulation

- That means that if we use a large enough sample, we can use the average of the Z_i's in order to estimate our expectation.

- In particular, we can use the following **simulation approximation** for a European Call

$$\hat{c}_0 \approx \frac{1}{N} \sum_{i=1}^{N} e^{-rT} (Z_i - K)^+$$

  where Z_i is the value of the asset at expiry on the i^{th} simulation path and N is the total number of simulation paths.

- The simulation is said to be **exact** because we can sample from the distribution directly. In most simulation applications this will not be possible.

Option Pricing via Exact Simulation: Comments

- Note that we are using Monte Carlo to compute some expectation or integral.
  - The expectation may depend on the asset at maturity, or may depend on the entire path of an asset.
  - The expectation may be based on a single asset or many.
- In the 1D case, if we know $\phi(S_T)$, quadrature methods (with error $\mathcal{O}(N^{-2})$), are preferable to simulation, which has error $\mathcal{O}(N^{-\frac{1}{2}})$, but this will not be the case in many dimensions.
- This limits the practical applications of exact simulation, although there are a few:
  - Forward Starting Options
  - Bermudan Options

Option Pricing via Exact Simulation: Comments

- As we will see later, we can also simulate the SDE in cases where $\phi(S_T)$ is unknown. This type of simulation is far more common, and can also be extended to multiple dimensions.
- In the case of Black-Scholes, that would mean simulating:

$$dS_t = rS_t\,dt + \sigma S_t\,dW \tag{20}$$

What is required to compute a Monte Carlo simulation?

- Generate random numbers from an arbitrary distribution. This can be broken down into two steps:
  - Generate random numbers from the simplest possible distribution (uniform)
  - Transform these random numbers to an arbitrary distribution of our choice.
- Generate a path for the underlying asset using these random numbers.
  - In the case of the exact simulation we only need one random number per path as we simulate from the current time to option expiry. (Why can we do this?)
  - More commonly, we will not know the underlying PDF but only the SDE, and as a result will need many timesteps to form the

path of asset. In these cases, many random numbers will be required per path.

- Apply the payoff function in the expectation to each underlying path and obtain the payoff along each path.
- Calculate the expectation via Monte Carlo Integration by averaging the payoffs over the full set of draws.
- In the next few lectures, we will go through each of these steps in detail and show you how to build simulation algorithms.

Monte Carlo Simulation: Pseudocode

```
N = 10000
r = 0.025
T = 1.0
d = exp(-rT)

for j=1 to N
    U_j = GenerateRandomUniform()
    X_j = TransformRandomUniformToOurPdf(U_j)

    V_j = CalculatePayoff(T, X_j)
    V_j = d * V_j
end for loop

V = \frac{1}{N} \sum_{j=1}^{N} V_j
```

In the next few lectures, we will tell you how each of these functions is implemented.

Monte Carlo Simulation: Strengths & Weaknesses

- Stengths
  - May be the most intuitive and simplest method that we study.
  - Flexibility in handling different underlying models and payoff structures.
  - Ideal choice for working with multi-asset problems, as it does not suffer from the curse of dimensionality.
- Weaknesses
  - Rate of Convergence is $\mathcal{O}(N^{-\frac{1}{2}})$. This is notably slower than the other techniques that we have studied. As a result, **variance reduction techniques**, which we will touch on later in the course, are an important field of study.
  - Difficulty with Backward Induction & American Option Exercise Decisions.