FE610 Stochastic Calculus for Financial Engineers

Lecture 3. Calculaus in Deterministic and Stochastic Environments

Steve Yang

Stevens Institute of Technology

01/31/2012

Outline

The manner in which information flows in financial markets is more consistent with **stochastic calculus** than with **"standard calculus"**.

- For example, the relevant "time interval" may be different on different trading days.
- Numerical methods used in pricing securities are costly in terms of computer time. Hence, the pace of activity may make analysts choose coarser or finer time intervals depending on the level of volatility.

Some reasons behind developing a new calculus:

- A complicated random variable can have a very simple structure in continuous time, once the attention is focused on infinitesimal intervals.
- A "binomial" structure may be a good approximation to reality during an infinitesimal interval dt, but not necessarily in a large "discrete time" interval denoted by Δ.
- The main tool of stochastic calculus, Ito integral, may be more appropriate to use in financial markets than the Riemann integral used in standard calculus.

- **Functions:**

  Suppose A and B are two sets, and let f be a rule which associate to every element x of A, exactly one element y in B. Such a rule is called a *function* or a *mapping*. In mathematical analysis, functions are denoted by

$$f := A \rightarrow B \tag{1}$$

  or by

$$y = f(x), x \in A. \tag{2}$$

  If the set B is made of real numbers, then we say that f is a *real-valued function* and write

$$f := A \rightarrow R \tag{3}$$

  If A and B are themselves collections of functions, then f transforms a function into another, and is called an *operator*.

- **Random Functions:**

$$y = f(x), x \in A, \tag{4}$$

once the value of x is given, we get the element y. Often y is assumed to be a *real number*. Now consider the following significant alternation.

- There is a set W, where $\omega \in W$ denotes a *state of the world*. The function f depends on $x \in R$ and on $\omega \in W$:

$$f : R \times W \to R, \tag{5}$$

or

$$y = f(x, \omega), x \in R, \omega \in W, \tag{6}$$

where the notation $R \times W$ implies that one has to "plug in" to $f(\cdot)$ two variables, one from the set W, and the other from R.

The function $f(x, \omega)$ has the following property: Given a $\omega \in W$, the $f(\cdot, \omega)$ becomes a function of x only. Thus, for different values of $\omega \in W$ we get different functions of x. The $f(x, \omega)$ can be called *random function* or *stochastic process*.

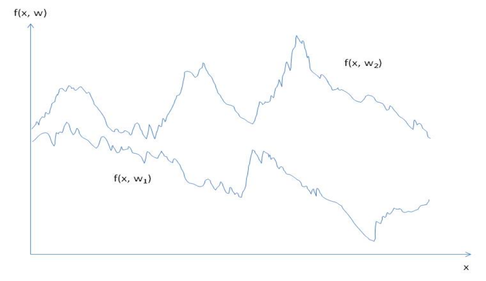

Figure : 1 - **Randomness of a stochastic process is in terms of the trajectory as a whole.**

- *The Exponential Function*

  The infinite sum

$$1 + 1 + \frac{1}{2!} + \frac{1}{3!} + ... + \frac{1}{n!} + ... \tag{7}$$

  converges to an irrational number between 2 and 3 as $n \to \infty$. This number is denoted by the letter e. The *exponential function* is obtained by raising e to a power of x:

$$y = e^x, x \in R. \tag{8}$$

  It has the following properties (discounting asset prices):

$$\frac{dy}{dx} = e^{f(x)} \frac{df(x)}{dx}. \tag{9}$$

$$e^x e^y = e^{x+y}. \tag{10}$$

  Finally, if x is a random variable, the $y = e^x$ will be random.

- *The Logarithmic Function*

  The logarithmic function is defined as the inverse of the exponential function. Practitioners often work with the logarithm of asset prices (**log return**). Given

  $$y = e^x, x \in R, \tag{11}$$

  the natural logarithm of y is given by

  $$ln(y) = x, y > 0. \tag{12}$$

- *Functions of Bounded Variation*

  Suppose a time interval is given by $[0, T]$. We *partition* this interval into n subintervals by selecting the $t_i, i = 1, ..., n$, as

  $$0 = t_0 \leq t_1 \leq t_2 \leq ... \leq t_n = T. \tag{13}$$

  The $[t_i - t_{i-1}$ represents the length of the ith subinterval.

- *Functions of Bounded Variation*

  Now consider a function of time $f(t)$, defined on the interval $[0, T]$:

$$f : [0, T] \to R. \tag{14}$$

$$\text{We form the sum} \tag{15}$$

$$\sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|. \tag{16}$$

Given that uncountably many partitions are possible, the sum assumes uncountably many values. If these sums are bounded from above the function $f(\cdot)$ is said to be of *bounded variation*. It implies functions are not excessively "irregular".

$$V_0 = max \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| < \infty \tag{17}$$

- *An Example*

  Consider function

$$f(t) = \begin{cases} t\sin(\frac{\pi}{t}) & \text{when } 0 < t \leq 1 \\ 0 & \text{when } t = 0 \end{cases} \tag{18}$$

It can be shown that $f(t)$ is not of bounded variation.

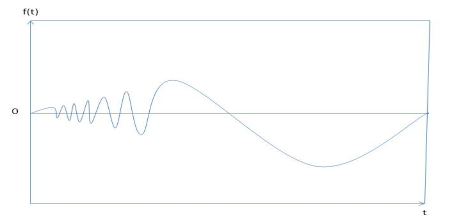

Figure : 2 - **Note that as $t \to 0$, f becomes excessively "irregular".**

Convergence and Limit

- **Convergence and Limit**

  Suppose we are given a *sequence*

$$x_0, x_1, x_2, \ldots, x_n, \ldots \tag{19}$$

  where x_n represents an object that changes as n is increased.

  In the case where x_n represents real numbers, we can state this more formally:

  DEFINITION: We say that a sequence of real numbers x_n converges to $x^* < \infty$ if for arbitrary $\epsilon > 0$, there exists a $N < \infty$ such that

$$|x_n - x^*| < \epsilon \text{ for all } n > N \tag{20}$$

  We call x^* the limit of x_n.

- **The Derivative**

  The notion of the derivative can be looked at in (at least) two different ways.

  - It is a way of defining rates of change of variables under consideration. In particular, if trajectories of asset prices are "too irregular", then their derivative may not exist.
  - The derivative is a way of calculating how one variable *responds* to a change in another variable. For example, given a change in the price of the underlying asset, we may want to know how the market value of an option written on it may move. These types of derivatives are usually taken using the *chain rule*.

  DEFINITION: Let $y = f(x)$ be a function of $x \in R$. Then the derivative of $f(x)$ with respect to x, if it exists, is formally denoted by the symbol f_x and is given by

$$f_x = lim_{\Delta \to 0} \frac{f(x+\Delta) - f(x)}{\Delta} \tag{21}$$

- **The Derivative (continued)**

  The variable x can represent any real-life phenomenon. Suppose it represents time. The Δ would correspond to a finite time interval. The $f(x)$ would be the value of y at time x, and the $f(x+\Delta)$ would represent the value of y at time $x+\Delta$.
- The numerator in (21) is the change in y during a time interval Δ. The ratio itself becomes the rate of change in y during the same interval.
- Why is a limit being taken in (21)? It is taken to make the ratio in (21) independent of the size of Δ, the time interval that passes.
- Making the ratio independent of the size of Δ, one pays a price. The derivative is defined for *infinitesimal* intervals. For larger intervals, the derivative becomes an *approximation* that deteriorates as Δ gets larger and larger.

- **The Derivative of Exponential Function**

  Consider the exponential function:

$$f(x) = Ae^{rx}, x \in R. \tag{22}$$

$$f_x = \frac{df(x)}{dx} = r[Ae^{rx}] = rf(x), \text{ or } \frac{f_x}{f(x)} = r. \tag{23}$$

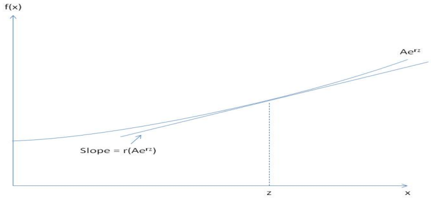

Figure : 3 - **The quantify f_x is the rate of change of $f(x)$ at point x.**

- **The Derivative as an Approximation**

  Let Δ be a finite interval. Then, using the definition in (21) and if Δ is "small", we can write approximately

$$f(x+\Delta) \cong f(x) + f_x \times \Delta. \tag{24}$$

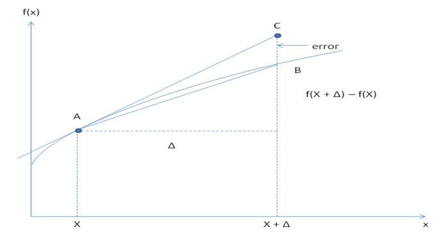

Figure : 4 - **The quality of approximation depends on the size of Δ and the shape of $f(\cdot)$.**

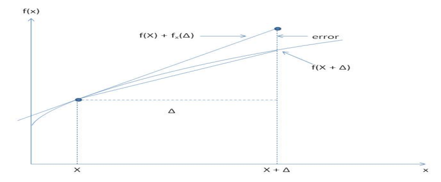

Figure : 5 – **Here, Δ is large. The approximation $f(x) + f_x\Delta$ is not near $f(x+\Delta)$.**

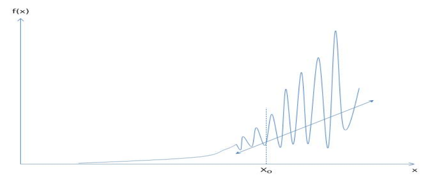

Figure : 6 – **When function $f(\cdot)$ is "irregular"/not smooth, the approximation is likely to fail.**

- *Example: High Variation*

  Consider the case that function $f(x)$ is continuous, but exhibits extreme variations even in small intervals Δ.

$$f(x+\Delta) \cong f(x) + f_x \times \Delta. \tag{25}$$

Here, not only is the prediction likely to fail, but even a satisfactory definition of f_x may not be obtained.

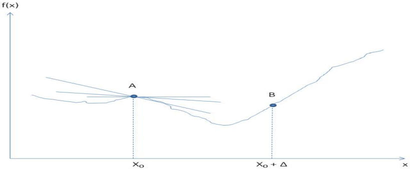

Figure : 6 - **When function $f(\cdot)$ is "irregular"/not smooth, the approximation is likely to fail.**

The Chain Rule

- The second use of the derivative is the chain rule. In the examples discussed earlier, $f(x)$ was a function of x, and x was assumed to represent time. The derivative was introduced as the response of a variable to a variation in time.

- In pricing derivative securities, we face a somewhat different problem. The price of a derivative asset, e.g., a call option, will depend on the price of the underlying asset, and price of the underlying asset depends on time.

- Hence, there is a chain effect. Time passes, new (small) events occur, the price of the underlying asset changes, and this affects the derivative asset's price. In standard calculus, the tool used to analyze these sorts of chain effects is known as the "chain rule".

- *The Chain Rule (continued)*

  Suppose in the example just given x was not itself the time, but a deterministic function of time, denoted by the symbol $t \geq 0$

  $$x_t = g(t). \tag{26}$$

  Then the function $f(\cdot)$ is called a *composite function*

  $$y_t = f(g(t)). \tag{27}$$

  DEFINITION: For f and g defined as above, we have

  $$\frac{dy}{dt} = \frac{df(g(t))}{dg(t)} \frac{dg(t)}{dt}. \tag{28}$$

  According to this, the chain rule is the product of two derivatives. 1). The derivative of $f(g(t))$ is taken with respect to $g(t)$. 2). The derivative of $g(t)$ is taken with respect to t.

- *The Integral*

  The integral is the mathematical tool used for calculating sums. In contrast to the Σ operator, which is used for sums of a countable number of objects, integrals denote sums of *uncountably infinite* objects.

  The general approach in defining integral is, in a sense, obvious. One would begin with an approximation involving a countable number of objects, and then take some limit and move into uncountable objects. Given that different types of limits, the integral can be defined in various ways.

- *The Riemann Integral*

  We are given a deterministic function $f(t)$ of time $t \in [0, T]$. Suppose we are interested in integrating this function over an interval $[0, T]$

$$\int_0^T f(s)ds, \tag{29}$$

- *The Integral (continued)*

  We partition the interval $[0, T]$ into n disjoint subintervals $t_0 = 0 < t_1 < ... < t_n = T$, then consider the approximation

$$\sum_{i=1}^{n} f(\frac{t_i + t_{i-1}}{2})(t_i - t_{i-1}), \tag{30}$$

DEFINITION: Given that

$$\max_i |t_i - t_{i-1}| \to 0, \tag{31}$$

the Riemann integral will be defined by the limit

$$\sum_{i=1}^{n} f(\frac{t_i + t_{i-1}}{2})(t_i - t_{i-1}) \to \int_0^T f(s)ds, \tag{32}$$

Where the limit is taken in a standard fashion.

- *The Integral (continued)*

  A better approximation can be achieved, if the base of the rectangles is small and the function $f(t)$ is smooth - that is, does not vary heavily in small intervals.

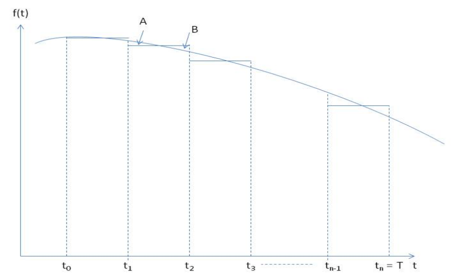

Figure : 8 - **When function $f(\cdot)$ is "irregular"/not smooth, the approximation is likely to fail.**

- *The Integral (continued)*

  In the standard calculus, using different heights for rectangles would not give different integral. But a similar conclusion cannot be reached in stochastic environments.

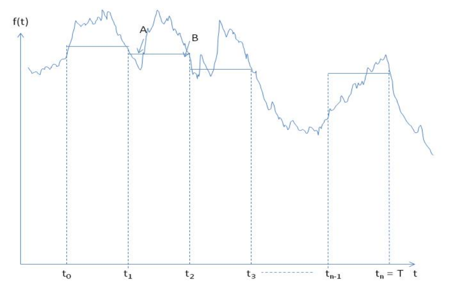

Figure : 9 – **Function $f(\cdot)$ shows step variations, and the approximation is likely to fail.**

- *The Integral (continued)*

  Suppose $f(W_t)$ is a function of a random variable W_t and that we are interested in calculating

$$\int_{t_0}^{T} f(W_s)dW_s, \tag{33}$$

  The choice of rectangles defined by (where W_t is a *martingale*)

$$f(W_{t_i})(W_{t_i} - W_{t_{i-1}}), \tag{34}$$

  will result in a different expression from the rectangles:

$$f(W_{t_{i-1}})(W_{t_i} - W_{t_{i-1}}), \tag{35}$$

  Then the expectation of the term in (36), conditional on information at time t_{i-1}, will vanish. This is because the future increments of a martingale will be unrelated to the current information set.

- *The Integral (continued)*

  Note that when $f(\cdot)$ depends on a random variable, the resulting integral itself will be random variable. In this sense, we will be dealing with *random integral*.

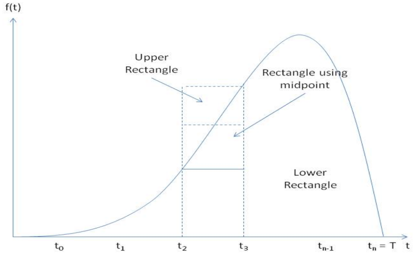

Figure : 9 - **In stochastic calculus, different definitions of approximating rectangles may lead to different results.**

- *The Stieltjes Integral*

  Define *differential df* as a small variation in the function $f(x)$ due to an infinitesimal variation in x

  $$df(x) = f(x + dx) - f(x). \tag{36}$$

  We have already discussed the equality

  $$df(x) = f_x(x)dx \tag{37}$$

  Now suppose we want to integrate a function $h(x)$ with respect to x

  $$\int_{x_0}^{x_n} h(x)dx \tag{38}$$

  $$\text{where the function } h(x) \text{ is given by} \tag{39}$$

  $$h(x) = g(x)f_x(x). \tag{40}$$

- *The Stieltjes Integral (continued)*

  The Stieltjes integral is defined as

$$\int_{x_0}^{x_n} h(x)dx, \tag{41}$$

$$\text{with } df(x) = f_x(x)dx \tag{42}$$

  This definition is not very different from that of the Riemann integral. If x represents time t, the Stieltjes integral over a partitioned interval $[0, T]$ is given by

$$\int_0^T g(s)df(s) \cong \sum_{i=1}^{n} g(\frac{t_i + t_{i-1}}{2})(f(t_i) - f(t_{i-1})). \tag{43}$$

  Because of these similarities, the limit as $max_i|t_i - t_{i-1}| \to 0$ of the right-hand side is known as the Riemann-Stieltjes integral.

The Riemann-Stieltjes Integral - Example

Example

We let

$$g(S_t) = aS_t, \tag{44}$$

where a is a constant. This makes $g(\cdot)$ a linear function of S_t. What is the value of the integral

$$\int_0^T aS_t dS(t) = 0.3, \tag{45}$$

If the Riemann-Stieltjes definition is used? Directly taking the integral gives

$$\int_0^T aS_t dS(t) = a\left[\frac{1}{2}S_t^2\right]_0^T \tag{46}$$

The Riemann-Stieltjes Integral - Example

Example

Due to the linearity of $g(\cdot)$, the area of the rectangle S_0ABS_T

$$a\left[\frac{S_T+S_0}{2}\right][S_T-S_0]=a\left[\frac{1}{2}S_T^2-\frac{1}{2}S_0^2\right]. \tag{47}$$

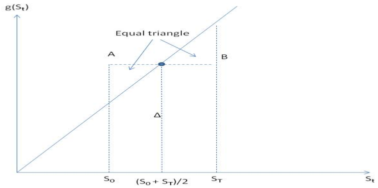

Figure : 9 – **Due to the linearity of $g(\cdot)$, a single rectangle is sufficent to replicate the area.**

- *Integration by Parts*

  Consider two differentiable functions $f(t)$ and $h(t)$, where $t \in [0, T]$ represents time. Then it can be shown that

$$\int_0^T f_t(t)h(t)dt \tag{48}$$

$$= [f(T)h(T) - f(0)h(0)] - \int_0^T h_t(t)f(t)dt, \tag{49}$$

  where $h_t(t)$ and $f_t(t)$ are the derivatives of the corresponding functions with respect to time. They are themselves functions of time t.

  The stochastic version of this transformation is very useful in evaluating Ito integrals. In fact, imagine that $f(\cdot)$ is random while $h(\cdot)$ is (conditionally) a deterministic function of time. Then, we can express *stochastic integrals* as a function of integrals with respect to a deterministic variable.

- *Partial Derivatives*

  Consider a call option, time to expiration affects the price (premium) of the call in two different ways. First, as time passes, the expiration date will approach, and the remaining life of the option gets shorter. This lowers the premium. But at the same time, as time passes, the price of the underlying asset will change. This will also affect the premium. We write

  $$C_t = F(S_t, t) \tag{50}$$

  where C_t is the call premium, S_t is the price of the underlying asset, and t is time. Now suppose we "fix" the time variable t and differentiate $F(S_t, t)$ with respect to S_t. The resulting partial derivative,

  $$\frac{\partial F(S_t, t)}{\partial S_t} = F_S, \tag{51}$$

- *Partial Derivatives (continued)*

  This effect is an abstraction, because in practice one needs some time to pass before S_t can change.

  The partial derivative with respect to time variable can be defined similarly as

  $$\frac{\partial F(S_t, t)}{\partial t} = F_t, \tag{52}$$

  Again, this shows the abstract character of the partial derivative. As t changes, S_t will change as well. But in taking partial derivatives, we behave as if it is a constant.

  Because of this abstract nature of partial derivatives, this type of differentiation cannot be used directly in representing actual changes of asset price in financial markets.

  They are useful in taking a total change and then splitting it into components that come from different sources.

Partial Derivatives - Example

Example

Consider a function of two variables:

$$F(S_t, t) = 0.3S_t + t^2, \tag{53}$$

where S_t is the price of a financial asset and t is time.

Taking the partial with respect to S_t involves simply differentiating $F(\cdot)$ with respect to S_t:

$$\frac{\partial F(S_t, t)}{\partial S_t} = 0.3, \tag{54}$$

Taking the partial with respect to t:

$$\frac{\partial F(S_t, t)}{\partial t} = 2t. \tag{55}$$

- *Total Derivatives*

  Let this total change be denoted by the differential dC_t. How much of this variation is due to a change in the underlying asset's prices? How much of the variation is the result of the expiration date getting nearer as time passes? Total differentiation is used to answer such questions.

  Let $f(S_t, t)$ be a function of the two variables. The the total differential is defined as

$$df = \left[\frac{\partial F(S_t, t)}{\partial S_t}\right] dS_t + \left[\frac{\partial F(S_t, t)}{\partial t}\right] dt. \tag{56}$$

  As t changes, S_t will change as well. But in taking partial derivatives, we behave as if it is a constant.

  Because of this abstract nature of partial derivatives, this type of differentiation cannot be used directly in representing actual changes of asset price in financial markets.

- *Taylor Series Expansion*

  Let $f(x)$ be a infinitely differentiable function of $x \in R$. And pick an arbitrary value of x; call this x_0.

  DEFINITIONS: The Taylor series expansion of $f(x)$ around $x_0 \in R$ is defined as

$$f(x) = f(x_0) + f_x(x_0)(x - x_0) + \frac{1}{2} f_{xx}(x_0)(x - x_0)^2$$
$$+ \frac{1}{3!} f_{xxx}(x_0)(x - x_0)^3 + \ldots \; = \sum_{i=0}^{\infty} \frac{1}{1!} f^i(x_0)(x - x_0)^i, \quad (57)$$

  where $f^i(x_0)$ is the ith order derivative of $f(x)$ with respect to x evaluated at the point x_0.

  We are not going to elaborate on why the expansion is valid, if $f(x)$ is continuous and smooth enough. Taylor series expansion is taken for granted.

Taylor Series - Example

Example

Consider the exponential function where t denotes time, T is fixed, $r > 0$ and $t \in [0, T]$:

$$B_t = 100e^{-r(T-t)}, \tag{58}$$

This function begins at $t = 0$ with a value of $B_0 = 100e^{-rT}$. Then it increases at a constant percentage rate r. As $t \to T$, the value of B_t approaches 100.

A first-order Taylor series expansion around $t = t_0$ will be

$$B_t \cong 100e^{-r(T-t_0)} + (r)100e^{-r(T-t_0)}(t - t_0), t \in [0, T], \tag{59}$$

Taylor series expansion of B_t shows that, as interest rates increase (decreases), the value of the bound decreases (increases).

- *Ordinary Differential Equations*

  The third major notion from standard calculus that we would like to review is the concept of an ordinary differential equation (ODE).

  For example, consider the expression

$$dB_t = -r_t B_t dt \text{ with know } B_0, r_t > 0. \tag{60}$$

  This expression states that B_t is a quantity that varies with t - i.e., changes in B_t are a function of t and of B_t. The equation is called an *ordinary differential equation*. Here, the percentage variation in B_t is proportional to some factor r_t times dt:

$$\frac{dB_t}{B_t} = -r_t dt. \tag{61}$$

  Now, we say that the function B_t defined by

- *Ordinary Differential Equations (continued)*

$$B_t = e^{\int_0^t r_u du}, \tag{62}$$

solves the ODE in (52) in that plugging it into (54) satisfies the equality (52). Thus, an ordinary differential equation is first of all an equation where there exist one or more *unknowns* that need to be determined.

$$dB_t = -r_t B_t dt. \tag{63}$$

the solution, with the condition $B_T = 1$ was

$$B_t = e^{\int_0^t r_u du}, \tag{64}$$

This example shows that the pricing functions for fixed income securities can be characterized as solutions of some appropriate differential equations.

ODE - Example

Example

In a simple equation $3x + 1 = x$,

the unknown is x, a number to be determined. Here the solution is $x = -1/2$.

In a *matrix equation* $Ax - b = 0$,

the unknown element is a vector. Under appropriate conditions, the solution would be $x = A^{-1}b$.

In an ordinary differential equation,

$$\frac{dx_t}{dt} = -ax_t + b, \tag{65}$$

where the unknown is x_t, a function. More precisely, it is a function of $t : x_t = f(t)$.