FE610 Stochastic Calculus for Financial Engineers

Lecture 7. Differentiation in Stochastic Environments

Steve Yang

Stevens Institute of Technology

02/28/2013

Outline

A Framework for Differentiation

The concept of differentiation deals with incremental changes in infinitesimal intervals. In applications to financial markets, changes in asset prices over incremental time periods are of interest. The natural framework to utilize for discussing differentiation is the stochastic differential equation (SDE):

$$dS(t) = a(S(t), t))dt + b(S(t), t)dW_t. \tag{1}$$

We will start with discrete and come up a formulation in continuous time. Consider a time interval $t \in [0, T]$, and the x axis is partitioned into n intervals of equal length h. We have

$$0 = t_0 < t_1 < ... < t_k < ... < t_n = T. \tag{2}$$

A Framework for Differentiation (continued)

Here we assume $t_k - t_{k-1} = h$ or $t_k = kh$ for all k.
Thus, we have the relation: $n = \frac{T}{n}$.

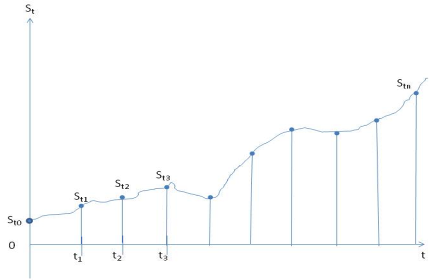

Figure : 2 - **Discrete Time Increments in Stochastic Environment**

A Framework for Differentiation (continued)

- We define the following quantities observed during these finite intervals:

$$S_k = S(kh) \text{ and} \tag{3}$$

$$\Delta S_k = S(kh) - S((k-1)h). \tag{4}$$

- Now pick a particular interval k. If the corresponding expectations exist, we define a random variable ΔW_k:

$$\Delta W_k = [S_k - S_{k-1}] - E_{k-1}[S_k - S_{k-1}]. \tag{5}$$

Here, $E_{k-1}[\cdot]$ represents the expectation conditional on information available at the end of interval $k-1$. The ΔW_k is the unpredictable part in $[S_k - S_{k-1}]$ given the information available at the end of the $(k-1)$th interval. The first term represents actual change in the asset price $S(t)$ during the kth interval.

A Framework for Differentiation (continued)

- We call unpredictable components of new information "innovations", and we note the following properties:

- ΔW_k is unknown at the end of the interval $(k-1)$. It is observed at the end of interval k. In measure theory, ΔW_k is said to be *measurable with respect to* I_k. That is, given the set I_k, one can tell the exact value of ΔW_k.
- Values of ΔW_k are unpredictable, given the information set of time $k-1$:

$$E_{k-1}[\Delta W_k] = 0, \text{ for all} k. \tag{6}$$

- ΔW_k represents changes in a martingale process and is called a *martingale difference*. The accumulated error process W_k will be given by

$$W_k = \Delta W_1 + ... + \Delta W_k = \sum_{i=1}^{k} \Delta W_i, \tag{7}$$

A Framework for Differentiation (continued)

- We can show that W_k is a martingale:

$$\begin{aligned} E_{k-1}[W_k] &= E_{k-1}[\Delta W_1 + ... + \Delta W_k] \\ &= [\Delta W_1 + ... + \Delta W_{k-1}] + E_{k-1}[\Delta W_k] \end{aligned} \tag{8}$$

The latter is true because $E_{k-1}[\Delta W_k]$ equals zeros and the $\Delta W_i, i = 1, ..., k-1$ are known given I_{k-1}.

$$W_k = \Delta W_1 + ... + \Delta W_k = \sum_{i=1}^{k} \Delta W_i, \tag{9}$$

In financial market, the important information contained in asset prices is indeed ΔW_k.

In particular, we want to show that under some reasonable assumptions, ΔW_k^2 and its infinitesimal equivalent dW_t^2 cannot be considered as "negligible" in Taylor-style approximations.

Assumptions and Proposition

- In this section, we deal with the innovation term ΔW_k and its square term (ΔW_k). We will use Merton's approach because it permits a better understanding of the economics behind the assumptions that will be made along the way.

  Let the (unconditional) variance of ΔW_k be denoted by V_k:

$$V_k = E_0[\Delta W_k^2]. \tag{10}$$

  The variance of cumulative errors is defined as:

$$V = E_0 \left[\sum_{k=1}^{n} \Delta W_k \right]^2 = \sum_{k=1}^{n} V_k, \tag{11}$$

  where the property that ΔW_k are uncorrelated across k is used and the expectations of cross product terms are set equal to zero.

Assumptions and Proposition (continued)

We now introduce some assumptions, following Merton (1990):

ASSUMPTION 1:

$$V > A_1 > 0, \tag{12}$$

where A_1 is independent of n. This assumption imposes a lower bound on the volatility of security prices.

* It says that when the period $[0, T]$ is divided into finer and finer subintervals, $n \to \infty$, and the variance of cumulative errors, V, will be positive.

** That is, more and more frequent observations of securities prices will not eliminate all the "risk". Clearly, most financial market participants will accept such an assumption. Uncertainty of asset prices never vanishes even when one observes the markets during finer and finer time intervals.

Assumptions and Proposition (continued)

ASSUMPTION 2:

$$V < A_2 < \infty, \tag{13}$$

where A_2 is independent of n. This assumption imposes an upper bound on the variance of cumulative errors and makes the volatility bounded from above.

* As the time axis is chopped into smaller and smaller intervals, more frequent trading is allowed. And such trading does not bring unbounded instability to the system.

** A large majority of market participants will agree with this assumption as well. After all, allowing for more frequent trading and having access to online screens does not lead to infinite volatility.

Assumptions and Proposition (continued)

We define

$$V_{\max} = \max_k [V_k, k = 1, ..., n]. \quad (14)$$

That is, $V_{\max}$ is the variance of the asset price during the most volatile subinterval.

ASSUMPTION 3:

$$\frac{K_k}{V_{\max}} > A_3, 0 < A_3 < 1, \quad (15)$$

with A_3 is independent of n.

According to this assumption, uncertainty of financial markets is not concentrated in some special periods. Whenever markets are open, there exists at least *some* volatility. This assumption rules out lotterylike uncertainty in financial market.

Assumptions and Proposition (continued)

PROPOSITION: Under assumption 1, 2, and 3, the variance of ΔW_k is proportional to h,

$$E[\Delta W_k^2] = \sigma_k^2 h. \tag{16}$$

where σ_k is a finite constant that does not depend on h. It may depend on the information at time $k-1$. According to this proposition, asset price become less volatile as h gets smaller.

PROOF: Use assumption 3:

$$V_k > A_3 V_{\text{max}}. \tag{17}$$

Sum both sides over all intervals:

$$\sum_{k=1}^{n} (V_k) > nA_3 V_{\text{max}}. \tag{18}$$

Assumptions and Proposition (continued)

Assumption 2 says that the left-hand side of this is bounded from above:

$$A_2 > \sum_{k=1}^{n} (V_k) > nA_3 V_{\max}. \tag{19}$$

Now divide both sides by nA_3:

$$\frac{1}{n}\frac{A_2}{A_3} > V_{\max}. \tag{20}$$

Note that $n = \frac{T}{h}$. Then,

$$\frac{1}{n}\frac{A_2}{A_3} > V_{\max} > V_k \text{ and } \frac{h}{T}\frac{A_2}{A_3} > V_k. \tag{21}$$

This gives an upper bound on V_k that depends only on h. We now obtain a lower bound that also depends only on h.

Assumptions and Proposition (continued)

We know that

$$\sum_{k=1}^{n} (V_k) > A_1 \tag{22}$$

is true. Then

$$nV_{\max} > \sum_{k=1}^{n} (V_k) > A_1. \tag{23}$$

Divide (23) by n. Use Assumption 3,

$$V_{\max} > \frac{A_1}{n} \text{ and } V_k > A_3 V_{\max} > \frac{A_1 A_3}{T} h. \tag{24}$$

Finally we have

$$\frac{h}{T}\frac{A_2}{A_3} > V_k > \frac{A_1 A_3}{T} h. \tag{25}$$

Assumptions and Proposition (continued)

Clearly the variance term V_k has upper and lower bounds that are proportional to h, regardless of what n is. This means that we should be able to find a constant σ_k depending only on k, such that V_k is proportional to h, and ignoring the (smaller) higher-order terms in h, write:

$$V_k = E[\Delta W_k]^2 = \sigma_k^2 h. \tag{26}$$

This proposition has several implications. An immediate one is the following. First, remember that if the corresponding expectation exist, one can always write

$$S_k - S_{k-1} = E_{k-1}[S_k - S_{k-1}] + \sigma_k \Delta W_k, \tag{27}$$

where ΔW_k now has variance h. After dividing both sides by h:

$$\frac{S_k - S_{k-1}}{h} = \frac{E_{k-1}[S_k - S_{k-1}]}{h} + \frac{\sigma_k \Delta W_k}{h}. \tag{28}$$

Assumptions and Proposition (continued)

But according to the proposition

$$E[\Delta W_k]^2 = h \tag{29}$$

Suppose we use this to justify the approximation:

$$\Delta W_k]^2 \cong h \tag{30}$$

- We take "limit" of the random variable:

$$\lim_{h \to 0} \frac{W_{(k-1)h+h} - W_{(k-1)h}}{h}. \tag{31}$$

Then, this could be interpreted as a time derivative of W_t. The approximation in (30) indicates that this derivative may not be well defined:

$$\lim_{h \to 0} \frac{W_{(k-1)h+h} - W_{(k-1)h}}{h} \to \infty. \tag{32}$$

Assumptions and Proposition (continued)

Figure 2 shows this graphically. We plot the function $f(h)$:

$$f(h) = \frac{h^{1/2}}{h}. \tag{33}$$

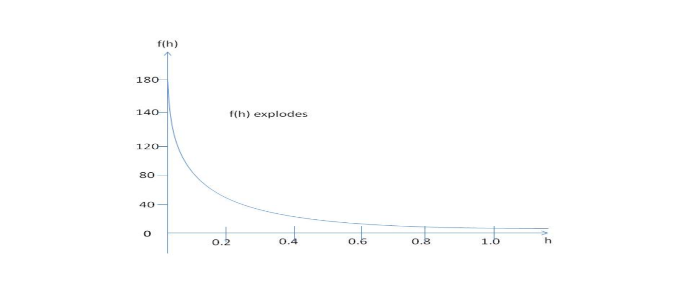

Figure : 2 – **As h gets smaller $f(h)$ goes to infinity.**

- First, we see that one can take any stochastic process S_t and write its variation during some finite interval h as

$$[S_k - S_{k-1}] = E_{k-1}[S_k - S_{k-1}] + \sigma_k \Delta W_k, \tag{34}$$

where the term ΔW_k is unpredictable given the information at the beginning of the time interval.

- Second, we showed that if h is "small", the unpredictable innovation term has a variance that is proportional to the length of the time interval, h:

$$Var(\Delta W_k) = h. \tag{35}$$

- Finally, we need to approximate the first term of (40),

$$E_{k-1}[S_k - S_{k-1}] = A(I_{k-1}, h). \tag{36}$$

The magnitude of this change depends on the latest information set and the length of the time interval.

Stochastic Differentials (continued)

- If $A(\cdot)$ is a smooth function of h, it will have a Taylor series expansion around $h = 0$,

$$A(I_{k-1}, h) = A(I_{k-1}, 0) + a(I_{k-1})h + R(I_{k-1}, h). \tag{37}$$

  Here, $a(I_{k-1})$ is the first derivative of $A(I_{k-1}, h)$ with respect to h evaluated at $h = 0$. The $R(I_{k-1}, h)$ is the remainder of the Taylor series expansion.
- Now, if $h = 0$, time will not pass and the predicted change in asset prices will be zero. In other words, $A(I_{k-1}, 0) = 0$.
- Thus, we can obtain the first-order Taylor approximation:

$$E_{k-1}[S_k - S_{k-1}] \cong a(I_{k-1}, kh)h. \tag{38}$$

  Hence, we can write a stochastic differential equation:

$$[S_{hk} - S_{(k-1)h}] \cong a(I_{k-1}, kh)h + \sigma_k[W_{kh} - W_{(k-1)h}]. \tag{39}$$

Stochastic Differentials (continued)

- Later, we let $h \to 0$ and obtain the infinitesimal version of this equation, which is the stochastic differential equation (SDE):

$$dS(t) = a(I_{k-1}, t)dt + \sigma_t dW(t). \tag{40}$$

This stochastic differential equation is said to have a drift $a(I_t, t)$ and a diffusion σ_t component.

Or in a different format as:

$$dS_t = a(S_t, t)dt + b(S_t, t)dW(t). \tag{41}$$

where dW_t is an *innovation* term representing unpredictable events that occur during the infinitesimal interval dt. The $a(S_t, t)$ and $b(S_t, t)$ are the drift and the diffusion coefficients, respectively. They are I_t-adapted.