# TEACHING NOTE 96-05:
# ITÔ'S LEMMA AND STOCHASTIC INTEGRATION

Version date: July 18, 2008 C:\Classes\Teaching Notes\TN96-05.doc

One of the most important mathematical results used in finance is Itô's Lemma. Though this result was found around 1950, it did not make its way into the financial models until 1973 when Black and Scholes discovered that it could be used to help find the price of an option.

Although there is a great deal of formal mathematic rigor to Itô's Lemma, the essential elements are quite simple. Let us begin, however, with a few basic statements from ordinary calculus. Recall that any differential in ordinary calculus is considered to have a limit of zero if raised to a power greater than 1.0. In other words, $dt^k \rightarrow 0$ if $k > 1$. Consider any well-behaved mathematical function like F(X,t).[1] Using a Taylor series expansion, the change in value of the function can be expressed as

$$dF = \frac{\partial F}{\partial X} dX + \frac{\partial F}{\partial t} dt + \frac{1}{2}\frac{\partial^2 F}{\partial X^2} dX^2 + \frac{1}{2}\frac{\partial^2 F}{\partial t^2} dt^2 + \frac{\partial^2 F}{\partial X \partial t} dXdt + \cdots$$

Because $dX^2 \rightarrow 0$, $dt^2 \rightarrow 0$ and $dXdt \rightarrow 0$, we write this as

$$dF = \frac{\partial F}{\partial X} dX + \frac{\partial F}{\partial t} dt,$$

It means that the change in F is a function of the changes in X and t. The change in X is multiplied by the partial derivative of F with respect to X and the change in t is multiplied by the partial derivative of F with respect to t. This is a formal way of stating that as X and t change, they induce a change in F. The changes in X and t are so small, however, that squared changes in X and t are zero in the limit and that their product is also zero in the limit.

In ordinary calculus, all variables are non-stochastic. This simply means that when we talk about a particular value of X, that value is known for certain. When X is stochastic, we leave the world of ordinary calculus and enter the world of stochastic calculus. There we cannot talk about a particular value of X; we must talk about a domain of possible values of X. We often summarize information like that in terms of expected values and variances. In the stochastic calculus, results are proven by demonstrating what happens when squared values of a

[1] By well-behaved, we mean that its first and second derivatives exist.

variable are multiplied by probabilities. A result is said to hold in "mean square limit."[2] A more formal statement of this concept is presented in later pages of this note.

Let us now propose that X is stochastic and follows an Itô process, such as $dW_t$ or a more generalized process such as

$$dX = \mu(X,t)dt + \sigma(X,t)dW_t.$$

Recall that $dW_t = \varepsilon_t\sqrt{dt}$. With the expressions $\mu(X,t)$ and $\sigma(X,t)$, we are allowing the expectation and variance of X to be functions of the level of X and time t.

Now suppose that we go back to the unspecified function F(X,t) and look at the Taylor series expansion when X is stochastic. While $dt^2$ is still zero because time is not stochastic, $dX^2$ is not zero because, as we noted in previous notes, $dW_t^2 = dt$. Note that dXdt is zero in the limit because $dXdt = (\mu(X,t)dt + \sigma(X,t)dW_t)dt = \mu(X,t)dt^2 + \sigma(X,t)dW_t dt = 0$, because $dW_t dt$ goes to zero in the limit.[3] This gives us

$$dF = \frac{\partial F}{\partial X}dX + \frac{\partial F}{\partial t}dt + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}dX^2.$$

Furthermore, we can state that $dX^2 = \mu(X,t)^2 dt^2 + \sigma(X,t)^2 dt + \mu(X,t)\sigma(X,t)dtdW_t$, which can be reduced to $dX^2 = \sigma(X,t)^2 dt$.[4] This gives us

$$dF = \frac{\partial F}{\partial X}dX + \frac{\partial F}{\partial t}dt + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}\sigma(X,t)^2 dt.$$

This is known as Itô's Lemma, being named for the Japanese mathematician who discovered it. It describes the stochastic process of a function F(X,t), which is driven by an Itô process X and time t. Note that by substituting for dX and $dX^2$ we could write the stochastic process for F as

$$dF = \left(\frac{\partial F}{\partial t} + \frac{\partial F}{\partial X}\mu(X,t) + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}\sigma(X,t)^2\right)dt + \frac{\partial F}{\partial X}\sigma(X,t)dW_t.$$

In this manner we see that the term in parentheses is the expected change in F and the variance is given as $(\partial F/\partial X)\sigma(X,t)$.[5] Note that the uncertainty in F comes from the uncertainty in $W_t$.

---

[2] The term "mean square limit" can be thought of somewhat like the concept of variance, which is the mean squared deviation around the expected value. It is approximately correct to say that a result in stochastic calculus holds when the variance converges to a finite value.

[3] From the definition of $dW_t$, the expression $dW_t dt$ will be $\varepsilon_t dt^{3/2}$. The 3/2 power on dt drives it to zero in the limit.

[4] See the previous footnote.

Because we often need to price derivative contracts, Itô's Lemma is widely used in finance. The price of a derivative is said to be "derived" from the price of the underlying asset and time. Thus, F(X,t) is a convenient specification of a derivative price. Suppose F is the price of an option or other derivative contract on an asset whose value is X. If we let that asset price evolve according to the Itô process, then we can be assured that the change in the option price is described by Itô's Lemma, as stated in the above equation.

In addition Itô's Lemma can be expressed in integral form. Let us restate the problem. We are given a random variable $X_t$, which follows the stochastic process:

$$dX_t = \mu(X(t), t)dt + \sigma(X(t), t)dW_t,$$

where we have subscripted t on X to reinforce that X takes on different values at different times t, depending on the evolution of $W_t$ through time. Applying Itô's Lemma to the function, $F(X_t,t)$, we obtain

$$dF(X_t, t) = \frac{\partial F(X_t, t)}{\partial X_t} dX_t + \frac{\partial F(X_t, t)}{\partial t} dt + \frac{1}{2} \frac{\partial^2 F(X_t, t)}{\partial X_t^2} dX_t^2.$$

Now suppose we integrate over the period time 0 to time t. This process is called *stochastic integration*, and is not defined the same way as standard or non-stochastic integration. The latter is done by dividing the area under a curve into rectangles. The area under the curve is approximately the sum of the areas of each of the rectangles. As the number of rectangles goes to infinity, the sum of the areas of the rectangles precisely approaches the area under the curve. Alternatively, standard integration can be viewed as adding up all of the infinitesimally small values dX/dt across values of t. Consequently, the derivative must be defined for all values of t. This process is called Riemann integration, or its variation Stieljes integration.

In a stochastic function, however, a derivative such as $dX_t/dt$ is not defined. To put it somewhat crudely, the zig-zaggedness of $X_t$ renders it impossible for the slope of a tangent line at any point to have a finite limit, which is a requirement for a derivative to exist. It is possible, however, to integrate a stochastic function by defining the process somewhat differently. Instead of being the limit of the area under each rectangle under the curve, a stochastic integral is defined as the *mean square limit*, meaning somewhat loosely that the integral is the expected value of the

---

[5]Remember that the expected value of $dW_t$ is zero; hence the expected value of dF comes from the first term on the right-hand side. The term in parentheses is the expected value; multiplying by dt scales it by the length of the time interval. The variance in the stochastic process comes from the second term on the right-hand side because the first term is the expected value, which is a constant. The variance is simply the square of whatever term is multiplied by $dW_t$ times the variance of $dW_t$, which is dt.

sum of the squared product of the volatility times the change in the stochastic variable. Such a limit will exist for the processes we typically encounter in finance. In this sense a stochastic integral is much more like a volatility measure. More formally the stochastic integral known as the Itô integral,

$$\int_0^t \sigma(X_\mu, \mu) dX_\mu ,$$

is

$$\lim_{n\to\infty} E\left(\sum_{k=1}^{n} \sigma(X_{k-1}, k)[X_k - X_{k-1}]\right),$$

where

$$\lim_{n\to\infty} E\left(\sum_{k=1}^{n} \sigma(X_{k-1}, k)[X_k - X_{k-1}] - \int_0^t \sigma(X_\mu, \mu) dX_\mu\right)^2 = 0 .$$

Conveniently, some of the properties of ordinary integration hold in stochastic integration. For example, by definition

$$\int_0^t dX_\mu = X_t - X_0 .$$

In other words, whether X is stochastic or not, the sum of the changes in X from $X_0$ to $X_t$ is, by definition, $X_t$ - $X_0$. In the special case where the volatility is constant, i.e., $\sigma(X_t,t) = \sigma$ for all t, we can pull the constant out and obtain

$$\int_0^t \sigma dX_\mu = \sigma[X_t - X_0].$$

Now let us write Itô's Lemma in its integral form. Using stochastic integration, we have

$$\int_0^t dF(X_\mu, \mu) = \int_0^t \frac{\partial F(X_\mu, \mu)}{\partial X_\mu} dX_\mu + \int_0^t \frac{\partial F(X_\mu, \mu)}{\partial \mu} d\mu + \int_0^t \frac{1}{2} \frac{\partial^2 F(X_\mu, \mu)}{\partial X_\mu^2} dX_\mu^2 .$$

This looks like it is going to be a problem because of the term $dX_u^2$, but we have previously noted that $dX_t^2 = \sigma(X_t,t)^2 dt$. Using that result and combining terms gives

$$F(X_t, t) - F(X_0, 0) = \int_0^t \left[\frac{\partial F(X_\mu, \mu)}{\partial \mu} + \frac{1}{2} \frac{\partial^2 F(X_\mu, \mu)}{\partial X_\mu^2} \sigma(X_\mu, \mu)^2\right] d\mu + \int_0^t \frac{\partial F(X_\mu, \mu)}{\partial X_\mu} dX_\mu ,$$

which is Itô's Lemma in integral formula, or sometimes called Itô's stochastic integral. Remember that either the differential or integral version of Itô's Lemma automatically implies that the other exists, so either can be used, and in some cases, one is preferred over the other.

Let us look at two more general cases of Itô's Lemma. First we look at the case where our random process is a function of two random processes. In other words let us say that X and Y follow stochastic differential equations driven by an Itô process. For right now, we do not have to specify those processes precisely. Consider a function F driven by X, Y and t. Then applying Itô's Lemma to F(X,Y,t), we obtain

$$dF = \frac{\partial F}{\partial X}dX + \frac{\partial F}{\partial Y}dY + \frac{\partial F}{\partial t}dt + \frac{1}{2}\frac{\partial^2 F}{\partial X^2}dX^2 + \frac{1}{2}\frac{\partial^2 F}{\partial Y^2}dY^2 + \frac{\partial^2 F}{\partial X \partial Y}dXdY .$$

Depending on the specifications of dX and dY, we can usually proceed to simplify this further. Note that there is no interaction term with dX and dt or with dY and dt as these will go to zero by the product of dX and dW or dY and dW.

Now consider two specific processes, X and Y, both with different but constant parameters.

$$dX = \mu_x dt + \sigma_x dW,$$
$$dY = \mu_Y dt + \sigma_Y dW.$$

Here both X and Y are driven by the same Brownian motion W. Now consider another process, Z, defined as

$$Z = XY.$$

We want to identify the stochastic process for Z. Applying Itô's Lemma to Z we obtain[6]:

$$dZ = \frac{\partial Z}{\partial X}dX + \frac{\partial Z}{\partial Y}dY + \frac{1}{2}\frac{\partial^2 Z}{\partial X^2}dX^2 + \frac{1}{2}\frac{\partial^2 Z}{\partial Y^2}dY^2 + \frac{\partial^2 Z}{\partial X \partial Y}dXdY .$$

Since Z is a very simple function of X and Y, we obtain the partial derivatives easily:

$$\frac{\partial Z}{\partial X} = Y, \quad \frac{\partial Z}{\partial Y} = X, \quad \frac{\partial^2 Z}{\partial X^2} = 0,$$
$$\frac{\partial^2 Z}{\partial Y^2} = 0, \quad \frac{\partial^2 Z}{\partial X \partial Y} = \frac{\partial}{\partial X}\left(\frac{\partial Z}{\partial Y}\right) = \frac{\partial}{\partial Y}\left(\frac{\partial Z}{\partial X}\right) = 1.$$

Substituting these results into the above, we obtain:

$$dZ = YdX + XdY + \sigma_X \sigma_Y dt .$$

[6]There is no term related to dt since Z is not directly determined by t.

This result reflects the fact that $dXdY = \sigma_x\sigma_y dt$. Now let us assume the two processes are driven by different Brownian motions:

$$dX = \mu_x dt + \sigma_x dW_X,$$
$$dY = \mu_Y dt + \sigma_Y dW_Y.$$

Let $\rho_{xy}$ be the correlation between the $W_x$ and $W_y$. We have $dXdY = \sigma_x\sigma_y dW_x dW_y$. Let us examine the expression $dW_x dW_y$. We know that the two Weiner increments are more specifically defined as:

$$dW_X = \varepsilon_X \sqrt{dt}$$
$$dW_Y = \varepsilon_Y \sqrt{dt}.$$

Now let us take the variance of their product. This was result is presented in TN00-05. We have

$$\text{Var}(dW_x dW_y) = 0,$$

$$E(dW_x dW_y) = dW_x dW_y = \rho_{xy} dt.$$

Going back to what we originally wanted,

$$dXdY = \sigma_x\sigma_y\rho_{xy} dt.$$

This is obviously the covariance between X and Y times dt. So

$$dZ = YdX + XdY + \sigma_x\sigma_y\rho_{xy} dt.$$

Of course if these are independent Brownian motions, the covariance term disappears.

These results provide only the basic tools that we shall use for pricing derivatives. The trick is to remove the uncertainty by forming a riskless hedge consisting of units of the derivative and units of the asset. When that is accomplished, this leads to an ordinary, non-stochastic partial differential equation, for which a closed-form solution sometimes exists and otherwise, a numerical solution can be obtained.[7]

## References

The references for TN96-04, Modeling Asset Prices as Stochastic Processes I, all generally apply here as well.

Good mathematical treatments are found in many places, such as

[7] A numerical solution is a solution to a differential equation that is obtained by solving the differential equation over a range of possible values of the variables. Usually a differential equation is restricted by certain conditions that define its value at certain combinations of the variables. These are called boundary conditions.

Karatzas, I. and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, 2nd. ed. New York: Springer-Verlag (1991), Chs. 3, 5.

For a fairly readable look from the finance perspective, see

Aucamp, D. C. and W. L. Eckardt, Jr. "An Intuitive Look at Itô's Lemma: A Pedagogical Note." *Financial Review* 16 (Spring, 1981), 41-50.

Baxter, M. and A. Rennie. *Financial Calculus.* Cambridge: Cambridge University Press (1996), Ch. 3.

Briys, E., M. Bellalah, H. M. Mai, F. de Varenne. *Options, Futures and Exotic Derivatives*. Chichester, U.K.: John Wiley and Sons (1998), Chs. 2.

Hull, J. C. *Options, Futures and Other Derivatives*, 5th ed. Upper Saddle River, New Jersey: Prentice-Hall (2003), Ch. 11.

Malliaris, A. G. and W. A. Brock. *Stochastic Methods in Economics and Finance* Amsterdam: North-Holland (1982), Ch. 2.

Neftci, Salih. *An Introduction to the Mathematics of Financial Derivatives*, 2nd ed. San Diego: Academic Press (2000), Chs. 9, 10.

Nielsen, L. T. *Pricing and Hedging of Derivative Securities*. Oxford, U.K.: Oxford University Press (1999), Ch. 2.

Wilmott, P., S. Howison, and J. DeWynne. *The Mathematics of Financial Derivatives.* Cambridge: Cambridge University Press (1995), Ch. 2.

The first finance application of Itô's Lemma was the classic Black-Scholes model,

Black, F. and M. Scholes. "The Pricing of Options and Corporate Liabilities." *The Journal of Political Economy* 81 (May-June, 1973), 637-654.