Ito's Lemma

Masaru Okada

October 10, 2025

The price of a stock option is a function of the underlying asset's price and time. Generally, it can be said that all derivatives are a function of the underlying's price process and time.

To understand this, one needs a solid grasp of the properties of functions of stochastic processes. Here I will summarize Ito's Lemma, a foundational concept discovered by the mathematician Kiyoshi Ito in 1951.

Consider a function $G = G(x,t)$. The change in G, denoted by ΔG, can be expressed in terms of the changes in its variables, Δx and Δt, as follows. (This is simply a Taylor expansion for a two-variable function.)

$$\begin{aligned}
\Delta G \;&= \frac{\partial G}{\partial x}\Delta x \;+\; \frac{1}{2!}\frac{\partial^2 G}{\partial x^2}\Delta x^2 \;+\; \frac{1}{3!}\frac{\partial^3 G}{\partial x^3}\Delta x^3 \;+\; \cdots && (1)\\
&\quad + \frac{\partial G}{\partial t}\Delta t \;+\; \frac{1}{2!}\frac{\partial^2 G}{\partial t^2}\Delta t^2 \;+\; \frac{1}{3!}\frac{\partial^3 G}{\partial t^3}\Delta t^3 \;+\; \cdots && (2)\\
&\quad + \frac{{}_2\mathrm{C}_1}{2!}\frac{\partial^2 G}{\partial x \partial t}\Delta x \Delta t \;+\; \frac{{}_3\mathrm{C}_1}{3!}\frac{\partial^3 G}{\partial x^2 \partial t}\Delta x^2 \Delta t \;+\; \frac{{}_3\mathrm{C}_2}{3!}\frac{\partial^3 G}{\partial x \partial t^2}\Delta x \Delta t^2 \;+\; \cdots && (3)\\
&= \sum_{n=1}^{\infty}\left(\Delta x \frac{\partial}{\partial x} \;+\; \Delta t \frac{\partial}{\partial t}\right)^n G && (4)
\end{aligned}$$

This equation is a general one, holding true for changes of any magnitude, not just infinitesimal ones. Building on this, let's now consider the case where the variable x is a stochastic process and the changes in x and t are infinitesimal.

Let's assume the variable x follows an Ito process, satisfying the equation:

$$dx \;=\; a(x,t)dt \;+\; b(x,t)dz$$

Here, dz represents a Wiener process. The functions $a = a(x,t)$ and $b = b(x,t)$ are the drift rate and the square root of the variance (b is the volatility) of the variable x, respectively.

Discretizing this equation gives us:

$$\Delta x \;=\; a\Delta t \;+\; b\varepsilon\sqrt{\Delta t}$$

Here, ε is a random variable that follows a standard normal distribution. (A standard normal distribution has a mean of zero and a standard deviation of 1.0.)

Squaring both sides of the equation yields:

$$\Delta x^2 \;=\; b^2\varepsilon^2\Delta t \;+\; 2ab\varepsilon\Delta t^{\frac{3}{2}} \;+\; a^2\Delta t^2$$

In the limit as $\Delta t \to 0$, we have:

$$\begin{aligned}
dx^2 \;&= b^2\varepsilon^2 dt \;+\; 2ab\varepsilon dt^{\frac{3}{2}} \;+\; a^2 dt^2 && (5)\\
&= b^2\varepsilon^2 dt && (6)
\end{aligned}$$

Here, we've dropped the higher-order terms in dt, keeping only the lowest-order term. Even with this simplification, dx^2 remains of the first order (the lowest order) with respect to dt, so it's a non-negligible term. (If even the lowest-order term were ignored, the resulting equation would not be a stochastic differential equation but a trivial conclusion.)

Now, let's consider this result, $dx^2 = b^2\varepsilon^2 dt$. As mentioned before, ε is a random variable following a standard normal distribution. Since a standard normal distribution has a mean (expected value) of zero and a variance of 1, we can write $E(\varepsilon) = 0$. With the variance being 1, we have:

$$E(\varepsilon^2) \ - \ E(\varepsilon)^2 \ = \ 1$$

Since $E(\varepsilon) = 0$, it follows that $E(\varepsilon^2) = 1$. The expected value of $\varepsilon^2\Delta t$ is $E(\varepsilon^2\Delta t) = \Delta t$. The variance of $\varepsilon^2\Delta t$ is:

$$E(\varepsilon^4\Delta t^2) \ - \ E(\varepsilon^2\Delta t)^2$$

It can be shown that this is on the order of Δt^2. (For now, let's accept this fact without proof and move on. This is a topic for a future discussion.) As the order of dt^2 drops, the variance of $\varepsilon^2 dt$ also drops to zero in the limit as $\Delta t \to 0$. With a variance of zero, dx^2 is no longer a stochastic variable. Furthermore, since its expected value is $E(\varepsilon^2 dt) = dt$, the result is:

$$dx^2 = b^2 dt$$

This is a key insight.

Using the result $dx^2 = b^2 dt$, we can expand $G = G(x,t)$ to the lowest order of dx and dt:

$$\begin{aligned}
dG \ &= \frac{\partial G}{\partial x}dx \ + \ \frac{1}{2!}\frac{\partial^2 G}{\partial x^2}dx^2 \ + \ \frac{1}{3!}\frac{\partial^3 G}{\partial x^3}dx^3 \ + \ \cdots && (7)\\
&\quad + \frac{\partial G}{\partial t}dt \ + \ \frac{1}{2!}\frac{\partial^2 G}{\partial t^2}dt^2 \ + \ \frac{1}{3!}\frac{\partial^3 G}{\partial t^3}dt^3 \ + \ \cdots && (8)\\
&\quad + \frac{{}_2\mathrm{C}_1}{2!}\frac{\partial^2 G}{\partial x \partial t}dxdt \ + \ \frac{{}_3\mathrm{C}_1}{3!}\frac{\partial^3 G}{\partial x^2 \partial t}dx^2 dt \ + \ \frac{{}_3\mathrm{C}_2}{3!}\frac{\partial^3 G}{\partial x \partial t^2}dxdt^2 \ + \ \cdots && (9)\\
&\simeq \frac{\partial G}{\partial x}dx \ + \ \frac{\partial G}{\partial t}dt \ + \ \frac{1}{2!}\frac{\partial^2 G}{\partial x^2}dx^2 && (10)\\
&= \frac{\partial G}{\partial x}dx \ + \ \frac{\partial G}{\partial t}dt \ + \ \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2 dt && (11)\\
&= \frac{\partial G}{\partial x}dx \ + \ \left(\frac{\partial G}{\partial t} \ + \ \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right) dt && (12)
\end{aligned}$$

This equation is known as Ito's Lemma.

Since x is an Ito process

$$dx \ = \ adt \ + \ bdz$$

we can substitute this into the equation above to change the expression from terms of dx, dt to terms of dt, dz:

$$\begin{aligned}
dG \ &= \frac{\partial G}{\partial x}dx \ + \ \left(\frac{\partial G}{\partial t} \ + \ \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right) dt && (13)\\
&= \frac{\partial G}{\partial x}(adt + bdz) \ + \ \left(\frac{\partial G}{\partial t} \ + \ \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right) dt && (14)\\
&= \frac{\partial G}{\partial x}bdz \ + \ \left(\frac{\partial G}{\partial x}a \ + \ \frac{\partial G}{\partial t} \ + \ \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2\right) dt && (15)
\end{aligned}$$

Thus, the drift (the coefficient of dt) for a general derivative $G = G(x, t)$ is:

$$\frac{\partial G}{\partial x}a \;+\; \frac{\partial G}{\partial t} \;+\; \frac{1}{2}\frac{\partial^2 G}{\partial x^2}b^2$$

and its volatility (the coefficient of dz) is:

$$\frac{\partial G}{\partial x}b$$

Now, let's consider the case where the stochastic process x is the stock price S. So, the derivative $G = G(S, t)$ is a stock derivative with S as the underlying asset. Since the stock price follows $dS = \mu S dt + \sigma S dz$, by comparing the coefficients with $dx = adt + bdz$, we see that $a = \mu S$ and $b = \sigma S$. This leads to:

$$dG \;=\; \frac{\partial G}{\partial x}\sigma S dz \;+\; \left(\frac{\partial G}{\partial x}\mu S \;+\; \frac{\partial G}{\partial t} \;+\; \frac{1}{2}\frac{\partial^2 G}{\partial x^2}\sigma^2 S^2\right) dt$$

This is the initial expression used in the derivation of the Black-Scholes-Merton equation.