

JUSTIFICATION IN THE DERIVATION OF THE BLACK-SCHOLES MODEL

6th February 2026

When working with stochastic processes, an interesting result is that $(dX)^2$ can be replaced by dt . In Article No.1, I used this fact to show Itô's Lemma, but here I want to justify the result properly. The purpose of this article is simple: to show why the quadratic variation of a Brownian motion leads to dt . I would classify this as a foundational idea that underpins the framework of the Black-Scholes model and much of modern quantitative finance.

WHY $(dX)^2$ BECOMES dt : A JUSTIFICATION USING STOCHASTIC CALCULUS

In Article No.1, I quoted the fact that we can replace $(dX)^2$ by dt , and this holds true for all such equations whenever we come across them. I will give an informal demonstration of the derivation. Consider the typical time interval of 0 to t and break this interval into n equal time steps of dt , where $n \in \mathbb{N}$ (i.e. a natural number). Consider the time at each interval to be t_j where:

$$t_j = \frac{jt}{n}$$

Where $j \in \mathbb{N}$ varying from 1 to n . Choose $dt = \frac{t}{n}$, so that each time interval is of length dt as required.

Let $dX(t_j)$ correspond to the value of the random variable at time t_j and let it be a random variable selected from a normal distribution with mean 0 and variance equal to dt (i.e. standard deviation equal to \sqrt{dt}). Therefore, we define the expectation of the random function $g(dX)$ as:

$$\mathbb{E}[g(dX)] = \int_{-\infty}^{\infty} g(x)f(x)dx,$$

Where

$$f(x) = \frac{1}{\sqrt{2\pi dt}} \exp\left(-\frac{x^2}{2dt}\right)$$

Note that $f(x)$ is just the probability density function (pdf) of the normal distribution with mean 0 and variance equal to dt . The expectation is the mean of the resulting function, which we can use this definition to calculate:

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{j=1}^n (dX(t_j))^2 - t\right)^2\right] &= \mathbb{E}\left[\left(\sum_{j=1}^n (dX(t_j))^2\right)^2 - 2t \sum_{j=1}^n (dX(t_j))^2 + t^2\right] \\ &= n\mathbb{E}\left[(dX)^4\right] + n(n-1)\left(\mathbb{E}\left[(dX)^2\right]\right)^2 - 2tn\mathbb{E}\left[(dX)^2\right] + t^2\mathbb{E}[1] \end{aligned} \quad (1)$$

By initially multiplying out the brackets, then by using the Linearity of Expectation and the fact that each $dX(t_j)$ comes from the same normal distribution for every value of j , i.e. $\mathbb{E}\left[(dX(t_j))^2\right] = \mathbb{E}\left[(dX(t_i))^2\right] \forall i, j \in \mathbb{N}$; also note that this fact is also true to the power of four $\forall i, j \in \mathbb{N}$. The first two terms in equation (1) come from the term:

$$\left(\sum_{j=1}^n (dX(t_j))^2\right)^2$$

When squaring this and expanding out the brackets, we obtain n terms of the form $(dX(t_j))^4$, thus we obtain the $n\mathbb{E}\left[(dX)^4\right]$ term, and we obtain $n(n-1)$ terms of the form $(dX(t_j))^2(dX(t_i))^2$ where $i \neq j$, thus we obtain the $n(n-1)\left(\mathbb{E}\left[(dX)^2\right]\right)^2$ term. Thus, we have written that $\mathbb{E}\left[(dX)^2\right] = \mathbb{E}\left[(dX(t_j))^2\right]$ and $\mathbb{E}\left[(dX)^4\right] = \mathbb{E}\left[(dX(t_j))^4\right] \forall j$. We also have used the result:

$$\begin{aligned}\mathbb{E} \left[(dX(t_j))^2 (dX(t_i))^2 \right] &= \mathbb{E} \left[(dX(t_j))^2 \right] \mathbb{E} \left[(dX(t_i))^2 \right] \\ &= \left(\mathbb{E} \left[(dX)^2 \right] \right)^2\end{aligned}$$

Which can be proved using the fact that random numbers $dX(t_i)$ and (t_j) are selected at different times (i.e. at t_i and t_j) and hence, are independent events. Perhaps, to give an example, will help the intuition behind why this works, consider a case where $n = 3$. Thus:

$$\begin{aligned}& \mathbb{E} \left[\left(\sum_{j=1}^n (dX(t_j))^2 \right)^2 - 2t \sum_{j=1}^n (dX(t_j))^2 + t^2 \right] \\ &= \mathbb{E} \left[\left((dX(t_1))^2 + (dX(t_2))^2 + (dX(t_3))^2 \right)^2 - 2t \left((dX(t_1))^2 + (dX(t_2))^2 + (dX(t_3))^2 \right) + t^2 \right] \\ &= \mathbb{E} \left[(dX(t_1))^4 + (dX(t_2))^4 + (dX(t_3))^4 + 2(dX(t_1))^2 (dX(t_2))^2 + 2(dX(t_1))^2 (dX(t_3))^2 + 2(dX(t_2))^2 (dX(t_3))^2 \right. \\ &\quad \left. - 2t \left((dX(t_1))^2 + (dX(t_2))^2 + (dX(t_3))^2 \right) + t^2 \right] \\ &= 3\mathbb{E} \left[(dX)^4 \right] + 6 \left(\mathbb{E} \left[(dX)^2 \right] \right)^2 - 6t\mathbb{E} \left[(dX)^2 \right] + t^2\mathbb{E}[1]\end{aligned}$$

And so, we can see in the case $n = 3$, we have the correct form.

Let us now evaluate equation (1). First note that:

$$\begin{aligned}\mathbb{E}[1] &= \int_{-\infty}^{\infty} f(x)dx = 1 \\ \mathbb{E} \left[(dX)^2 \right] &= \int_{-\infty}^{\infty} x^2 f(x)dx = dt \\ \mathbb{E} \left[(dX)^4 \right] &= \int_{-\infty}^{\infty} x^4 f(x)dx = 3(dt)^2\end{aligned}$$

Using integration by parts. Substituting these results back into equation (1), we obtain:

$$\mathbb{E} \left[\left(\sum_{j=1}^n (dX(t_j))^2 - t \right)^2 \right] = n(3(dt)^2) + n(n-1)(dt)^2 - 2tn(dt) + t^2$$

But since $dt = \frac{t}{n}$, we obtain:

$$\begin{aligned}\mathbb{E} \left[\left(\sum_{j=1}^n (dX(t_j))^2 - t \right)^2 \right] &= n \left(3 \left(\frac{t}{n} \right)^2 \right) + n(n-1) \left(\frac{t}{n} \right)^2 - 2tn \left(\frac{t}{n} \right) + t^2 \\ &= \frac{2t^2}{n}\end{aligned}$$

Let $n \rightarrow \infty$, so:

$$\mathbb{E} \left[\left(\sum_{j=1}^n (dX(t_j))^2 - t \right)^2 \right] = \frac{2t^2}{n} \rightarrow 0$$

as $n \rightarrow \infty$. Since the left-hand side (LHS) corresponds to an integral over the square of a quantity, the unique way that this integral is 0, in the limit $n \rightarrow \infty$, is when:

$$\sum_{j=1}^n (dX(t_j))^2 = t$$

as $n \rightarrow \infty$. This is the mean square limit.

Define the stochastic integral as:

$$\int_0^t h(t) dX = \lim_{n \rightarrow \infty} \sum_{j=1}^n h(t_j) dX(t_j)$$

for a function $h(t)$. Intuitively speaking, we can think of the integral on the LHS as an infinite sum of infinitesimally small areas under the curve. Thus:

$$\lim_{n \rightarrow \infty} \sum_{j=1}^n (dX(t_j))^2 = \int_0^t (dX)^2$$

by choosing $h(t) = dX(t)$. Thus:

$$\int_0^t (dX)^2 = t$$

can be rewritten as:

$$\int_0^t (dX)^2 = \int_0^t dt$$

Therefore, we obtain:

$$\int_0^t ((dX)^2 - dt) = 0$$

And so, we can conclude that $(dX)^2 = dt$ when written under an integral.

CONCLUSION

The key idea is that Brownian motion builds up variance linearly with time, which means $(dX)^2$ behaves like dt when we look at the limit $dt \rightarrow 0$. With that in mind, the rule that $(dX)^2 = dt$ isn't a trick, it follows directly from the fact that Brownian increments are of order $(dt)^{\frac{1}{2}}$ (noting that $dX = \phi(dt)^{\frac{1}{2}}$). This result is what makes Itô's Lemma work and why models like the Black-Scholes model can be derived cleanly. With this result established, we can use the rule without deriving it each time.

ALEXANDER FLANN