Informal derivation of the Black-Scholes PDE

Derek Huang

January 16, 2019

1 Introduction

The famous Black-Scholes-Merton equation, first published in 1973, is perhaps the most famous options pricing model that exists today. Although key assumptions of the model have been empirically disproven as inaccurate representations of financial markets, from a student standpoint, the model has much learning value. It is an interesting application of no-arbitrage arguments, an exercise in the risk-neutral valuation method, and an example of how mathematical rigor and financial intuition can be combined.

In this brief article, the Black-Scholes partial differential equation will be derived as an educational example, starting from basic assumptions. Solving the PDE is more arduous, requiring a change of variables into dimensionless parameters, the conversion of the Black-Scholes PDE into a permutation of the heat equation, and some longer calculations. For the sake of brevity, these will be omitted.

2 The Lognormal Random Walk

One of the key assumptions followed by the Black-Scholes model is that the underlying asset S follows a **stochastic process** $\{S_t\}$. The underlying distribution of S is assumed to be **constant**, with the return dS_t/S_t, the percentage change in the asset price at time t, assumed to satisfy the following SDE:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t$$

The differential term dW_t is the change in a standard Brownian motion W_t, where initial value $W_0 = 0$ a.s., $W_t \sim \mathcal{N}(0,\ t)$, and with each increment dW_t independent from the last. The t subscript for S and the Wiener process emphasizes the fact that the subscripted parameters are time-dependent. Since the underlying distribution of S, and therefore dS/S is assumed constant, μ and σ, which are respectively the average rate of return of S and the standard deviation of returns on S, are both constant. We can rewrite the equation on the previous page in terms of a differential change in S at time t, dS_t, as follows:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{2.1}$$

If we omit the stochastic term $\sigma S_t dW_t$, we can see that

$$\frac{dS_t}{S_t} = \mu dt$$

By simply integrating and exponentiating both sides, we have

$$S_t = e^{\mu t + C}$$

Setting $S_0 = e^C$, we can write

$$S_t = S_0 e^{\mu t}$$

Our simple, non-stochastic model for S_t shows that the final solution for S_t, including the stochastic term, should have an exponential form. We know that from its definition, the Wiener process W_t is normally distributed, so intuitively we can guess that S_t follows the **lognormal distribution**. Later, we can explicitly show this by applying Ito's lemma to solve for S_t.

3 The Risk-Neutral Measure

The motivation for changing the original probability measure $\mathbb{P}$, which the Wiener process W_t is under, is because in real life, different assets have differing amounts of risk, and individual risk preferences vary. The riskier an asset, the more return investors demand, but each investor prices risk differently. It can therefore be said that investors demand different **risk premia**. With relation to the lognormal model for S previously discussed, another way of stating this is that each investor may have a different estimate of μ, the growth rate and average return of S. However, by pricing under an equivalent probability measure where μ is irrelevant in the dynamics of S, a single price for a derivative on S can be calculated, independent of investors' individual estimate of μ. A rigorous discussion of assumptions behind risk-neutral pricing can be had, starting from the Arrow-Debreu market model, but is omitted for the sake of simplicity.

Referring back to the Brownian motion W_t under probability measure $\mathbb{P}$ that drives S_t, we can define a new Brownian motion W_t^Q given by the equation

$$W_t^Q = W_t + \frac{\mu - r}{\sigma} t \tag{3.1}$$

By Girsanov's theorem, which will not be rigorously treated in this brief example, since W_t is a continuous-time **martingale**, there exists another probability measure Q, under which the new Brownian motion W_t^Q is also a martingale.

The ratio $(\mu - r)/\sigma$ is known as the **market price of risk**. The derivation of the ratio is outside the scope of this exercise, but the ratio can be interpreted intuitively as the ratio of the asset's excess return over its volatility. Therefore, one can interpret W_t^Q as a standard Brownian motion adjusted for the price of risk, under probability measure Q, our unique **risk-neutral measure** $\mathbb{Q}$ for S_t. The technicalities of changing measure will be omitted for brevity.

Using the rules of stochastic calculus, we can rewrite (3.1) as

$$dW_t^Q = dW_t + \frac{\mu - r}{\sigma} dt \tag{3.2}$$

We can now describe the dynamics of S_t under our new probability measure $\mathbb{Q}$. Substituting (3.2) back into equation (2.1), we can derive

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t \left(dW_t^Q - \frac{\mu - r}{\sigma} dt \right) \\ &= \mu S_t dt - (\mu - r) S_t dt + \sigma S_t dW_t^Q \\ &= r S_t dt + \sigma S_t dW_t^Q \end{aligned} \tag{3.3}$$

It becomes clear that when we change our probability measure from real-world $\mathbb{P}$ to risk-neutral measure $\mathbb{Q}$, that under $\mathbb{Q}$, S earns on average the rate r, the **risk-free rate**, which we assume is known and constant. Intuitively, this makes sense, as after adjusting the future probabilities of the asset's evolution to account for all investor risk premia, r represents a **riskless** average rate of return on S. Now that the dynamics of S_t have been adjusted to fall under the risk-neutral probability measure $\mathbb{Q}$, we can price a derivative on S irrespective of individual investors' risk preferences or estimates of μ. In the next section, we will apply Ito's lemma to find a stochastic differential equation for a derivative on S, and also show its application in finding an analytic solution for (2.1).

4 Ito's Lemma

From a mathematical standpoint, Ito's lemma is the stochastic calculus counterpart to the chain rule. It is a necessary tool to apply to stochastic differential equations due to the random Wiener process component they contain, as ordinary rules of calculus are not sufficient. We will omit rigorous discussion of how to prove the lemma and instead focus on defining its use and application.

Consider the following stochastic differential equation of the form

$$dX_t = A(t)dt + B(t)dW_t \tag{4.1}$$

Here $A(t)$ and $B(t)$ are known, time-dependent functions, and as usual, dW_t is the differential change in a standard Wiener process. This is a typical Ito **drift-diffusion** process. Supposing that we have a twice-differentiable function $f(X,\ t)$, we can expand f in a Taylor series as

$$df = \frac{\partial f}{\partial X} dX_t + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial X^2} dX_t^2 + \frac{1}{2} \frac{\partial^2 f}{\partial t^2} dt^2 + \dots \tag{4.2}$$

From equation (4.1), we can write dX_t^2 as

$$dX_t^2 = A(t)^2 dt^2 + 2A(t)B(t)dtdW_t + B(t)^2 (dW_t)^2$$

By Ito's lemma, whose proof is beyond the scope of this simple derivation, as $dt \to 0$, the terms dt^2, $dtdW_t$, and any higher-order terms tend to 0 much faster than dt or $(dW_t)^2$. Intuitively, for small time steps, it is clear that dt^2 tends to 0 much faster than dt. However, the behavior of $(dW_t)^2$ is not as obvious.

Recall that when we defined W_t, the Wiener process under probability measure $\mathbb{P}$, we stated that $W_t \sim \mathcal{N}(0,\ t)$. In other words, W_t follows the normal distribution, with mean 0 and variance t.

We can informally write the differential change in W_t at time t, dW_t, as

$$dW_t = z\sqrt{dt} \tag{4.3}$$

In the above usage, z is a normally distributed random variable with zero mean and unit variance. That is, $z \sim \mathcal{N}(0,\ 1)$, and by inspection, $dW_t \sim \mathcal{N}(0,\ dt)$. In terms of order notation, it thus becomes clear from (4.3) that dW_t^2 is also $O(dt)$, and as $dt \to 0$, tends to 0 similarly.

Thus, applying Ito's lemma to (4.2), we can shorten the expression to

$$df = \frac{\partial f}{\partial X} dX_t + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial X^2} dX_t^2 \tag{4.4}$$

Our expression for dX_t^2 can also be simplified to

$$dX_t^2 = B(t)^2 dt \tag{4.5}$$

Substituting (4.5) and (4.1) into (4.4), we have

$$\begin{aligned} df &= \frac{\partial f}{\partial X}(A(t)dt + B(t)dW_t) + \frac{\partial f}{\partial t} dt + \frac{1}{2} \frac{\partial^2 f}{\partial X^2} B(t)^2 dt \\ &= \left(\frac{\partial f}{\partial t} + A(t)\frac{\partial f}{\partial X} + \frac{1}{2} B(t)^2 \frac{\partial^2 f}{\partial X^2} \right) dt + B(t)\frac{\partial f}{\partial X} dW_t \end{aligned} \tag{4.6}$$

Equation (4.6) is the form we expect for df after the application of Ito's lemma. Having defined the dynamics of S_t under the risk-neutral probability measure $\mathbb{Q}$, we now turn towards applying to lemma to find an expression describing the dynamics of a derivative on S, in this case an **option**, denoting its value by V. To do so, we must make some assumptions regarding the form of V, the value of the option, so that we can apply Ito's lemma.

First, we will make V a function of time-dependent parameters S and t, which can state explicitly by writing $V(S,\ t)$. Note that r and σ are not included, as we have assumed them both to be constant in our model for S defined in (3.3). Second, we assume that V is a twice-differentiable function with regards to S and t. With these two assumptions, we can now write an expression for dV in the form specified by Ito's lemma, where $A(t) = rS_t$ and $B(t) = \sigma S_t$, as

$$dV = \left(\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2\frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S_t\frac{\partial V}{\partial S}dW_t^Q \tag{4.7}$$

Note that $A(t)$ is the drift coefficient for S under $\mathbb{Q}$, while $B(t)$ is its diffusion coefficient. By Ito's lemma, dS_t^2 has become

$$dS_t^2 = \sigma^2 S_t^2 dt \tag{4.8}$$

This form of dV is exactly what we expect after the application of the lemma. One can already see several of the terms present in the final form of the Black-Scholes PDE in the expression for dV we have just derived. However, there is still a dW_t^Q term present in the expression, which indicates that there is still randomness that flows into the system for V. However, after looking at the dynamics of option V with S in a **hedged portfolio** and using no-arbitrage arguments under $\mathbb{Q}$, we can derive a fully deterministic PDE for the price of an option V on an asset S. The resulting equation is the Black-Scholes PDE.

Before we proceed, we can also quickly show how to apply Ito's lemma to find an analytic solution for S. Recalling our definition of dS_t in (2.1), under $\mathbb{P}$, suppose we choose $f(S) = \ln S_t$. By Ito's lemma, the form will be

$$df = \left(A(t)\frac{\partial f}{\partial S} + \frac{1}{2}B(t)^2\frac{\partial^2 f}{\partial S^2}\right)dt + B(t)\frac{\partial f}{\partial S}dW_t$$

Note that we omit the t partial derivative, as f is only a function of S. Substituting $\ln S_t$ for f, and again setting $A(t) = \mu S_t$ and $B(t) = \sigma S_t$, we have

$$\begin{aligned} d\ln S_t &= \left(\mu S_t\frac{1}{S_t} + \frac{1}{2}\sigma^2 S_t^2(-1)\frac{1}{S_t^2}\right)dt + \sigma S_t\frac{1}{S_t}dW_t \\ &= \left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma dW_t \end{aligned}$$

Integrating both sides, we can easily write

$$\ln S_t = \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t + C$$

Exponentiating both sides, and setting $S_0 = e^C$, we have

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \tag{4.9}$$

Equation (4.9) is our analytic solution for S, achieved through the application of Ito's lemma. In the next section, we will return to our expression for dV, and show how to eliminate the stochastic term from the expression to obtain a deterministic PDE governing the value V of an option.

5 The Hedged Portfolio

One of the other assumptions present in the derivation of the Black-Scholes PDE is that instantaneous, cost-less, **continuous time trading** and hedging is possible, where one may trade fractional shares of arbitrary precision. In other words, for any timestep dt, considering a portfolio Π, we can minimize the differential change in the portfolio $d\Pi$ instantaneously through hedging activities without incurring transactions costs. Of course, in reality, trading, even high-frequency computer trading, is done discontinuously, with discrete amounts of securities, and involves transactions costs and the market bid-ask spread. However, the continuous-time assumption allows for an interesting result, which is that all randomness driven by dW_t^Q in our expressions for dV and $d\Pi$ can be hedged away, resulting in an equation of only deterministic terms.

Consider a hedged portfolio Π, which we will define as

$$\Pi = V - \Delta S_t \tag{5.1}$$

As with our derivation of dV, we implicitly assume that Π is twice-differentiable and dependent on time-dependent parameters S and t. The portfolio Π is long one option of value V, and short Δ shares of the asset S at time t. For conceptual purposes, it is easier to think of S_t as an underlying cash equity such as a stock. The values of fixed income securities are significantly influenced by levels of interest rates, which thus play a significant role in their valuation and of derivatives on these instruments and introduce additional risk.

Informally, it is intuitively clear that $d\Pi$ can be represented in terms of dV and dS_t, which we can make clear by writing that $d\Pi = dV + \Delta dS_t$.

More formally, in expanded Taylor series form, we can write $d\Pi$ as

$$\begin{aligned} d\Pi &= \frac{\partial V}{\partial S} dS_t + \frac{\partial V}{\partial t} dt - \Delta \frac{dS}{dt} dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2} dS^2 + \frac{1}{2}\frac{\partial^2 V}{\partial t^2} dt^2 + \dots \\ &= \left(\frac{\partial V}{\partial S} - \Delta\right) dS_t + \frac{\partial V}{\partial t} dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2} dS^2 + \frac{1}{2}\frac{\partial^2 V}{\partial t^2} dt^2 + \dots \end{aligned}$$

We invoke Ito's lemma again, zeroing dt^2 and $dtdW_t^Q$ terms, and replacing $(dW_t^Q)^2$ with dt. Simultaneously, we substitute in our descriptions of dS and dS^2 under risk-neutral measure $\mathbb{Q}$ in (3.3) and (4.8), and rearranging, we have

$$\begin{aligned} d\Pi &= \left(\frac{\partial V}{\partial S} - \Delta\right) dS_t + \frac{\partial V}{\partial t} dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2} dS^2 \\ &= \left(\frac{\partial V}{\partial S} - \Delta\right)(rS_t dt + \sigma S_t dW_t^Q) + \frac{\partial V}{\partial t} dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S_t^2 dt \\ &= \left(\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S} - rS_t\Delta + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S_t \left(\frac{\partial V}{\partial S} - \Delta\right) dW_t^Q \end{aligned} \tag{5.2}$$

This is the form we expect after the application of Ito's lemma, and it looks very similar to the expression that we had for dV. But here we make an interesting observation. Suppose we define Δ, the fraction of S we are short in Π, as

$$\Delta = \frac{\partial V}{\partial S} \tag{5.3}$$

Substituting this choice of Δ back into (5.2), we see that

$$d\Pi = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}\right) dt \tag{5.4}$$

Notice that the term dW_t^Q no longer appears in our expression for $d\Pi$. That is, with our specific choice of Δ it is possible to **hedge away all risk** at any time-step dt! In other words, we can eliminate the stochastic term from our expression for $d\Pi$, and make $d\Pi$ wholly a function of deterministic variables. Since $d\Pi$ is now deterministic at each dt, the relationship between S and its option V in Π implies that there is only **one price** for V.

Recalling our expressions for dV and dS_t under $\mathbb{Q}$, equations (4.7) and (3.3) respectively, note that if we directly substitute them into $d\Pi = dV - \Delta dS_t$ and keep our choice of Δ, we will achieve exactly the same result we had with the Taylor series expansion. The dW_t^Q terms will cancel, and we will have a fully deterministic expression describing $d\Pi$. However, we still need a way to represent $d\Pi$ in known terms, which we will cover in the next section using no-arbitrage arguments under our risk-neutral probability measure $\mathbb{Q}$.

6 Solving $d\Pi$ for dV Under $\mathbb{Q}$

Consider again equation (3.3), the dynamics of S under measure $\mathbb{Q}$, which we have informally treated so far. A minor result from (3.3) is that the growth rate μ for any asset S is r, the risk-free interest rate. Conceptually, if we have a set of tradable instruments Ω, under our risk-neutral probability measure $\mathbb{Q}$, then $\forall S \in \Omega$, $\mu(S) = r$. Since the portfolio Π is also being considered under our risk-neutral measure $\mathbb{Q}$, and is also in Ω, then $\mu(\Pi) = r$ as well.

Therefore, we can represent the dynamics for Π under $\mathbb{Q}$ as

$$d\Pi = r\Pi dt + \sigma_\Pi \Pi dW_t^Q$$

Recalling that (5.4), our expression for $d\Pi$, contains only deterministic terms, then it is **necessary and sufficient** that $\sigma_\Pi = 0$, thus eliminating the stochastic dW_t^Q term from our equation for $d\Pi$ above. Therefore, we can describe the dynamics of the hedged portfolio Π under risk-neutral measure $\mathbb{Q}$ as simply

$$d\Pi = r\Pi dt$$

Substituting this equality back into (5.4), we now have

$$r\Pi dt = \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} \right) dt$$

We now substitute in (5.1), our definition of the hedged portfolio Π, and (5.3), our definition of Δ. After dividing dt on both sides, we have

$$\begin{aligned} r(V - \Delta S_t) &= \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} \\ rV - r\frac{\partial V}{\partial S} S_t &= \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} \end{aligned}$$

Rearranging the terms, our result becomes

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2} + rS_t \frac{\partial V}{\partial S} - rV = 0 \tag{6.1}$$

This is the famous Black-Scholes PDE. It contains no stochastic terms because of what we have demonstrated with the hedged portfolio Π, composed of V and S. In our particular definition of Π, if we are short Δ of S for each V we are long, the $d\Pi$ for every dt can be written as a deterministic PDE, with no stochastic terms. Under our risk-neutral probability measure $\mathbb{Q}$, where $\forall S \in \Omega$, $\mu(S) = r$, and realizing that $\Pi \in \Omega$ and that $d\Pi = r\Pi dt$ since $d\Pi$ is deterministic, we can therefore derive (6.1). To solve for the closed-form solution to the PDE, we will need to introduce initial and boundary conditions, perform change of variables, and fit the PDE into the heat equation. However, the closed-form derivation is a different exercise for a different time.