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:

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

S_t can be explicitly solved for and written in integral form,

$$\ln S_t = \mu t + \int_0^t \sigma dW_s + C$$
$$S_t = e^C \exp\left(\mu t + \int_0^t \sigma dW_s\right)$$

Setting $S_0 = e^C$ to represent the price of the asset at $t = 0$,

$$S_t = S_0 \ \exp\left(\mu t + \int_0^t \sigma dW_s\right) \tag{2.2}$$

It is intuitively evident that the lower bound for S_0 and t is 0. Since it is clear that $\ln S_t \sim \mathcal{N}(\mu, \sigma^2)$, that the natural logarithm of S_t is normally distributed with mean μ and variance σ^2, S_t therefore follows the **lognormal distribution**.

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.

4 Ito's Lemma

Having defined the dynamics of S_t under the risk-neutral probability measure $\mathbb{Q}$, we now turn towards finding a way to describe the dynamics of a derivative, in this case an **option**, denoting its value by V. To do so, we must make some further assumptions, which follow naturally from the work done so far.

First, the value of the option V is 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.

Second, we assume that V is a twice-differentiable function with regards to S and t, as then we can perform a Taylor series expansion on V describing at any point $(S_i,\ t_i)$, where $i \in \{0\} \cup \mathbb{N}$, the sensitivities of V to S and t at time t_i.

Performing the Taylor series expansion, we can write dV as

$$dV = \frac{\partial V}{\partial S} dS_t + \frac{\partial V}{\partial t} dt + \frac{1}{2} \frac{\partial^2 V}{\partial S^2} dS_t^2 + \frac{1}{2} \frac{\partial^2 V}{\partial t^2} dt^2 + \dots \tag{4.1}$$

Substituting (3.3) into our results above, we can write the term dS_t^2 as

$$dS_t^2 = (rS_t dt + \sigma S_t dW_t^Q)^2 = r^2 S_t^2 dt^2 + 2r\sigma S_t^2 dt dW_t^Q + \sigma^2 S_t^2 (dW_t^Q)^2 \tag{4.2}$$

Here we invoke the powerful **Ito's lemma** to simplify our expression for dV, removing some less important terms, and making the expression for dV finite.

By Ito's lemma, whose proof is beyond the scope of this simple derivation, as $dt \to 0$, the terms dt^2 and $dt dW_t^Q$ tend to 0 much faster than dt or $(dW_t^Q)^2$. Although intuitively, for small time steps, it is clear that dt^2 tends to 0 much faster than dt, the behavior of $dt dW_t^Q$ and $(dW_t^Q)^2$ is not obvious upon inspection.

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 $O(dt)$. That is, dW_t^2 is of the same order as dt, and as $dt \to 0$, tends to 0 at a similar rate. Recalling equation (3.2), we can rewrite $(dW_t^Q)^2$ as

$$(dW_t^Q)^2 = \left(dW_t + \frac{\mu - r}{\sigma} dt \right)^2$$

Expanding,

$$
\begin{aligned}
(dW_t^Q)^2 &= \left(dW_t + \frac{\mu - r}{\sigma}dt\right)^2 \\
&= dW_t^2 + 2\left(\frac{\mu - r}{\sigma}\right)dtdW_t + \left(\frac{\mu - r}{\sigma}\right)^2 dt^2
\end{aligned}
\tag{4.4}
$$

Again, by Ito's lemma, as $dt \to 0$, the $dtdW_t$ and dt^2 terms in $(dW_t^Q)^2$ will tend to 0 much faster, while dW_t^2 will behave similarly to dt, being $O(dt)$. Since μ, r, and σ are constants, we can treat $(\mu - r)/\sigma$ as a negligible constant. Thus, since dW_t^2 is $O(dt)$, $(dW_t^Q)^2$ is also $O(dt)$. A very important result, the proof of which is beyond the scope of this informal derivation is that as $dt \to 0$, with probability 1, $dW_t^2 \to dt$. By extension, as $dt \to 0$, $(dW_t^Q)^2 \to dt$.

Going back to (4.1) and (4.2), by Ito's lemma we can thus zero any terms containing dt^2 or $dtdW_t^Q$ and replace all occurrences of $(dW_t^Q)^2$ with dt.

Therefore, as $dt \to 0$, dS_t^2 can be rewritten as

$$
dS_t^2 = \sigma^2 S_t^2 dt \tag{4.5}
$$

Our Taylor series expansion of dV can thus be shortened to

$$
dV = \frac{\partial V}{\partial S}dS_t + \frac{\partial V}{\partial t}dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}dS_t^2
$$

Substituting in dS_t and dS_t^2, we can rearrange the terms in dV such that

$$
\begin{aligned}
dV &= \frac{\partial V}{\partial S}(rS_t dt + \sigma S_t dW_t^Q) + \frac{\partial V}{\partial t}dt + \frac{1}{2}\frac{\partial^2 V}{\partial S^2}\sigma^2 S_t^2 dt \\
&= \left(\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S^2}\right)dt + \sigma S_t \frac{\partial V}{\partial S}dW_t^Q
\end{aligned}
\tag{4.6}
$$

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. This equation is the Black-Scholes PDE.

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 (4.5), 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 (4.6) 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.