Bluff Your Way Through Black-Scholes

Saurav Sen

December 2000

Contents

What is Black-Scholes?

Black-Scholes is a framework for pricing options (and corporate liabilities, according to the title of the 1973 paper by Fisher Black and Myron Scholes which started everything).

The Classical Black-Scholes Model

The so-called classical Black-Scholes model is a simplification of the general Black-Scholes framework. It is very well-known and widely used in practice, even though it has some obvious shortcomings. The main reason for this seems to be that the cost/benefit ratio of using a significantly more accurate version is too high for most everyday purposes. Specifically, the quantity that is of greatest interest is usually the volatility of stock returns, σ. Classical Black-Scholes assumes a constant σ, and empirical evidence suggests that the first generalisation which is significantly more accurate for out-of-sample data is a stochastic volatility model, where σ is determined as the solution to another stochastic differential equation (SDE). Additionally, classical Black-Scholes often provides a fairly good insight into option pricing anyway. For this and other reasons, we concentrate on classical Black-Scholes in the present course.

Some Useful Background Mathematics

1. **Brownian Motion:** Brownian motion can be defined in several equivalent ways. Here is one simple definition: The process $\{W_t : t \geq 0\}$ is called a Brownian motion with respect to the probability measure $\mathbb{P}$ if

   - $W_0 = 0$ and W_t has continuous sample paths.
   - $W_t \sim N^{\mathbb{P}}(0, t)$, i.e. W_t is normally distributed with mean 0 and variance t under the probability measure $\mathbb{P}$.
   - $W_t - W_s \sim N^{\mathbb{P}}(0, t-s)$ and is independent of the history of the process until time s. In words, Brownian motion has independent, Gaussian increments.

   The covariance function for a Brownian motion is calculated as follows. If $t > s$,

   $$\begin{aligned} \mathbb{E}[W_t W_s] &= \mathbb{E}\left[(W_t - W_s) W_s + W_s^2\right] \\ &= \mathbb{E}[W_t - W_s]\,\mathbb{E}[W_s] + \mathbb{E}\left[W_s^2\right] \\ &= s \end{aligned}$$

   The first term on the second line follows from independence of increments. Similarly, if $s > t$, then we can show that $\mathbb{E}[W_t W_s] = t$. Thus we have the result:

   $$\mathbb{E}[W_t W_s] = \min\{t, s\}$$

   Brownian motion is a martingale. A martingale is a constant expectation process. M_t is called a $\mathbb{P}-$ martingale if

   $$\mathbb{E}_t^{\mathbb{P}}[M_T] = M_t$$

   For a Brownian motion,

   $$\begin{aligned} \mathbb{E}_t^{\mathbb{P}}[W_T] &= \mathbb{E}_t^{\mathbb{P}}[W_T - W_t + W_t] \\ &= \mathbb{E}_t^{\mathbb{P}}[W_T - W_t] + W_t \\ &= W_t \end{aligned}$$

   The first term on the second line has expected value zero, and the second term is known at time t, hence we can write W_t instead of $\mathbb{E}_t^{\mathbb{P}}[W_t]$. For other properties of Brownian motion, see the references and exercise sheets.

2. **Ito Integrals:** Ito integrals are one way of assigning meaning to the symbol

   $$I = \int_0^T f(t, W_t)\, dW_t$$

They are defined in a manner similar to Riemann integrals in deterministic calculus, i.e. as the limit of an approximation by discrete sums.

$$I \approx \sum_{n=0}^{N-1} f\left(t_n, W_{t_n}\right)\left\{W_{t_{n+1}} - W_{t_n}\right\}$$

Note that the function is evaluated at the left end point t_n of each sub-interval. For deterministic integrals, the choice does not matter - we could choose any point in the sub-interval and the sum would converge to the same value (provided the function was integrable). However, for stochastic integrals they can give different answers. The Ito definition uses the left end-point. The Stratanovich definition uses the midpoint of each subinterval, $\left(t_n + t_{n+1}\right)/2$. The Stratanovich definition has the advantage that rules of classical Newtonian calculus translate directly to the stochastic setting. However, the Ito definition is more relevant in the context of finance, since it is non-anticipating (We know prices at the start of each time period - not halfway through into the future). The main thing to note about Ito integrals is that I is a random variable (unlike the deterministic case). Additionally, since I is essentially the limit of a sum of normal random variables, I is normally distributed too, and can be characterised by its mean and variance. The two main results used by us are:

$$\mathbb{E}\left[\int_0^T f\left(t, W_t\right) dW_t\right] = 0$$

$$\mathbb{E}\left[\left(\int_0^T f\left(t, W_t\right) dW_t\right)\left(\int_0^T g\left(t, W_t\right) dW_t\right)\right] = \int_0^T \mathbb{E}\left[f\left(t, W_t\right) g\left(t, W_t\right)\right] dt$$

A consequence of the second property is that

$$\mathbb{E}\left[\left(\int_0^T f\left(t, W_t\right) dW_t\right)^2\right] = \int_0^T \mathbb{E}\left[f\left(t, W_t\right)^2\right] dt$$

This is called the Ito isometry.

3. **Ito's Lemma:** Ito's lemma forms the basis for defining a set of calculus-like rules for stochastic processes. In a simple form, it says that if

$$dX_t = \mu_t dt + \sigma_t dW_t$$

where μ_t and σ_t are $\mathcal{F}_t-$adapted, i.e. known for sure at time t, and if $f\left(t, x\right)$ is continuously differentiable once in the first argument and twice in the second, then $f\left(t, X_t\right)$ is also an $\mathcal{F}_t-$ adapted process, and

$$df\left(t, X_t\right) = \left\{\frac{\partial f}{\partial t}\left(t, X_t\right) + \mu_t \frac{\partial f}{\partial x}\left(t, X_t\right) + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2}\left(t, X_t\right)\right\} dt + \sigma_t \frac{\partial f}{\partial x}\left(t, X_t\right) dW_t$$

One way of 'deriving' Ito's lemma is to Taylor-expand f, use the fact that $dW^2 \rightarrow dt$ as $dt \rightarrow 0$, and retain terms to order dt. See the exercise sheets for examples on how to apply Ito's lemma.

4. **Moment-Generating Functions (MGF's):** You can find information on MGF's in any textbook on basic probability. If X is a random variable, then its MGF is defined as

$$\begin{aligned} M_X(\theta) &:= \mathbb{E}\left[e^{\theta X}\right] \\ &= \mathbb{E}\left[1 + \theta X + \tfrac{1}{2!}\theta^2 X^2 + \tfrac{1}{3!}\theta^3 X^3 + ... + \tfrac{1}{n!}\theta^n X^n + ...\right] \\ &= 1 + \theta\mathbb{E}[X] + \tfrac{1}{2!}\theta^2\mathbb{E}\left[X^2\right] + \tfrac{1}{3!}\theta^3\mathbb{E}\left[X^3\right] + ... + \tfrac{1}{n!}\theta^n\mathbb{E}[X^n] + ... \end{aligned}$$

The reason for the name should now be obvious: The $k-$th moment of X is the $k-$th derivative of the MGF, evaluated at $\theta = 0$. One result that is particularly relevant in finance is the MGF of a normal distribution. If $X \sim N(\mu, \sigma^2)$, then

$$M_X(\theta) = e^{\theta\mu + \frac{1}{2}\theta^2\sigma^2}$$

Here is a proof:

$$X \sim N\left(\mu, \sigma^2\right) \Rightarrow X = \mu + \sigma\phi$$

where $\phi \sim N(0, 1)$. Hence

$$M_X(\theta) = \mathbb{E}\left[e^{\theta X}\right] = \mathbb{E}\left[e^{\theta\{\mu+\sigma\phi\}}\right] = e^{\theta\mu}\mathbb{E}\left[e^{\theta\sigma\phi}\right]$$

The MGF of X is therefore equal to the MGF of ϕ, with θ replaced by $\theta\sigma$. So we calculate the (easier) MGF of ϕ :

$$\begin{aligned} \mathbb{E}\left[e^{\theta\phi}\right] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\theta x}e^{-x^2/2}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\theta x - x^2/2}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(x^2 - 2\theta x + \theta^2 - \theta^2\right)}dx \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-\theta)^2 + \frac{1}{2}\theta^2}dx \\ &= e^{\frac{1}{2}\theta^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-u^2/2}du \text{ where } u = x - \theta \\ &= e^{\frac{1}{2}\theta^2} \end{aligned}$$

Thus

$$M_X(\theta) = \mathbb{E}\left[e^{\theta X}\right] = e^{\theta\mu}\mathbb{E}\left[e^{\theta\sigma\phi}\right] = e^{\theta\mu + \frac{1}{2}\theta^2\sigma^2}$$

5. **Kolmogorov Backward Equation:** Kolmogorov's backward equation relates the solution of the PDE

$$F_t + \frac{1}{2}\sigma(t, x)^2 F_{xx} + \mu(t, x) F_x + r(t) F = 0$$
$$F(T, x) = \psi(x)$$

with the solution of the SDE

$$dx = \mu(t,x)\,dt + \sigma(t,x)\,dW$$

Consider the function

$$H(t,x) = F(t,x)\exp\left\{\int_0^t r(s)\,ds\right\} = F(t,x)\,R(t) \in \mathcal{F}_t$$

Since R is a function of time alone, the usual product rule of differentiation can be applied

$$\begin{aligned} dH &= FdR + RdF \\ &= rFRdt + R\left\{\left(F_t + \frac{1}{2}\sigma^2 F_{xx} + \mu F_x\right)dt + \sigma F_x dW\right\} \\ &= R\left[F_t + \frac{1}{2}\sigma^2 F_{xx} + \mu F_x + rF\right] + \sigma R F_x dW \end{aligned}$$

The first term is zero because F satisfies the PDE. In integrated form

$$H(T,x(T)) - H(t,x(t)) = \int_t^T \sigma(s,x)\,R(s)\,F_x(s,x)\,dW_s$$

Taking expectations, we have, since $\mathbb{E}[dW] = 0$,

$$\begin{aligned} H(t,x(t)) &= \mathbb{E}_t\left[H(T,x(T))\right] \\ &\Rightarrow F(t,x)\exp\left\{\int_0^t r(s)\,ds\right\} = \mathbb{E}_t\left[F(T,x)\exp\left\{\int_0^T r(s)\,ds\right\}\right] \\ &\Rightarrow F(t,x) = \mathbb{E}_t\left[\psi(x)\exp\left\{\int_t^T r(s)\,ds\right\}\right] \end{aligned}$$

In other words, if x is a solution of the SDE

$$dx = \mu(t,x)\,dt + \sigma(t,x)\,dW$$

then the solution of the PDE

$$\begin{aligned} F_t + \frac{1}{2}\sigma(t,x)^2 F_{xx} + \mu(t,x)\,F_x + r(t)\,F = 0 \\ F(T,x) = \psi(x) \end{aligned}$$

is given by

$$F(t,x) = \mathbb{E}_t\left[\psi(x)\exp\left\{\int_t^T r(s)\,ds\right\}\right]$$

Note that this works even when r is stochastic! Thus we have a recipe for solving PDE's: form the appropriate SDE and then take expectations as above. For more on Kolmogorov's backward equation, see the model solutions.

Assumptions in the Classical Black-Scholes Model

- The market consists of a continuous-trading, perfectly elastic economy. Perfect elasticity means that the process driving the stock price is unaffected by swings in demand, or equivalently that all investors are price-takers. Assets are stock (risky), cash (riskless) and a derivative (defined in terms of the stock price at maturity).
- Efficient Market Assumption: All information about the past performance of the stock is contained in the stock price S, which is observable to all market participants. This means that the stock price is a Markov process.
- There are no transaction costs.
- Stock can be infinitely subdivided and short selling (selling something you do not own) is allowed, i.e. fractional and negative holdings of stock are permitted.
- The risk-free continuously compounded interest rate r is constant.
- Stock prices are assumed to evolve according to the SDE:

$$\frac{dS}{S} = \mu dt + \sigma dW$$

  where μ and σ are constants and W is a standard Brownian motion.
- The price of an option at time t, which we write as $V(t, S)$ is a deterministic function of t and S.

The Equation $dS/S = \mu dt + \sigma dW$

The equation

$$\frac{dS}{S} = \mu dt + \sigma dW$$

is often referred to as the geometric Brownian motion assumption in Black-Scholes. This is because the solution to the equation is

$$S_t = S_0 e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W}$$

which looks like geometric growth driven by a drifting Brownian motion. To see this, define $Y = f(t, S) = \log S$. Then using Ito's lemma, we have:

$$f_t = 0; f_S = 1/S; f_{SS} = -1/S^2$$

$$\begin{aligned} dY = df(t, S) &= \left(f_t + \mu S f_S + \tfrac{1}{2}\sigma^2 S^2 f_{SS}\right) dt + \sigma S f_S dW \\ &= \left(\mu - \tfrac{1}{2}\sigma^2\right) dt + \sigma dW \end{aligned}$$

i.e. $d \log S = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW$

We can now integrate this sum of linear differentials to get

$$\log S_t - \log S_0 = \left(\mu - \tfrac{1}{2}\sigma^2\right) t + \sigma W_t$$
$$\text{i.e. } S_t = S_0 e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W}$$

Note that unlike the deterministic case, the derivative of $\log S$ is not $1/S$. See model solutions to the exercises for more on this point.

μ represents the average rate of growth of the stock and σ the uncertainty. It is easy to check, using the moment generating function technique, that

$$\mathbb{E}_t S_T = S_t e^{\mu(T-t)}$$

Deriving the Black-Scholes Equation: Method I

The stock price and riskless cash evolve as follows

$$dS/S = \mu dt + \sigma dW$$
$$dB/B = r dt$$

Using Ito's lemma,

$$dV = \left(V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + \mu S V_S\right) dt + \sigma S V_S dW$$

Want to construct a self-financing replicating portfolio (x, y) of x cash and y stock for this option. The value of this portfolio is

$$\pi = xB + yS$$

In continuous time, self-financing means

$$d\pi = x dB + y dS$$

To replicate, we want V and π to follow the same SDE's.

$$d\pi = (xrB + y\mu S)\, dt + y\sigma S dW$$

Matching the volatility terms of dV and $d\pi$, we get $y = V_S$. Now matching drift terms, we get

$$V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} = xrB$$

With this choice, $\pi = V$, so we can write $xB = \pi - yS = V - SV_S$. Substituting this into the equation above gives:

$$V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0$$

which is the Black-Scholes equation.

Deriving the Black-Scholes Equation: Method II

Here is a method given by Wilmott, Howison and Dewynne.

$$dS/S = \mu dt + \sigma dW$$

Ito's lemma implies, as before:

$$dV = \left(V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + \mu S V_S\right) dt + \sigma S V_S dW$$

Now we construct a portfolio

$$\Pi = V - \Delta S$$

consisting of one option and short Δ of the stock. Then

$$\begin{aligned} d\Pi &= dV - \Delta dS \\ &= \left(V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + \mu S \left(V_S - \Delta\right)\right) dt + \sigma S \left(V_S - \Delta\right) dW \end{aligned}$$

Choose $\Delta = V_S$ to knock out the dW term. Now the portfolio grows at a deterministic rate. By no-arbitrage, this rate must equal r, i.e.

$$d\Pi = \left(V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS}\right) dt = r\Pi dt = r\left(V - SV_S\right) dt$$

Rearranging this equation gives Black-Scholes

$$V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0$$

Comparison of Methods I and II

In method 1, we chose a self-financing portfolio of cash and stock to replicate the option. In method 2, we chose a portfolio of option and stock to replicate cash.

What is Delta-Hedging?

In finance jargon, the option delta refers to the amount of stock needed to hedge the option position. This is shown more explicitly in method 2, where $\Delta = V_S$ is the amount of stock that needs to be *sold* to make the hedged portfolio Π instantaneously riskless. In method 1, the delta is given by $y = V_S$ and here it is the amount of stock that needs to be *bought*, since we are replicating the option and not cash as in method 2. In the perfect Black-Scholes world where uncertainty is driven by a standard Brownian motion and time is continuous, delta-hedging is a perfect strategy. More realistically, the delta-hedging eliminates leading order risk. See Bjork for a description of Gamma-neutral hedging.

Deriving the Black-Scholes Equation: Method III

In the market measure $\mathbb{P}$,

$$dS/S = \mu dt + \sigma dW = rdt + \sigma \left(dW + \frac{\mu - r}{\sigma} dt \right)$$

If $\frac{\mu-r}{\sigma}$ is sufficiently well-behaved (and it is for the constant-parameter classical Black-Scholes case), we can write $dW + \frac{\mu-r}{\sigma}dt$ as $d\tilde{W}$, a standard Brownian motion under an equivalent (and in this case unique) probability measure $\tilde{\mathbb{P}}$. For the interested reader, this is a consequence of Girsanov's theorem. Under this new measure,

$$dS/S = rdt + \sigma d\tilde{W}$$

Since the average rate of growth of the stock under this new measure is r, we recognise $\tilde{\mathbb{P}}$ to be the risk-adjusted measure. In fact, it can be verified that the relative price process $Z = B^{-1}S$ is a martingale under this measure. Now if $V(t, S)$ is the price of an option, then Ito's lemma gives:

$$dV = \left(V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S \right) dt + \sigma S V_S dW$$

In the risk-adjusted measure, all assets grow at the average rate r, i.e. we can write

$$\tilde{\mathbb{E}}\left[dV\right] = \left(V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S \right) dt = rV dt$$

Cancelling dt and rearranging gives the Black-Scholes equation

$$V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV = 0$$

Market Price of Risk or Sharpe Ratio

In the market measure, μ represents the average rate of growth of the stock. Since the stock is risky, we expect $\mu > r$. The difference $\mu - r$ measures the amount of risk inherent in the stock. σ, on the other hand, measures volatility, or market-wide risk. The quantity

$$\frac{\mu - r}{\sigma}$$

therefore represents the excess return over the risk-free rate for the stock, normalised by market volatility. This is called the market price of risk.

Solving the Black-Scholes PDE

The Black-Scholes PDE is

$$\begin{aligned} V_t + \tfrac{1}{2}\sigma^2 S^2 V_{SS} + rSV_S - rV &= 0 \\ V\left(T, S\left(T\right)\right) &= \psi\left(S\left(T\right)\right) \end{aligned}$$

Here, $V(t,S)$ is the price of the derivative at time t, and its value at time T is a deterministic function of the stock price at that time. We recognise this as a Kolmogorov backward PDE, and so the solution is simply given as follows:

$$V(t,S) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\psi(S(T))\right]$$

Here, we use the measure $\mathbb{Q}$ since the corresponding SDE is

$$\frac{dS}{S} = rdt + \sigma dW$$

The drift is r, which means we must be in the risk-adjusted measure. The solution of this SDE is

$$S(T) = S(t)\exp\left\{\left(r - \tfrac{1}{2}\sigma^2\right)(T-t) + \sigma W_{T-t}\right\}$$

Thus under the measure $\mathbb{Q}$, $S(T)$ is lognormally distributed with parameters $\left(r - \frac{1}{2}\sigma^2\right)(T-t)$ and $\sigma^2(T-t)$, conditional on $S(t)$. In principle at least, we can now evaluate the expectation and price the derivative in one step.

There is of course another way to solve the Black-Scholes PDE, by transforming it to a standard heat equation. A very good outline of this method can be found in Wilmott, Howison & Dewynne, so it is not repeated here.

Pricing a European Call Using Black-Scholes

A European call option is a contract that gives its holder the right (but no obligation) to buy a unit of the underlying stock for a specified price K, called the strike price, at a specified future date T called the maturity date, but not before that (hence, 'European'). The payoff from a European call is

$$\psi(S) = \max\{S-K, 0\}$$

Black and Scholes showed that the price of such an option is given by

$$C(t,S;T,K) = S\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

where $\Phi(.)$ stands for the cumulative distibution of a standard Gaussian random variable $N(0,1)$, and

$$d_{1,2} = \frac{\log(S/K) + \left(r \pm \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}$$

To prove this, we use the Kolmogorov solution

$$C(t,S;T,K) = e^{-r(T-t)}\mathbb{E}_t^{\mathbb{Q}}\left[\max\{S-K,0\}\right]$$

Write $w_X(x)$ for the probability density function of the random variable X. Then

$$\begin{aligned}
C &= e^{-r(T-t)} \int_0^\infty \max\{S-K,0\}\, w_S(S)\, dS \\
&= e^{-r(T-t)} \int_K^\infty (S-K)\, w_S(S)\, dS \\
&= e^{-r(T-t)} \int_K^\infty S w_S(S)\, dS - K e^{-r(T-t)} \int_K^\infty w_S(S)\, dS \\
&= \text{Term I - Term II}
\end{aligned}$$

$$\begin{aligned}
\text{Term II} &= K e^{-r(T-t)} \Pr\left[S(T) > K\right] \\
&= K e^{-r(T-t)} \Pr\left[\log S(T) > \log K\right] \\
&= K e^{-r(T-t)} \Pr\left[\log S(t) + \left(r - \tfrac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}\,\phi > \log K\right] \\
&= K e^{-r(T-t)} \Pr\left[\phi > \frac{\log(K/S(t)) - \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\
&= K e^{-r(T-t)} \Pr\left[\phi > \frac{-\log(S/K) - \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\
&= K e^{-r(T-t)} \left(1 - \Phi(-d_2)\right) \\
&= K e^{-r(T-t)} \Phi(d_2)
\end{aligned}$$

$$\begin{aligned}
\text{Term I} &= e^{-r(T-t)} \int_K^\infty S w_S(S)\, dS \\
&= e^{-r(T-t)} \int_{\log K}^\infty e^Y w_Y(Y)\, dY
\end{aligned}$$

where $Y = \log S$. Now since Y is normally distributed,

$$\begin{aligned}
\text{Term I} &= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma^2(T-t)}} \int_{\log K}^\infty \exp\left\{Y - \frac{\left[Y - \left(r - \frac{1}{2}\sigma^2\right)(T-t)\right]^2}{2\sigma^2(T-t)}\right\} dY \\
&= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_Y^2}} \int_{\log K}^\infty \exp\left\{Y - \frac{\left[Y - \mu_Y\right]^2}{2\sigma_Y^2}\right\} dY \\
&= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_Y^2}} \int_{\log K}^\infty \exp\left\{-\frac{1}{2\sigma_Y^2}\left(-2\sigma_Y^2 Y + Y^2 - 2\mu_Y Y + \mu_Y^2\right)\right\} dY \\
&= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_Y^2}} \int_{\log K}^\infty \exp\left\{-\frac{1}{2\sigma_Y^2}\left(Y^2 - 2\left(\mu_Y + \sigma_Y^2\right)Y + \left(\mu_Y + \sigma_Y^2\right)^2 - \left(\mu_Y + \sigma_Y^2\right)^2 + \mu_Y^2\right)\right\} dY \\
&= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_Y^2}} \int_{\log K}^\infty \exp\left\{-\frac{1}{2\sigma_Y^2}\left(\left[Y - \left(\mu_Y + \sigma_Y^2\right)\right]^2 - \left(\mu_Y + \sigma_Y^2\right)^2 + \mu_Y^2\right)\right\} dY \\
&= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_Y^2}} \int_{\log K}^\infty \exp\left\{-\frac{1}{2\sigma_Y^2}\left(\left[Y - \left(\mu_Y + \sigma_Y^2\right)\right]^2 - 2\mu_Y\sigma_Y^2 - \sigma_Y^4\right)\right\} dY \\
&= \frac{e^{-r(T-t)}}{\sqrt{2\pi\sigma_Y^2}} e^{\mu_Y + \frac{1}{2}\sigma_Y^2} \int_{\log K}^\infty \exp\left\{-\frac{1}{2\sigma_Y^2}\left[Y - \left(\mu_Y + \sigma_Y^2\right)\right]^2\right\} dY
\end{aligned}$$

Now we have

$$\begin{aligned}\mu_Y + \frac{1}{2}\sigma_Y^2 &= \log S(t) + \left(r - \frac{1}{2}\sigma^2\right)(T-t) + \frac{1}{2}\sigma^2 (T-t) \\ &= \log S(t) + r(T-t)\end{aligned}$$

Therefore,

$$e^{-r(T-t)} e^{\mu_Y + \frac{1}{2}\sigma_Y^2} = e^{\log S(t)} = S(t)$$

Additionally, if we write

$$\begin{aligned}U &= Y - \left(\mu_Y + \sigma_Y^2\right) \\ &= Y - \log S(t) - \left(r + \frac{1}{2}\sigma^2\right)(T-t)\end{aligned}$$

Then

$$\begin{aligned}&\frac{1}{\sqrt{2\pi\sigma^2 (T-t)}} \int_{\log K}^{\infty} \exp\left\{-\frac{1}{2\sigma_Y^2}\left[Y - \left(\mu_Y + \sigma_Y^2\right)\right]^2\right\} dY \\ &= \frac{1}{\sqrt{2\pi\sigma_Y^2}} \int_{\log K - \log S(t) - \left(r + \frac{1}{2}\sigma^2\right)(T-t)}^{\infty} \exp\left\{-\frac{U^2}{2\sigma_Y^2}\right\} dU\end{aligned}$$

Now if we write $Z = U/\sigma_Y$, then the above integral becomes

$$\begin{aligned}&\frac{1}{\sqrt{2\pi}} \int_{\frac{\log K/S(t) - \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}}^{\infty} e^{-\frac{1}{2}Z^2} dZ \\ &= \Pr\left[\phi > \frac{\log K/S(t) - \left(r + \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right] \\ &= \Pr\left[\phi > -d_1\right] \\ &= 1 - \Phi(-d_1) \\ &= \Phi(d_1)\end{aligned}$$

Therefore,

$$\text{Term I} = S\Phi(d_1)$$

Adding Term I and Term II, we have the result

$$C(t, S; T, K) = S\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2)$$

Why is there no μ in the Black-Scholes formula?

Black-Scholes prices derivatives in terms of the underlying stock. If we wanted to price the stock itself, we'd need to know μ. Stock returns are assumed to be normally distributed, and a normal distribution can be described completely by its mean and variance. Since we ultimately take expectations, we can knock out the first moment of returns, and so the price of a derivative is now explicitly dependent on the stock price itself and the second moment of returns, i.e. the volatility.

A useful analogy is that the definition of a metre itself might be fairly tricky, but once it is unambiguously defined, we can measure the length of a room precisely in terms of metres. Black-Scholes does not attempt to price stocks. Rather, it uses the price of the underlying as given and prices derivatives in terms of this underlying stock price.

References

1. 'Arbitrage Theory in Continuous Time' by Tomas Bjork covers basic stochastic calculus and the Black-Scholes model. See this book also for a discussion on the 'greeks' - delta, gamma, vega and rho.

2. 'The Mathematics of Financial Derivatives' by Wilmott, Howison and Dewynne has a good overview of the PDE techniques for solving the Black-Scholes equation.

3. 'Stochastic Calculus' by Bernt Oksendal has an excellent introduction to Ito calculus.