1 Deriving the Black–Scholes equation

Consider a financial asset/share with price S_t at time t that is modelled to evolve in time according to the log-normal SDE

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t. \tag{1}$$

Here the expected growth rate μ and the volatility σ are constants. We now ask a deceptively simple question: if at time t the asset price $S_t = s$ and if we plan to sell the asset at the later time $T > t$, then what is the fair present value of the asset? This innocuous-looking question leads to the most important result in mathematical finance.

1.1 The central conceit of mathematical finance

Let's try a very reasonable approach. At time T the asset price will be S_T and this is the amount of money we will obtain at that time. However, the present value of this future income must be discounted with the risk-free interest rate r, say, as otherwise we could do better by investing that amount of money in a risk-free bond. So, if we denote the fair present value of our asset by the function $u(s,t)$ then it is entirely reasonable to view the expectation

$$u(s,t) = \mathbb{E}\left[S_T e^{-r(T-t)} \,|\, S_t = s\right] \quad \text{for} \quad T \geq t \tag{2}$$

as the fair present value of our asset. However this will quickly lead to absurd consequences, so (2) is in fact wrong. Indeed, if (2) were correct then $u(s,t)$ would be the solution to the following PDE problem:

$$u_t + L_{s,t}u - ru = u_t + \mu s u_s + \frac{\sigma^2}{2} s^2 u_{ss} - ru = 0 \quad \text{with} \quad u(s,T) = s. \tag{3}$$

This is the backward equation with a discount term added and a boundary condition stating that the terminal pay-off is simply the asset price. Notably, the generator for a log-normal process involves only homogeneous derivative terms su_s and s^2u_{ss}, which is a reflection of the fact that our equations must be invariant under the simple scaling transformation $s \to \lambda s$ for arbitrary $\lambda > 0$. This must be so, because a price value such as s can always be changed in a inconsequential way by rescaling the currency (think of converting Lire to Euros). The mathematical upshot is that homogeneous derivative terms often give rise to solutions that consists of simple power laws or logarithms, which is why such solutions appear so often in mathematical finance. The problem is then usually to fit such solutions to boundary data that is not of the same form, but in the present case the boundary data is also a simple power law $u = s$, so this is easy. Hence, our problem is easily solved by assuming that $u = sf(t)$, which yields

$$u(s,t) = se^{(\mu-r)(T-t)} \tag{4}$$

as the solution to (3). Now, this is a horrible solution unless $\mu = r$. Suppose $\mu > r$, then (4) means that the fair value of the asset at time t *exceeds* the market price of the asset at the same time, and vice versa for $\mu < r$. This is hard to square with having a market at all. Moreover, as $T \to \infty$ the fair value goes to either infinity or zero. Clearly, this cannot be right, despite the seemingly reasonable assumptions underlying (2).

1.2 Arbitrage-free pricing of derivatives

To rectify the faults of (4) requires a new idea, which is the idea of arbitrage-free pricing. Demanding that a market is arbitrage-free (i.e., that it should not be possible to make a profit without risk) leads to many conclusions, but we will use only a single one: the rate of return on any risk-free asset must equal the global risk-free rate of return r.

Let us know consider the value of a fairly general derivative based on the asset S_t, which we denote by V_t. We will assume that there exists a function $v(s,t)$ such that

$$V_t = v(S_t, t). \tag{5}$$

In other words, our derivative depends only on the present asset price and on time; this class of derivative includes a lot of the standard financial options as well as the example of deferred asset selling in §1.1. Using Itô's formula, the derivative V_t evolves in time as

$$dV_t = [v_t(S_t,t) + L_{S_t,t}\, v(S_t,t)]\, dt + \sigma S_t v_s(S_t,t)\, dW_t. \tag{6}$$

Here the explicit notation is meant to highlight that v and its derivatives are evaluated along the trajectory S_t. Clearly, the evolution of V_t involves risk due to the final, random term. However, if we define an auxiliary derivative Y_t as a linear combination of V_t and S_t then we can arrange that this Y_t evolves without risk at the present time t. Specifically, if we set

$$Y_t = V_t - \triangle S_t \quad \text{with} \quad \triangle = \text{const} \tag{7}$$

then

$$\begin{aligned} dY_t &= [v_t + L_{S_t,t}\, v]\, dt + \sigma S_t v_s\, dW_t - \triangle[\mu S_t\, dt + \sigma S_t\, dW_t] \\ &= \left[v_t + \mu S_t v_s + \frac{\sigma^2}{2} S_t^2 v_{ss} - \triangle \mu S_t\right] dt + [\sigma S_t (v_s - \triangle)]\, dW_t. \end{aligned} \tag{8}$$

We can eliminate the noise term by setting

$$\triangle = v_s(S_t, t). \tag{9}$$

For constant $\triangle$ this only works at this single instant in time, of course, but that is sufficient to deduce the main result. So, with this choice of the "hedge" $\triangle$ we can make Y_t evolve risk-free at this instant in time, so therefore $dY_t = rY_t dt$ must hold by the principle of arbitrage-free markets. Combining this with (8) we find that this implies the relation

$$v_t + rS_t v_s + \frac{\sigma^2}{2} S_t^2 v_{ss} - rv = 0. \tag{10}$$

In the most famous cancellation of financial theory, the expected growth rate μ has dropped out! Now, this same construction can be made for arbitrary time t and for arbitrary values of S_t as well. Therefore we conclude that (10) must in fact hold for at all times and for all possible trajectories, in other words we can replace S_t by s and obtain the celebrated Black–Scholes equation

$$v_t + rsv_s + \frac{\sigma^2}{2} s^2 v_{ss} - rv = 0 \tag{11}$$

for the fair value of a derivative of the form (5). In the standard setting where there is a pay-off $\Phi(S_T)$ at the final time, the derivative problem consists of (11) together with the terminal condition

$$v(s,T) = \Phi(s). \tag{12}$$

Notice that this problem would be compatible with (2) and (3) only if the expected rate of growth of the asset were r instead of μ. The point of view of retaining (2) by replacing μ with r in (1) is called the "risk-neutral" view, but to me it is unclear whether there is any meaning attached to it beyond this formal observation.

Finally, the linear payoff $\Phi = s$ that occurred in (3) now leads to the trivial solution $v = s$, which unlike (4) does not depend on μ, r, or T! In other words, according to arbitrage-free pricing, the fair present value of an asset is its present market price, as it should be. As an amusing aside, in this example the hedge $\triangle$ is in fact constant in time and equal to one, so the risk-free portfolio $Y_t = S_t - S_t = 0$ that arose in the derivation of (11) is actually empty.