Black and Scholes-Merton Model

I. Derivation of Black-Scholes Formula

Model Assumptions:

- $dS = \mu S dt + \sigma S dZ$
- No dividends
- Markets are friction-free:
  - ✓ No taxes
  - ✓ No Transaction costs
  - ✓ No short sales restrictions
  - ✓ Assets are divisible at will
  - ✓ Continuous trading
- $r_{borrowing} = r_{lending} = \text{constant}$

Derivative Pricing:

a. Let f(S,t) be the price of the derivative & use Ito's lemma:

$$df(S,t) = \left\{ \frac{\partial f(S,t)}{\partial t} + \frac{\partial f(S,t)}{\partial S}\mu S + \frac{1}{2}\frac{\partial^2 f(S,t)}{\partial S^2}\sigma^2 S^2 \right\} dt + \left\{ \frac{\partial f(S,t)}{\partial S}\sigma S \right\} dZ$$

b. Form a portfolio containing -1 unit of the derivative and $\Delta \equiv \frac{\partial f(S,t)}{\partial S}$ shares of the underlying stock S.

c. At time t (now), the portfolio value is: $\Pi \equiv \Delta S - f(S,t)$

d. Finally, compute the change in portfolio value:

$$
\begin{aligned}
d\Pi(S,t) &= \Delta\, dS - df(S,t) \\
&= \Delta \mu S dt + \Delta \sigma S dZ - \left\{ \frac{\partial f(S,t)}{\partial t} + \frac{\partial f(S,t)}{\partial S}\mu S + \frac{1}{2}\frac{\partial^2 f(S,t)}{\partial S^2}\sigma^2 S^2 \right\} dt - \left\{ \frac{\partial f(S,t)}{\partial S}\sigma S \right\} dZ \\
&= \frac{\partial f(S,t)}{\partial S}\mu S dt + \frac{\partial f(S,t)}{\partial S}\sigma S dZ - \left\{ \frac{\partial f(S,t)}{\partial t} + \frac{\partial f(S,t)}{\partial S}\mu S + \frac{1}{2}\frac{\partial^2 f(S,t)}{\partial S^2}\sigma^2 S^2 \right\} dt - \left\{ \frac{\partial f(S,t)}{\partial S}\sigma S \right\} dZ \\
&= -\left\{ \frac{\partial f(S,t)}{\partial t} + \frac{1}{2}\frac{\partial^2 f(S,t)}{\partial S^2}\sigma^2 S^2 \right\} dt
\end{aligned}
$$

Since there is no stochastic component, the portfolio is thus riskless. Therefore we must have: $d\Pi(S,t) = r\,\Pi\, dt$ since the portfolio must earn exactly the risk-free rate (by non-arbitrage argument).

In other words,

$$
-\left\{ \frac{\partial f(S,t)}{\partial t} + \frac{1}{2}\frac{\partial^2 f(S,t)}{\partial S^2}\sigma^2 S^2 \right\} = r\left\{ \frac{\partial f(S,t)}{\partial S} S - f(S,t) \right\}
$$

or: $$\frac{\partial f(S,t)}{\partial t} + \frac{\partial f(S,t)}{\partial S} rS + \frac{1}{2}\frac{\partial^2 f(S,t)}{\partial S^2}\sigma^2 S^2 - r f(S,t) = 0$$

Note that since we did not specify what type of derivative was used, the equation above must be satisfied by EVERY derivative written on stock S.

Given an underlying process for S, what differentiates various derivatives written on that stock S is the type of initial/boundary conditions:

a. For a call option with exercise price K: $f(S,T) = \max(0, S-K)$

b. For a put option with exercise price K: $f(S,T) = \max(0, K-S)$

c. For a futures contract with delivery price K: $f(S,T) = S - K$

In the case of a European Call option, the solution is:

$$C(S,K,t,T) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

- N(.) is the cumulative normal distribution
- $d_1 = \frac{Ln(S) - Ln(K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$
- $d_2 = d_1 - \sigma\sqrt{T-t}$

(Black-Scholes formula)

By Put-Call parity, the value of a put option can easily be obtained: $P(S,K,t,T) = C(S,K,t,T) - S + Ke^{-r(T-t)}$

Finally, note that the non-arbitrage argument is very general and can be used to price any derivative.

Exercise: Prove graphically that at expiration we have:

$$P(S,K,T,T) = C(S,K,T,T) - S + Ke^{-r(T-T)}$$

In other words, show graphically that at expiration we have:

$$P(S,K,T,T) + S = C(S,K,T,T) + K$$

II. Extensions of the Model

Relaxing the Log-Normal Assumption:

In reality, stock returns exhibit outliers too often to be consistent with the assumption made in part I.

1. Constant Elasticity of Variance Model:

a. Instead of $dS_t = \mu S_t dt + \sigma S_t dZ_t$, assume
$dS_t = \mu S_t dt + \sigma S_t^{1-\alpha} dZ_t$ where α is a constant parameter.

b. This allows the instantaneous rate of return dS/S to have a diffusion component that depends on the level of S, the stock price. Why?
Because $\frac{dS_t}{S_t} = \mu dt + \sigma S_t^{-\alpha} dZ_t = \mu dt + \frac{\sigma}{S_t^{\alpha}} dZ_t$
For α >0, the volatility of dS/S is INCREASING when S is DECREASING, a convenient property since this is consistent with empirical observations.

c. Except for some special cases where α =0,1/2, or 1, there is usually no closed-form solution.

2. Jump-diffusion model:

a. Jumps are often modeled as a Poisson process where the intensity or probability of a jump is determined by λ. The probability of a jump between time t and t+dt is thus λ dt and the probability of a jump not occurring is therefore 1-λ dt.

b. If a jump happens and nothing else (pure jump), the stock price jumps from S_t to $S_{t+\Delta t}=Y_tS_t$, so $\Delta S_t=(Y_t-1)S_t$ where Y_t-1 is the random jump size (percentage price change or return in decimal format). Y_t is often lognormal (i.e. Y_t=exponential(X) where X is normally distributed) so that Y_t-1 has a range between -1 and infinity, consistent with a return (since percentage returns vary between -100% and infinity).

c. The jump-diffusion model is a diffusion model where a Poisson jump is added:

Replace $dS_t = \mu S_t dt + \sigma S_t dZ_t$ with the following:

$dS_t = (\mu - \lambda\kappa)S_t dt + \sigma S_t dZ_t + S_t dq_t$ where

- μ = expected return of the stock

- q = Poisson process generating the jumps, with dq independent of dZ and with the various jumps independent over time, and with
$$\begin{cases} dq_t = (1)(Y_t - 1) = Y_t - 1 & \text{with } p = \lambda dt \\ dq_t = (0)(Y_t - 1) = 0 & \text{with } p = 1 - \lambda dt \end{cases}$$

- $\kappa = E[Y-1]$ = average jump size as a percentage (but in decimal format) of the stock price S.

- Ln(Y) is normally distributed with variance δ^2.

Note that $\frac{dS_t}{S_t} = (\mu - \lambda\kappa)dt + \sigma dZ_t + dq_t$

Therefore $E(dS_t / S_t) = E[(\mu - \lambda\kappa)dt] + E[\sigma dZ_t] + E[dq_t]$

so $E(dS_t / S_t) = \mu dt - \lambda\kappa dt + 0 + (\lambda dt)\overbrace{E[Y_t - 1]}^{\kappa} + (1 - \lambda dt)(0) = \mu dt$ ✓

Therefore:

$$\begin{cases} dS_t = (\mu - \lambda\kappa)S_t dt + \sigma S_t dZ_t \ \text{ if the Poisson event does not occur} \\ dS_t = (\mu - \lambda\kappa)S_t dt + \sigma S_t dZ_t + (Y_t - 1)S_t \ \text{ if the Poisson event occurs} \end{cases}$$

In order to generate dq_t values in a computer program, one must be able to generate 1s and 0s with probabilities p=λdt and 1-p=1-λdt. One must therefore generate 1s and 0s from the Bernoulli distribution. The Bernoulli distribution is itself a special case of the Binomial distribution where the number of trials n=1.
The binomial (discrete) distribution represents the probabilities of k successes in n trials where each trial has a success probability p.
In Matlab, the Binomial distribution random number generator function is binornd(n,p,rows,columns) and returns a (rows x columns) matrix of randomly generated k values given n trials and a probability of success p for each trial. Therefore by setting n=1, the only two possible values for the number of successes k are 0 and 1, just like in the Bernoulli distribution.

The dq values can therefore be generated as:
dq = binornd(1,λdt,T/dt,1) [Y_t-1]
This will yield a (T/dt x 1) column of zeros (with probability 1-λdt) and ones (with probability λdt) that get multiplied by [Y_t-1].

Relaxing the Constant Interest Rate Assumption:

In practice, unless the option is deep in-the-money and far away from maturity, the stochasticity of interest rates does not matter much for pricing purposes.

Relaxing the Constant Volatility Assumption:

The volatility parameter can itself be made a stochastic process, as in Heston (1993):

- Stock price: $dS_t = \mu S_t\, dt + \sqrt{v_t}\, S_t\, dW_t^S$
- Variance $v_t \equiv \sigma_t^2$: $dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t}\, dW_t^v$

where

- μ is the rate of return of the underlying asset/stock.
- θ is the long variance, or long-run average price variance; as t tends to infinity, the expected value of ν_t tends to θ.
- κ is the rate at which ν_t reverts to θ (speed of mean reversion).
- ξ is the volatility of the volatility, or 'vol of vol', and determines the variance of ν_t.

Note that the stochastic volatility feature makes it more difficult to obtain a closed-form solution.

III. Volatility Parameter Estimation

• Volatility is the only unobserved variable but is nevertheless a very important one as option prices are sensitive to it.

• We thus need ways to attempt to measure it.

A. Standard Deviation of Returns

Recall that the Black-Scholes setup implies:

a. $\text{Ln}\left(S_{t+\tau}/S_t\right) = \left\{\text{Ln}\left(S_{t+\tau}\right) - \text{Ln}\left(S_t\right)\right\} \sim N\left(\left\{\mu - \sigma^2/2\right\}\tau,\ \sigma^2\tau\right)$

b. Thus record all the stock prices S_{t_i} at τ intervals, and by computing the log-returns between the intervals, calculate:

$$\tau\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left\{\mathrm{Ln}\left(S_{t_i+\tau}/S_{t_i}\right) - \overline{\mathrm{Ln}\left(S_{t+\tau}/S_t\right)}\right\}^2$$

c. Since this yields the variance over a τ interval, the last step is to compute:

$$\left\{\hat{\sigma}^2 \text{ over one year}\right\} = \frac{1}{\tau}.\left\{\hat{\sigma}^2 \text{ over interval } \tau \text{ year}\right\}$$

(In order to annualize the variance)

Note that the more data we have, the more precise the estimate. To increase the amount of data, one can either increase the time period or increase the frequency of data points recorded. However, too long a period is not a good thing as volatility changes over time. Also, too high a frequency may pick up intraday microstructure perturbances that can bias the results.

B. Estimation of Stochastic Volatility Models

Section III A dealt with cases where the volatility is assumed to be constant. How do we estimate parameters in the stochastic case? Answer: it all depends on the model.

1. Example of a *simple* stochastic volatility model:

$$\begin{cases} dS_t = \mu S_t dt + \sigma_t S_t dZ_t \\ d\sigma_t = \delta\left(\theta - \sigma_t\right)dt + \kappa dW_t \end{cases}$$

a. The volatility is said to follow an Ornstein-Uhlenbeck process, and it can be shown that the conditional distribution of $\sigma_{t+\tau} | \sigma_t$ exhibits a normal distribution with the following characteristics:

- $E\left[\sigma_{t+\tau} | \sigma_t\right] = \theta + e^{-\delta\tau}\left(\sigma_t - \theta\right)$
- $V\left[\sigma_{t+\tau} | \sigma_t\right] = \frac{\kappa^2}{2\delta}\left\{1 - e^{-2\delta\tau}\right\}$

b. Notice that the distribution is known with certainty (the normal distribution is perfectly defined by two parameters only: the mean and the variance), and so if the volatility was observed we could perform a Maximum-Likelihood estimation.

c. A Quasi-Maximum Likelihood estimation (maximizing a function related to the logarithm of the likelihood function but not exactly equal to it, often a simplified/approximated version of it) is also possible.

By Taylor first-order expansion, we have $e^{-\delta\tau} = 1 - \delta\tau + o(\tau)$, so $E\left[\sigma_{t+\tau} | \sigma_t\right] = \delta\tau\theta + \left(1 - \delta\tau\right)\sigma_t$ is an AR(1) model that can be estimated in the following manner:

- Estimate the time series of volatilities from the data.

- Then regress: $\hat{\sigma}_{t+\tau} = \underbrace{\delta\tau\theta}_{\equiv\omega} + \underbrace{\left(1 - \delta\tau\right)}_{\equiv\rho}\hat{\sigma}_t + \upsilon_t$ to obtain estimates for ω and ρ. Then, knowing τ and the ρ regression estimate gives us δ. Also, knowing τ and δ and the ω regression estimate gives us θ. Finally, the variance of the residuals along with knowledge of δ and τ yields an estimate of κ.

d. Generalized Method of Moments (GMM)

$$\sigma_{t+\tau} - E\left[\sigma_{t+\tau} \mid \sigma_t\right] = \sigma_{t+\tau} - \left[\theta + e^{-\delta\tau}\left(\sigma_t - \theta\right)\right] = \varepsilon_{t+\tau}$$

$$\varepsilon_{t+\tau} \approx N(0, V\left[\sigma_{t+\tau} \mid \sigma_t\right])$$

$$\mathrm{Var}(\varepsilon_{t+\tau}) = V\left[\sigma_{t+\tau} \mid \sigma_t\right] = \frac{\kappa^2}{2\delta}\left\{1 - e^{-2\delta\tau}\right\}$$

$$\mathrm{Var}(\varepsilon_{t+\tau}) = E(\varepsilon_{t+\tau}^2) - [E(\varepsilon_{t+\tau})]^2 = E(\varepsilon_{t+\tau}^2)$$

$$f_t(\Theta) = \begin{bmatrix} \varepsilon_{t+\tau} \\ \varepsilon_{t+\tau} \times \sigma_t \\ (\varepsilon^2_{t+\tau} - V\left[\sigma_{t+\tau} \mid \sigma_t\right]) \\ (\varepsilon^2_{t+\tau} - V\left[\sigma_{t+\tau} \mid \sigma_t\right]) \times \sigma_t \end{bmatrix} \quad \text{so} \quad E\left[f_t(\Theta)\right] = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$g_T(\Theta) = \frac{1}{T}\sum_{t=1}^{T} f_t(\Theta) \quad \text{T vectors of size (4x1) averaged over time}$$

$$J_T(\Theta) = g_T^{'}(\Theta) W_T(\Theta) g_T(\Theta)$$

The quadratic form $J_T(\theta)$ is being minimized by the parameters. Note that the matrix $W_T(\theta)$ is called the "weighing" matrix.

Regardless of the method used, once the parameters have been estimated, they can then be plugged into the stochastic volatility option pricing model (closed-form formula if it exists, or in a numerical pricing technique such as Monte Carlo valuation for example) and be used to price options.