The Stock Market

Daniel Lyalin

Introduction

Figure 1: Image by Mandy Li

Invented by Fischer Black and Myron Scholes, the Black-Scholes stock valuation model is used by many major corporations, including General Motors and 3M, to estimate the future prices of European style stock options. The model relies on a solution of the Black-Scholes partial differential equation known as the Black-Scholes Formula, a formula reliant on the properties of Brownian motion. Hence, to understand the derivation and usage of the model, we must understand Brownian motion, starting with probability distributions.

Probability Distributions

The simplest and first major probability distribution is the Bernoulli distribution. This is characterized as a discrete random variable with boolean outcomes of *0* and *1*, corresponding to either a "failure" or "success." The chances of success, or *P(1)*, are usually denoted with the lowercase "p," which has to be a discrete, decimal number between *0* and *1*. If the probability of success is p, then the probability of failure, or *P(0)*, is denoted as $1-p$, or the decimal complementary to p. The mean, or expected value, of the Bernoulli distribution is hence p, as the expected value of a Bernoulli variable with $P(1) = p$ and $P(0) = q$ (or $1-p$) can be written as :

$$E(X) = P(1) \cdot 1 + P(0) \cdot 0 = p \cdot 1 + q \cdot 0 = p \tag{1}$$

The variance of a Bernoulli distribution is pq, as:

$$E[X^2] = P(1) \cdot 1^2 + P(0) \cdot 0^2 = p \cdot 1^2 + q^2 = p \tag{2}$$

$$Var[X] = E[X^2] - E[X]^2 = p - p^2 = p(1-p) = pq \tag{3}$$

Following from the Bernoulli distribution is the binomial distribution, a discrete random variable that measures the number of successes in a sequence of independent experiments; it is a variation of the Bernoulli distribution repeated several times, with each repetition having a boolean outcome with probabilities p and $1-p$. The binomial distribution requires a fixed number of trials n, boolean outcomes for each trial p and q $(1-p)$, and independence. As both n and p can be controlled, the probability mass function, or the probability of the binomial distribution yielding discrete random variable with value X, for the binomial distribution, where k is the number of expected successes, is:

$$P(k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{4}$$

As an example, the probability of landing heads six times when tossing a fair coin *10* times is written as:

$$P(6) = \binom{10}{6} \cdot 0.5^{0.6} \cdot 0.5^4 = 0.21 \tag{5}$$

The mean value of the binomial distribution is np, as it is the summation of the expected probability p of success for each event, of which there are n. Thus, p added to itself n times can be written as np. Similarly, the variance of the binomial distribution is the sum of the variances of each experiment. As each experiment is modeled by a Bernoulli distribution, the summation of variances is $np(1-p)$.

An expansion of the binomial distribution to continuous variables is the uniform continuous distribution, which models the probability of getting X in range $[a, b]$ where probability is uniformly distributed along the range. The probability mass function for the uniform continuous distribution is:

$$P(X) = \frac{1}{b-a}$$

for all X in range $[a, b]$. For example, the probability $P(4 < X < 15)$ in range $[3, 40]$ is:

$$P(4 \leq X \leq 15) = (15-4) \cdot \frac{1}{40-3} = \frac{11}{37} \tag{6}$$

The expected value of the uniform continuous distribution can be found by taking the average of a and b: $\frac{1}{b-a}$. Finding the variance is more involved:

$$Var[X] = E[X^2] - E[X]^2 \tag{7}$$

$$\begin{aligned}
E[X^2] &= \int_{-\infty}^{\infty} x^2 f(x)\,dx \\
&= \int_a^b x^2 \cdot \frac{1}{b-a}\,dx \\
&= \frac{1}{b-a} \cdot \int_a^b x^2\,dx \\
&= \frac{1}{b-a}[\frac{x^3}{3}]_a^b \\
&= \frac{1}{b-a} \cdot \frac{1}{3} \cdot [b^3 - a^3] \\
&= \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} \\
&= \frac{b^2+ab+a^2}{3}
\end{aligned}$$

$$\begin{aligned}
E[X]^2 &= \left(\frac{b+a}{2}\right)^2 \\
&= \frac{b^2+2ab+a^2}{4}
\end{aligned}$$

$$\begin{aligned}
Var[X] &= \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4} \\
&= \frac{4b^2+4ab+4a^2-3b^2-6ab-3a^2}{12} \\
&= \frac{b^2-2ab+a^2}{12} \\
&= \frac{(b-a)^2}{12}
\end{aligned}$$

Expanding from the uniform continuous distribution, we arrive at the normal distribution. This is a continuous probability distribution with continuous real random variables that models data symmetric about a mean value. The probability density function of the normal distribution, or the function that gives the relative likelihood of a random value taking a certain value in its sample space, is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \tag{8}$$

where σ is the standard deviation of the data and μ is the mean of the data. A special case of the normal distribution, known as the standard normal distribution, is characterized by $\sigma = 1$ and $\mu = 0$, with the probability density function changing to be:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \tag{9}$$

Using Code and Confidence Intervals

To get the next element of probability necessary to understanding Brownian motion, and to see an example of how Python can be used to find values, let's try to estimate the value of pi.

If we draw a circle centered at the origin with radius r, then we know that its area is modelled by the equation $A = \pi r^2$. Following from this, if we can find the area of the circle without using pi, we can divide that area by the r^2 to get back pi. To estimate the area of this circle, we can use Python.

The first in doing this is generating two arrays of x- and y-values in the range $[-r, r]$, such that the two points furthest from the origin are $(-r, -r)$ and (r, r). For simplicity, we can use the unit circle, where $r = 1$. In Python, we'd write this as:

```
import numpy as np

x = np.random.rand(10000000)
y = np.random.rand(10000000)
```

Next, we can use the distance formula to create a new array filled with the distances from each of our x-y pairs to the origin. Because the origin's x- and y-value are *0*, the distance formula can be simplified to $x^2 + y^2$. In Python:

```
distance = x ** 2 + y ** 2
```

Finally, we can count the number of values within the array that are less than or equal to r, which indicates points that fall within or on the circumference of the circle. This summation returns the area of the circle, as the area equals the total amount of points enclosed within the circle. If we take the count and divide the sum by r^2, and multiply the product by *4* to account for quadrants other than the first (by squaring our point values, we effectively limit ourselves to positive values in the first quadrant), we get π. In Python:

```
np.count_nonzero(distance <= 1) / n * 4
```

The first time I ran this, Python returned: *3.140942*, an approximation of π accurate to the hundredth place. If you increase n to a greater value, the approximation will become increasingly accurate. Formulaically, this algorithm can be represented as:

$$\pi = \lim_{r \to \infty} \frac{1}{r^2} \sum_{x=-r}^{r} \sum_{y=-r}^{r} \begin{cases} 1 & \text{if } \sqrt{x^2 + y^2} \leq r \\ 0 & \text{if } \sqrt{x^2 + y^2} > r \end{cases} \tag{10}$$

The first apparent issue with this approach is that, by virtue of generating random values, our answers are noisy and inconsistent. We can fix this using confidence intervals. These are ranges within which we know our answer must lie. Using the variance (σ^2) of our approximations, the number of trials (n) we ran, and the average value of our approximations ($\bar{x}$), the confidence intervals can be expressed as:

$$\bar{x} - 2.58 \cdot \sqrt{\frac{\sigma^2}{n}} \text{for the lower bound} \tag{11}$$

$$\bar{x} + 2.58 \cdot \sqrt{\frac{\sigma^2}{n}} \text{for the upper bound} \tag{12}$$

In Python, this can be found using the following code:

```
answers = np.empty(100)
for i in range(100):
    answers[i] = approximate_pi()
      variance = np.var(answers)
mean = np.mean(answers)
lower = mean - 2.58 * np.sqrt(variance / 100)
upper = mean + 2.58 * np.sqrt(variance / 100)
```

Note the use of the constant in *2.58* in both equations. This value is derived from the standard distribution. In statistics, z-scores measure how many standard deviations away from the mean a data point lies; within the normal distribution, 99% of all values have a z-score between -*2.58* and *2.58* standard deviations of the mean. Hence, adding (or subtracting) *2.58* allows you to determine a range within which the value of pi lies with 99% certainty.

Markov Chains

Though we can now effectively calculate probabilities for single states, Markov Chains are useful in that they allow us to work with a sequence of states where the next state is dependent only on the current state. Because of this, the probability of getting any state next is dependent on the current state and time elapsed; the Markov property, for which Markov Chains are named, is exactly this: Markov Chains are "memory-less." Apart from the Markov property, Markov Chains have two other essential properties: state spaces and step probabilities.

The state space of a Markov Chain is the set of all possible states that the current state can transition to. In a dice roll, for example, the state space would be: $X_t = 1, 2, 3, 4, 5, 6$. Our current state within the Markov Chain in this example would be the value of the previous roll. As states as are ordered by time, our current state would be written as X_t and the next two states would be X_{t+1} and X_{t+2}, respectfully.

The step probabilities of a Markov Chain are the set of probabilities for transitioning from one state to another. Probabilities are written as $P(X_{t+1}|X_t)$. For example, $P(X_{t+1} = 2|1)$ denotes the probability of rolling a two given we just rolled a one. For every given state, we need a set of step probabilities for transitioning from that state to every state in the state space. For example, if we rolled a one and our set of step probabilities was $(0, 1, 0, 0, 0, 0)$, then we could deduce that we are guaranteed to roll a two next. In reality, the step probabilities for transitioning from a one would be $(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6})$.

The step probabilities are stored in a transition matrix P_t characterized by the state space S where element (i, j) is such that:

$$(P_t)_{ij} = \mathbb{P}(X_{t+1} = j|X_t = i) \tag{13}$$

For example, p_{12} would represent the probability of transitioning from a one to a two.

Each row of the transition matrix is called a probability vector, and the elements of each probability vector must sum up to one. An example probability matrix, for example for transitioning from one type of weather to any other, could look like:

$$P = \begin{bmatrix} .5 & .25 & .25 \\ .5 & 0 & 0.5 \\ .25 & .25 & .5 \end{bmatrix} \tag{14}$$

It should be noted that the transition matrix is not necessarily symmetric, meaning $p_{cr} \neq p_{rc}$; this means the probability of transiting from state r to state c is not necessarily equal to the probability of transitioning from state c to state r.

Stochastic Processes and Brownian Motion

Markov Chains are a discrete process in that they have a finite number of states; stochastic processes, on the other hand, are a process defined by random continuous transitions, meaning that random transitions occur between states that are elements of real, rather than integral, numbers. In stochastic processes, time is also treated as continuous.

Brownian motion is a stochastic process defined by the summation of many standard, normal random variables. Mathematically, it is defined by the formula: $Z_t = \sum_{i=0}^{t} N_i$, where N_i is a value of the cumulative function at i of the standard normal distribution. Any given state in Brownian motion is defined as Z_t, which indicates the value of Z at time t. The two defining characteristics of Brownian motion are:

1. $Z_0 = 0$
2. $Z_{t+s} - Z_s = N(0, t)$

The first of these characteristics is an assumption used to simplify the process, though the first state can be equal to any number without violating the rules of Brownian motion. The second characteristic means that for any $0 \leq s, t \leq \infty$ the distribution of the increment $Z_{t-s} - Z_s$ has the same distribution as $Z_t - Z_0 = Z_t$, that distribution being the normal distribution with mean *0* and variance t. The second characteristic is best understood through its derivation.

Let's start with a set of continuous values characteristic of Brownian motion. If we were to discretize the values into n amounts separated by Δt time and select two random times t and $t + \Delta t$, then the change from the first state to the second could be written as:

$$Z_{t+\Delta t} - Z_t = \Delta Z_t \tag{15}$$

This means that the change between the two states is equal to the change since the first state. This ΔZ_t can be rewritten as the product of a random variable scaled against $\sqrt{\Delta t}$. We chose to take the square root because it both grows and decreases at a slower rate than our time steps, preventing our function from freezing or growing too fast and preserving continuity. Our random variable will be taken from the standard normal distribution, meaning the equation will be:

$$\Delta Z_t = N_t \sqrt{\Delta t} \tag{16}$$

where N_t represents the value of the random variable drawn from the standard normal distribution at time t. The change since time zero can thus be written as:

$$\Delta Z_0 = N_0 \sqrt{\Delta t} \tag{17}$$

Knowing this, we know that the value at our first discrete time step is equal to:

$$Z_{\Delta t} = Z_0 + \Delta Z_0 \tag{18}$$

In other words, the value at our first time step is equal to the value at the previous step plus the change since that step. As we've decided that the value of the first step is always equal to zero, the value at the first time step has to be equal solely to the change since the first time step. Hence:

$$Z_{\Delta t} = N_0\sqrt{\Delta t} \tag{19}$$

The value at the second step must then equal the value at the first time step before it plus the change since the first step, or:

$$Z_{2\Delta t} = \Delta Z_{\Delta t} + Z_{\Delta t} = (N_0 + N_{\Delta t})\sqrt{\Delta t} \tag{20}$$

The value at any time step can be expressed as the summation of the values of the normal distribution at all the previous times multiplied by $\sqrt{\Delta t}$, similarly to how the value at the second time step could be expressed as the values of the normal distribution at time one and time zero multiplied by $\sqrt{\Delta t}$. Knowing that the change since any time step equals $N_t\sqrt{\Delta t}$, the expected value of change is always zero, or the mean value of the standard normal distribution, and the variance is $\sqrt{\Delta t}$.

Our final time, T, can be written as the summation of all the previous random variables times $\sqrt{\Delta t}$. While the expected value is still zero, the variance is now $(\sqrt{\Delta t})^2 n$, or $\Delta t n$, which is equal to T. Because of this, the value at T equals the normal distribution with mean zero and variance T. Working backwards from this, we can deduce the property $Z_t - Z_s = N(0, t-s)$. This makes sense, as we would expect later times to have greater variance, as randomness accumulates over time.

Another explanation of the properties of Brownian motion, as well as proofs, can be found at:

https://galton.uchicago.edu/~lalley/Courses/313/BrownianMotionCurrent.pdf

In differential form, Brownian motion can be expressed as:

$$\Delta Z_t = Z_{t+\Delta t} - Z_t = N(0, \Delta t) \tag{21}$$

In calculus, this is written as:

$$dZ_t = N_t = \text{Normal}(0, dt) \tag{22}$$

with dt approaching zero in the limit to account for continuous time. It should be noted that Brownian motion as we've defined it is a stationary process in that its probability distribution does not change with time; this implies that mean and variance remain constant. However, we can redefine Brownian motion such that it "drifts" depending on the mean, meaning both the mean and variance of the Normal distribution are dependent on the time. This is written as:

$$\Delta Z_t = \mu\Delta t + N_t \tag{23}$$

and alternatively as:

$$dZ_t = \mu dt + N_t \tag{24}$$

This implies:

$$Z_T = \text{Normal}(\mu T, T) \tag{25}$$

As we'll be using Brownian motion to model stocks, it would be more effective to write this equation using percentages, which can be done using Geometric Brownian Motion:

$$\frac{\Delta S_t}{S_t} = \mu \Delta t + N_t \tag{26}$$

This is alternatively written as:

$$\frac{dS_t}{S_t} = \mu dt + N_t \tag{27}$$

In this equation, μ now means percentage, rather than number, growth and S_t is now log-normal, meaning it is always positive and has a long tail. Finally, we can scale the normal term to scale deviation by multiplying it by a constant:

$$\frac{\Delta S_t}{S_t} = \mu \Delta t + \sigma N_t \tag{28}$$

Alternatively:

$$\frac{dS_t}{S_t} = \mu dt + \sigma N_t \tag{29}$$

This form can be rewritten as:

$$dS_t = S_t(\mu dt + \sigma N_t) \tag{30}$$

Black-Scholes Equation

The Black-Scholes Equation rests on several assumptions:

- The system contains one risky asset.
- The system contains one risk-free asset.
- The rate of return on the risk-free asset is constant.
- The returns of the risky asset have a path modelled by dilated geometric Brownian motion.
- The risky asset does not pay a dividend.
- There are no other risk-free assets available.
- It is possible to borrow and lend any amount of cash at the same rate as the return rate of the risk-free asset.
- It is possible to buy and sell any amount of stock.
- There are no transaction costs.

The Black-Scholes equation is a partial differential equation used to calculate the price evolution of European call and put options. Call options give the owner the right to buy a stock at an agreed-upon "strike price," and put options the right to sell at that strike price. If the price of the stock is above the strike price, then the call option owner can make a profit by buying at the strike price and selling at the actual, higher price; similarly, if the price of the stock is less than the strike price of a put option, then buying the stock and selling it at the option price results in profit. The variables used in the equation are:

- r, the risk-free interest rate
- σ, the standard deviation, or *volatility*, of the returns of the stock
- S, the stock price
- t, time
- V, the option price as a function of S and t
- μ, the average rate of growth, or *drift*, of the option

The equation is written as follows:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\mu^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{31}$$

It can be better understood if it is written as:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\mu^2 S^2 \frac{\partial^2 V}{\partial S^2} = rV - rS\frac{\partial V}{\partial S} \tag{32}$$

The left side of the equation models the change in the price of V with respect to time plus the convexity of the option's value with respect to the price of the stock (convexity is the second derivative of the stock price with respect to interest rates). The right hand side is the risk-free return of a long and a short position in the option ; a long position is when one owns the stock, a short when one borrows. Over any infinitesimal time interval, the losses and gains of the various terms cancel out and result in a risk-free investment,

To derive the equation, we need to use Ito's lemma, an identity used in Ito calculus (the calculus of stochastic processes) to find the derivative of a time-dependent stochastic process. Ito's lemma is written as:

$$df = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{\sigma_t^2}{2}\frac{\partial^2 f}{\partial x^2}\right) dt + \sigma_t \frac{\partial f}{\partial x} dB_t \tag{33}$$

The derivation of Ito's lemma begins by defining *f(t,x)* to be a twice-differentiable scalar function; the Taylor expansion of this function is written as:

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}dx^2 + \cdots \tag{34}$$

If we take our equation for geometric Brownian motion $dX_t = \mu_t dt + \sigma_t B_t$ and substitute X_t in the place of x and the formula for dx, then the Taylor expansion becomes:

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}(\mu_t dt + \sigma_t dB_t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\mu_t^2 dt^2 + 2\mu_t \sigma_t dt dB_t + \sigma_t^2 dB_t^2) + \cdots \tag{35}$$

As we take the limit of this equation as dt approaches zero, the terms dt^2 and $dtdB_t$ approach zero faster than dB^2, meaning we can set them equal to zero. Brownian motion has a property called quadratic variance, which means that

$$\int_0^t (dB(s))^2 = t \tag{36}$$

and

$$(dB(t))^2 = dt \tag{37}$$

For this reason, we can substitute the dt term for dB^2 and factor out the remaining dt and dB terms to be left with Ito's lemma:

$$df = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{\sigma_t^2}{2}\frac{\partial^2 f}{\partial x^2}\right) dt + \sigma_t \frac{\partial f}{\partial x} dB_t \tag{38}$$

If we take our equation for geometric Brownian motion $\frac{dS_t}{S_t} = \mu dt$ and replace N_t with dW to represent any stochastic variable (a fancy name for a random variable), we get $\frac{dS_t}{S_t} = \mu dt + \sigma dW$. To find how V, the payoff of a stock option, changes with respect to S and t, we can substitute Ito's lemma to get:

$$dV = \left(\mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt + \sigma S \frac{\partial V}{\partial S} dW \tag{39}$$

An extra component we need to finish the derivation is the delta-hedge portfolio, a type of financial portfolio that remains financially stable over long periods of time and hence has a delta, or change, of zero. The portfolio hedges risk, which gives its name. The formula for the value a delta-hedge portfolio with one risky and one risk-free asset is:

$$\Pi = -V + \frac{\partial V}{\partial S} S \tag{40}$$

Over the period of time $[t, \Delta t]$, the total change in the portfolio's value is:

$$\Delta\Pi = -\Delta V + \frac{\partial V}{\partial S}\Delta S \tag{41}$$

If we discretize both equations for $\frac{dS}{S}$ and dV by changing differentials, which are infinitesimally small variations, for deltas, finite minute variations, we get:

$$\Delta S = \mu S \Delta t + \sigma S \Delta W \tag{42}$$

$$\Delta V = (\mu S \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2})\Delta t + \sigma S \frac{\partial V}{\partial S}\Delta W \tag{43}$$

If we now appropriately substitute these into the formula for the delta-hedge portfolio, we get:

$$\Delta\Pi = \left(-\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\Delta t \tag{44}$$

Because the ΔW term has disappeared from the equation, the portfolio lacks randomness, and is now effectively riskless; because of this, the rate of return of the formula can be used to model any risk-free asset. If the rate of return of the risk-free asset is r, then over the same time period we have:

$$r\Pi\Delta t = \Delta\Pi \tag{45}$$

If we substitute the original formula for the delta-hedge portfolio for the right side of the equation, we get:

$$\left(-\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)\Delta t = r\left(-V + S\frac{\partial V}{\partial S}\right)\Delta t \tag{46}$$

Dividing both sides by t, distributing the r on the right side, and moving the right side to the left, we get the Black-Scholes PDE.

Black-Scholes Formula

The Black-Scholes Formula is a solution of the Black-Scholes Partial Differential Equation that relies on the following boundary conditions:

$$C_{E,T} = max(0, S_T - X) \tag{47}$$

$$P_{E,T} = max(0, X - S_T) \tag{48}$$

where T is the time at which C, a European call option, and P, a European put option, expire. These two formulas provide the value of each of the options when they expire. With these boundary conditions, the Black-Scholes PDE solves to the Black-Scholes Formula:

$$C_E(S,t) = N(d_1)S - N(d_2)Xe^{-rT} \tag{49}$$

where S is the price of the security, T is the date of expiration, t is the current date, X is the exercise price, r is the rate of return for the risk-free asset, and σ is the volatility. The d_1 and d_2 are, respectfully, defined as:

$$d_1 = \frac{\ln\frac{S}{X} + (r + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \tag{50}$$

$$d_1 = \frac{\ln\frac{S}{X} + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \tag{51}$$

Notice that the Black-Scholes Formula uses the value of the cumulative function of the standard normal distribution at the values d_1 and d_2, as our initial equations for Brownian motion did. As we've established, this implies that N will always fall in the range *[0, 1]*.

To program this formula, we'd start by creating and populating a NumPy array with the values of a stock option at various times. We can then find the returns on that stock by taking $\ln\frac{S_t}{S_{t-1}}$, which in code would be:

```
returns = numpy.log(stock[1:] / stock[:-1])
```

We can find the variance of these returns with $\sigma = \sqrt{\frac{1}{n} \cdot \Sigma(r - \bar{r})^2}$, or with the code:

```
variance = numpy.sqrt(numpy.nanvar(returns))
```

Finally, we can create a function that takes all the necessary variables and returns the predicted value of the option:

```
import scipy.stats.norm.cdf

def returns(stock):
  return numpy.log(stock[1:] / stock[:-1])

def variance(returns):
  return numpy.sqrt(numpy.nanvar(returns))

def sim(stock, T, t, X, r):
  var = variance(stock)
  S = stock[-1]
  denominator = var * numpy.sqrt(T - t))
  d1_numerator = (numpy.log(S/X) + (r + (var ** 2 / 2))* (T - t))
  d1 = d1_numerator / denominator
  d2_numerator = (numpy.log(S/X) + (r - (var ** 2 / 2)) * (T - t))
  d2 = d2_numerator / denominator
  return S * cdf(d1) - X * numpy.exp(-r*T) * cdf(d2)
```

Conclusion

For their achievement of separating stock options from their underlying risks, Robert Merton and Myron Scholes were awarded the Noble Prize in Economics in 1973; Fischer Black was acknowledged as a contributor. Though alternatives exist, it is still a highly popular method for valuing company options, primarily because of the simplicity of its approach. Eighty percent of the fifty largest companies in the world use the Black-Scholes Model. However, a newer option becoming increasingly popular is the binomial "lattice" model, which uses a binomial tree to model the various paths an option's price may take. For more on that approach, you can read: https://en.wikipedia.org/wiki/Lattice_model_(finance)

References

1. "Approximations of π." Wikipedia, Wikimedia Foundation, 29 Jan. 2021, en.wikipedia.org/wiki/Approximations_of_%CF%80

2. "Black–Scholes Equation." Wikipedia, Wikimedia Foundation, 26 Dec. 2020, en.wikipedia.org/wiki/Black%E2%80%93Scholes_equation#Derivation_of_the_Black%E2%80%93Scholes_PDE.

3. “Google Colaboratory.” Google, Google, colab.research.google.com/drive/1UK_uuLmVqo3cB-KbrEcm_gGswYwV-qiS#scrollTo=teys8zf08_r-.

4. “Itô’s Lemma.” Wikipedia, Wikimedia Foundation, 19 Nov. 2020, en.wikipedia.org/wiki/It%C3%B4%27s_lemma#Geometric_Brownian_motion.

5. “Markov Chain.” Wikipedia, Wikimedia Foundation, 26 Jan. 2021, en.wikipedia.org/wiki/Markov_chain.

6. Milton, Adam. “What Are Call and Put Options?” The Balance, www.thebalance.com/call-and-put-options-definitions-and-examples-1031124.

7. “Probability Distribution.” Wikipedia, Wikimedia Foundation, 10 Jan. 2021, en.wikipedia.org/wiki/Probability_distribution.

8. Probability. 26 Jan. 2021, en.wikipedia.org/wiki/Probability.

9. “Random Variable.” Wikipedia, Wikimedia Foundation, 27 Dec. 2020, en.wikipedia.org/wiki/Random_variable.

10. “Stochastic Process.” Wikipedia, Wikimedia Foundation, 25 Jan. 2021, en.wikipedia.org/wiki/Stochastic_process#:~:text=A%20stochastic%20or%20random%20process%20can%20be%20defined%20as%20a,an%20element%20in%20the%20set.

11. Veisdal, Jørgen. “The Black-Scholes Formula, Explained.” Medium, Cantor’s Paradise, 4 July 2020,
medium.com/cantors-paradise/the-black-scholes-formula-explained-9e05b7865d8a.