Black Scholes using the Central Limit Theorem

Rajeshwari Majumdar, Phanuel Mariano, Lowen Peng, and Anthony Sisti
University of Connecticut

Abstract

The standard method of deriving the Black-Scholes European call option pricing formula involves stochastic differential equations. In this project we provide an alternate derivation using the Lindeberg-Feller central limit theorem under some technical assumptions. This method allows us obtain the Black-Scholes formula using undergraduate probability. Theoretical results are supplemented with market simulations.

Introduction

A **financial option** is a contract that gives the option holder the right to buy or sell an asset for a certain price, called the **strike price**. The specific type of option we studied was the **European call option**. It entitles the holder to purchase a unit of the underlying asset for a certain strike price at a pre-determined expiration time. The Black-Scholes model was proposed by Fischer Black and Myron Scholes in their 1973 paper entitled "The Pricing of Options and Corporate Liabilities," in which they derived a formula for the value of a "European-style" option in terms of the price of the stock. We consider an alternative approach to the derivation of the Black-Scholes European call option pricing formula using the central limit theorem. We utilize the **Lindeberg-Feller Central Limit Theorem** which states:

Theorem 1. *Suppose for each n and $i = 1, \dots n$, X_{ni} are independent and have mean 0. Let $S_n = \sum_{i=1}^{n} X_{ni}$. Suppose that $\sum_{i=1}^{n} \mathbb{E}[X_{ni}^2] \to \sigma^2$ for $0 < \sigma^2 < \infty$. Then, the following two conditions are equivalent:*

1. *S_n converges weakly to a normal random variable with mean 0 and variance σ^2, and the triangular array $\{X_{nj}\}$ satisfies the condition that*

$$\lim_{n\to 0} \max_{i} \mathbb{E}\left[X_{ni}^2\right] = 0.$$

2. *(Lindeberg Condition) For all $\epsilon > 0$,*

$$\sum_{i=1}^{n} \mathbb{E}\left[X_{ni}^2; |X_{ni}| > \epsilon\right] \to 0.$$

The Black Scholes Formula

The **Black-Scholes European option pricing model** is written,

$$C = X_0 N(d_+) - K e^{-rt} N(d_-) \tag{1}$$

where C is the discounted price of the call option, X_0 is the initial value of the underlying security, K is the strike price, t is the expiration date, N is the standard Gaussian CDF, and

$$d_{\pm} = \frac{1}{\sigma} \log\left[e^{rt} X_0 / K\right] \pm \frac{1}{2}\sigma.$$

Example 1. *Consider the purchase of a European call option on a stock with a present value of 50 Euros and a strike price of 52 Euros under the following conditions: r = 4%, t = 1 (year), σ = 15%. To calculate the price of this option we use Equation 1. To that end, we first find d_+ and d_-.*

$$d_+ = \frac{\log\left[e^{0.04(1)} 50/52\right]}{0.15} + \frac{1}{2}(0.15) = 0.0802 \quad \textit{and} \quad d_- = \frac{\log\left[e^{0.04(1)} 50/52\right]}{0.15} - \frac{1}{2}(0.15) = -.0698.$$

We then have

$$\begin{aligned} C_0 &= 50N(0.0802) - 52e^{-(0.04)(1)} N(-0.0698) \\ &= 50(.532) - 52(0.96)(0.472) \\ &= 3.04. \end{aligned}$$

Thus, from Black-Scholes model, the price of this call option would be 3.04 Euros.

Pricing the European Call Option

We can derive the Black-Scholes formula on the premise that our prices follow a log-normal distribution.

Lemma 1. *Assume there are no opportunities for arbitrage and the risk-free interest is r. Given a European call option with expiration t and strike K, let X_t be the time-t price of the underlying security, where $X_t = X_0 e^{Y_t}$ for some $Y_t \sim \mathcal{N}(\mu_{Y_t}, \sigma^2_{Y_t})$. Then the discounted price of the call option, denoted C, is given by*

$$C = e^{-rt} \mathbb{E}_{Q_t}\left[\max(X_t - K, 0)\right] = X_0 N(d_+) - K e^{-rt} N(d_-)$$

Proof. The price of a call option should be its discounted expected profit, that is,

$$C = e^{-rt} \mathbb{E}_{Q_t}\left[\max(X_t - K, 0)\right];$$

otherwise, there would be an opportunity for arbitrage, violating our initial assumption. The natural probability measure to study the price of a security is the time-t cash pricing measure Q_t. Expected profit is thus a straightforward computation:

$$\begin{aligned} \mathbb{E}_{Q_t}[\max(X_t - K, 0)] &= \int_{\Omega} \max(X_t - K, 0)\, \mathrm{d}Q_t \\ &= \int_{\{X_t \geq K\}} X_t - K\, \mathrm{d}Q_t \\ \implies \mathbb{E}_{Q_t}[\max(X_t - K, 0)] &= e^{rt} X_0 N(d_+) - K N(d_-), \end{aligned} \tag{2}$$

and by discounting this formula, we obtain $E_{Q_t}\left[\max(X_t - K, 0)\right] = X_0 N(d_+) - K e^{-rt} N(d_-)$. □

Log-Normality of Prices

We use the Lindeberg-Feller Central Limit Theorem to prove the premise that our prices follow a log-normal distribution under the following assumptions.

Assumption 1. *For each t, the random variable $Y_t = \log X_t$ has mean 0 and finite variance.*

Assumption 2. *The differences $Y_t - Y_s$ are independent for disjoint intervals $[s, t]$; for intervals of equal length, they are i.i.d.*

Assumption 3. *For every $\epsilon > 0$, $n\mathbb{E}\left[\left(Y_{t/n} - Y_0\right)^2; \left|Y_{t/n} - Y_0\right| > \epsilon\right] \to 0$.*

Theorem 2. *Under Assumptions 1, 2, and 3, for every $t > 0$, $Y_t = \log X_t$ is a normal random variable with variance $\sigma^2 t$.*

In addition to the Lindeberg-Feller central limit theorem, the proof of Theorem 2 makes use of the following lemma.

Lemma 2. *Suppose $f : [0, \infty) \to [0, \infty)$ satisfies $f(x + y) = f(x) + f(y)$. There exists a constant C such that $f(x) = Cx$ for all $x \geq 0$.*

Proof. (Theorem 2) We first show that

$$\mathrm{Var}\left[Y_t\right] = \sigma^2 t. \tag{3}$$

by Assumption 2, $\mathrm{Var}\left[Y_{t+s} - Y_0\right] = \mathrm{Var}\left[Y_s - Y_0\right] + \mathrm{Var}\left[Y_t - Y_0\right]$. With $f(u) = \mathrm{Var}\left[Y_u - Y_0\right]$, since f is non-negative, by Lemma 2 , $f(t) = tf(1)$, where $f(1) = \mathrm{Var}\left[Y_1 - Y_0\right] = \sigma^2$, thus establishing Equation 3. Now with $X_{ni} = Y_{ti/n} - Y_{t(i-1)/n}$ we obtain, by telescopic cancellation,

$$Y_t - Y_0 = \sum_{i=1}^{n} X_{ni}, \tag{4}$$

where the dependence of X_{ni} on t is supressed for notational convenience. Clearly, X_{ni} has mean 0. By the assumption of stationarity, X_{ni} is distributionally equal to $Y_{t/n} - Y_0$. Consequently, by Equation 3,

$$\mathbb{E}\left[X_{ni}^2\right] = \sigma^2 \frac{t}{n}, \tag{5}$$

implying $\sum_{i=1}^{n} \mathbb{E}\left[X_{ni}^2\right] = \sigma^2 t$.

Finally, by the consequence of the assumption of stationarity noted above,

$$\sum_{i=1}^{n} \mathbb{E}\left[X_{ni}^2; |X_{ni}| > \epsilon\right] = n\mathbb{E}\left[\left(Y_{t/n} - Y_0\right)^2; \left|Y_{t/n} - Y_0\right| > \epsilon\right],$$

whence the Lindeberg condition follows from Assumption 3. □

Simulations

The Black-Scholes stochastic differential equation is written $\mathrm{d}X_t = \mu X_t \mathrm{d}t + \sigma X_t \mathrm{d}W_t$ with the unique solution being a geometric Brownian motion

$$X_t = X_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right) \tag{6}$$

We might expect, then, that the approximate process,

$$S_{t+1} = S_0 \prod_{j=0}^{t} (1 + \mu + \sigma\epsilon_j) \tag{7}$$

for an appropriate choice of i.i.d. random variables ϵ_t, serves as a discrete simulation of geometric Brownian motion. By Equation 6, we expect the Gaussian associated to S_t to have parameters approximately $(\mu - \sigma^2/2)t$ and $\sigma^2 t$ respectively. We can compute the theoretical mean and variance of the random variable Y_t using discrete intervals.

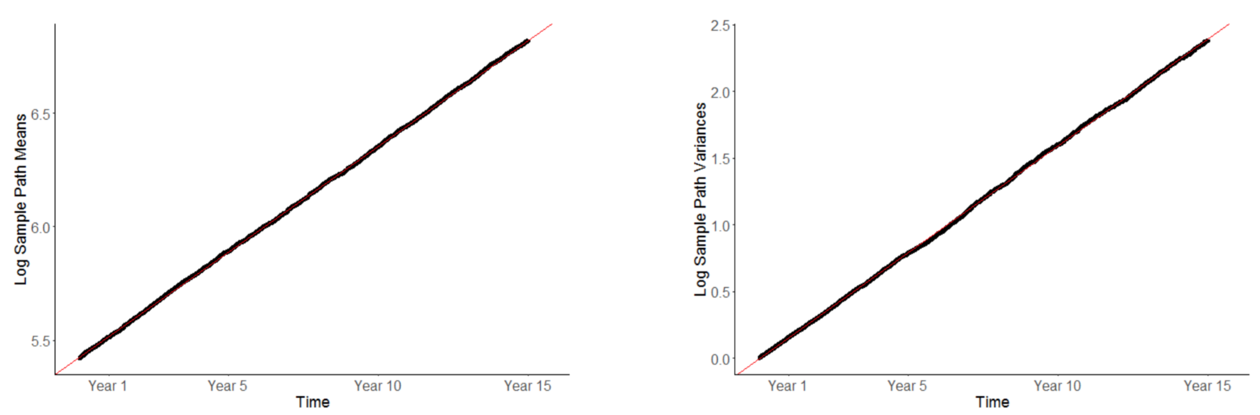
$$\mathbb{E}[Y_t] = \log S_0 + t(-r + \mathbb{E}[\log(1 + \mu + \sigma\varepsilon_i)]). \tag{8}$$

$$\mathrm{Var}(Y_t) = \mathbb{E}[(Y_t)^2] - \mathbb{E}[Y_t]^2 = \left(\mathbb{E}[\log^2(1 + \mu + \sigma\varepsilon_i)] - \left(\mathbb{E}[\log(1 + \mu + \sigma\varepsilon_i)]\right)^2\right)t. \tag{9}$$

Using parameters obtained from real market values, we compute the following theoretical means and variances for $\varepsilon_i \sim$ Rademacher before running the actual market simulations.

$\varepsilon_i \sim$ Rademacher	$\mathbb{E}[Y_t] = 6.8326$	Slope $= 0.0002580$
	$\mathrm{Var}(Y_t) = 2.370$	Slope $= 0.0004329$

Our objective is to verify the following: The mean of our log-sample paths at time t approximate to $(\mu - \sigma^2/2)t$, the variance of our log-sample paths at time t approximate to $\sigma^2 t$ and at each time t, X_t is log-normal. From Figures 1(A) and 1(B), we note that the slopes given by our toy model, both simulated and theoretical, closely match the drift and volatility of the geometric Brownian motion.



(A) Log Sample Path Means (B) Log Sample Path Variances

Figure 1: Simulated slopes for $\varepsilon_i \sim$ Rademacher.

Figure 2 shows the log prices at the end of the 15th year. The data is normally distributed.

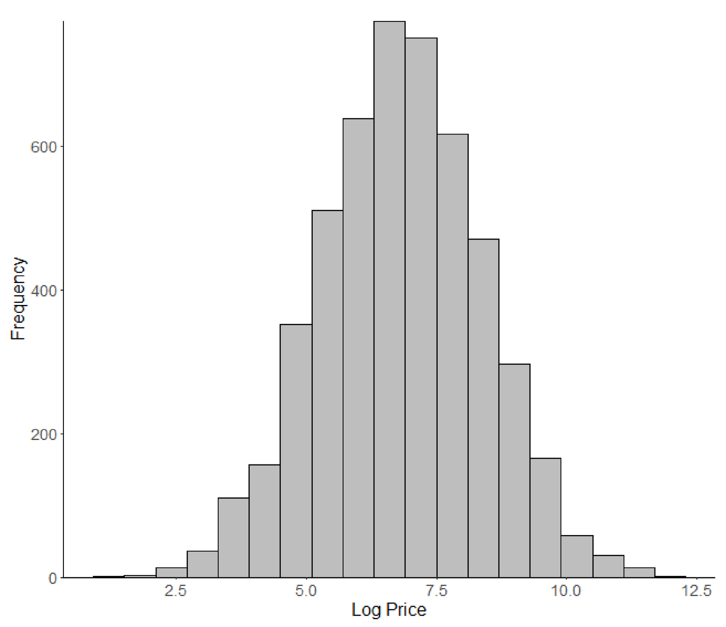


Figure 2: Histogram of log prices at Year 15 with $\varepsilon_i \sim$ Rademacher.