1 Introduction

These notes are intended for students taking a course in Financial Engineering or Advanced Engineering Economy, as a supplement to lectures. They are not intended as a substitute for an in-depth study of the theory of stochastic processes applied to finance.

Section 2 introduces stochastic processes. The properties of stationary and independent increments are defined, and illustrated with the example of the Poisson Process. Section 3 introduces Brownian Motion, a fundamental building block of financial engineering and mathematics. Basic properties are covered as well as a few notable results. While Brownian Motion itself cannot be used to model the stochastic evolution of stock prices, functions of Brownian Motion can. Brownian Motion is inexorably linked to the normal distribution. In these notes, the notation $X \sim N(\mu, \sigma^2)$ means the random variable $X$ has the normal distribution with mean $\mu$ and variance $\sigma^2$. Its density function is given by

$$
\phi(y) := \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}
$$

(1)

and its cumulative distribution function is given by

$$
\Phi(x) := P(X \leq x) = \int_{-\infty}^{x} \phi(y)dy.
$$

(2)

Section 4 introduces the Random Walk, a simple model of a stochastic process. It is shown that the Random Walk converges in a certain sense to Brownian Motion. It turns out that event probabilities associated with Brownian Motion can be determined by analyzing a “limiting” Random Walk, and vice-versa. Section 5 introduces Geometric Brownian Motion, which is the most ubiquitous model of stochastic evolution of stock prices. Basic properties are established. Section 6 revisits the Binomial Lattice, and shows that the choices for the parameters we have been using will approximate Geometric Brownian Motion when the number of periods is sufficiently large and the period length tends to zero. Section 7 provides a number of numerical examples of the results obtained up to this point.

Section 8 derives the famous Black-Scholes Call Option formula. We know from our earlier work that the value of the option may be obtained as a discounted expectation. Since the distribution of the stock price at maturity (under the risk-neutral world) is determined in Section 6, we are in position to perform the expectation calculation. Section 9 illustrates the Black-Scholes formula with a number of examples. Section 10 introduces one-dimensional Ito Processes, stochastic processes that encompass Geometric Brownian Motion processes. Essentially, Ito Processes may be represented as a stochastic integral involving functions of Brownian
Motion. We will (heuristically) derive the famous Ito-Doeblin’s Formula, otherwise known as Ito-Doeblin’s Lemma. Examples of the use of Ito-Doeblin’s Formula are provided. Section 11 revisits the Black-Scholes formula by first deriving the famous Black-Scholes equation using Ito-Doeblin’s Lemma. The Black-Scholes formula solves a partial differential equation, which involve the sensitivities (or the “Greeks”) of the call option value to the stock price and the time to maturity. Section 12 closes with an application of continuous-time stochastic optimal control to a real options valuation problem.

2 Stochastic Processes

Stochastic processes are used to model a variety of physical phenomena; for example, a stock price over time or a project value over time. Each outcome \( \omega \) of a continuous-time stochastic process \( \{X(t) : t \geq 0\} \) may be identified with a function \( X(\omega, \cdot) \) defined on \([0, \infty)\), whose graph 
\[
\{(t, X(\omega, t)) : t \geq 0\}
\]
is called a sample path. The index \( t \) is commonly identified with time. To ease notational burdens the symbol \( \omega \) is often suppressed and one writes \( X(t) \) or \( X_t \) for the random value of the process at time \( t \). It is important to always keep in mind that \( X_t \) is a random variable.

Stochastic processes differ in the properties they possess. Often, properties are imposed for natural reasons or simply for analytical convenience. Two such convenient properties are stationary and independent increments.

**Definition 1** A stochastic process \( \{X(t) : t \geq 0\} \) possesses independent increments if the random variables
\[
X(t_1) - X(t_0), \ X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})
\]
are independent random variables for any \( 0 \leq t_0 < t_1 < \cdots < t_n < \infty \), and stationary increments if the distribution of \( X(t + s) - X(t) \) depends only on difference \((t + s) - t = s\).

**Example 1** A stochastic process that possesses stationary and independent increments is the counting process \( N(t) \) associated with a Poisson Process. A counting process \( N(t) \) records the number of arrivals that have entered a system by time \( t \). Each sample path is a nondecreasing step function that always starts at zero, i.e, \( N(0) = 0 \), and its discontinuities define the arrival times of the events. At most one arrival may enter the system at any point in time and there are always a finite number of arrivals in any finite interval of time (almost surely). \(^2\) Thus, each sample path does not have too many “jumps” (discontinuity points). For the Poisson (counting) process

\[^2\]The expression “almost surely” means “except for a small number of sample paths whose total probability of ever occurring is zero”.

2
Process, the (discrete) random variable \( N(t+s) - N(t) \) is Poisson with parameter \( \lambda s \). Recall that a discrete random variable \( X \) is Poisson with parameter \( \lambda \) if it only takes on non-negative integer values and its probability mass function is given by

\[
P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}
\]

for each non-negative integer \( k \). The mean and variance of \( X \) equal \( \lambda \). The parameter \( \lambda \) defines the expected rate of arrivals per unit time.

### 3 Brownian Motion

Another example of a stochastic process that possesses stationary and independent increments is Brownian Motion, one of the most famous and important stochastic processes of probability and finance.

#### 3.1 Definition

**Definition 2** Standard Brownian Motion is a stochastic process \( \{B(t) : t \geq 0\} \) with the following properties:

(i) it always starts at 0, i.e., \( B(0) = 0 \);

(ii) it possesses stationary and independent increments;

(iii) each sample path \( B(\omega, \cdot) \) is continuous (almost surely); and

(iv) \( B(t+s) - B(t) \sim N(0, s) \).

A stochastic process \( \{X(t) : t \geq 0\} \) is called \( (\mu, \sigma) \) Brownian Motion if it may be represented as

\[
X(t) = X(0) + \mu t + \sigma B(t),
\]

where \( B \) is standard Brownian Motion and \( X(0) \) is independent of \( B \). By property (iv)

\[
X(t+s) - X(t) = \mu s + \sigma B(s) \sim N(\mu s, \sigma^2 s).
\]

The parameter \( \mu \) is called the drift and the parameter \( \sigma^2 \) is called the variance of the process \( X \). The expression Brownian Motion will be used to describe all such processes \( X \); the symbol \( B \) will be reserved for standard Brownian Motion in which \( \mu = 0 \) and \( \sigma^2 = 1 \).
3.2 Interpretation

It was the French mathematician Bachelier who used Brownian Motion (in his doctoral dissertation, 1900) to describe the price movements of stocks and commodities. Brownian Motion is an example of a continuous-time, continuous state space, Markov process. Each increment $X(t + s) - X(t)$ is independent of the history of the process up to time $t$. Alternatively, if we know $X(t) = x$, then any knowledge of the values of $X(\tau)$ for $\tau < t$ will have no effect on the probability law governing $X(t + s) - X(t)$.

There are two serious problems with using Brownian Motion to model stock price movements:

- First, since the price of the stock has a normal distribution it could theoretically become negative, which can never happen in practice due to limited liability.
- Second, the price difference over an interval of fixed length has the same distribution regardless of the price at the beginning of the period. For example, the model assumes the probability the stock price will increase by 10 over the next month will be the same whether the initial price is 100 or 20. In the first case, the increase is 10% whereas in the second case it is 50%. It is more reasonable to assume the probability of a fixed percentage increase is the same.

From the financial modelling perspective, one may ask: Why study Brownian Motion? It turns out that modelling the percentage change in the stock price over an interval as a normal random variable underlies the most ubiquitous basic model of stock price dynamics, which we shall cover shortly. This particular model will be seen as a simple “transform” of Brownian Motion, and so its properties are fundamentally tied to the properties of Brownian Motion.

Remark 1 It is a famous theorem of probability that the Brownian Motion process exists. In a subsequent section we shall heuristically show that Brownian Motion may be seen as an appropriate limit of a particular type of discrete-time stochastic process called a Random Walk.

Remark 2 Another important theorem states that if $X$ is a stochastic process with stationary and independent increments whose sample paths are continuous (almost surely), then $X$ must be Brownian Motion. This means that Brownian Motion could be defined by stationary and independent increments and sample path continuity alone, with the normal distribution property of each increment being a consequence of these assumptions. Note that the Poisson Process, which has stationary and independent increments, is not Brownian Motion since its sample paths are discontinuous.
3.3 A few useful results

The structure of Brownian Motion permits closed-form expressions for a variety of results, though some require advanced techniques to establish. We note a few of these results now.

3.3.1 Covariance

The covariance of the random variables $B_{t+s}$ and $B_t$ is determined as follows:

$$COV(B_{t+s}, B_t) = E[B_{t+s}B_t] - E[B_{t+s}]E[B_t]$$

$$= E[B_{t+s}B_t]$$

$$= E[(B_{t+s} - B_t + B_t)B_t]$$

$$= E[(B_{t+s} - B_t)B_t] + E[B_t^2]$$

$$= E[B_{t+s} - B_t]E[B_t] + E[B_t^2]$$

$$= t,$$

where the second line follows since the means are zero; the third line applies a simple but very useful identity; the fourth line follows by the linearity of the expectation operator $E$; the fifth line follows by the independent increments of Brownian Motion; and the last line follows since each increment has zero mean and $E[B_t^2] = Var(B_t)$. In other words, for any times $\tau$ and $\tau'$

$$COV(B_\tau, B_{\tau'}) = E[B_\tau B_{\tau'}] = \min\{\tau, \tau'\}.$$  \hfill (12)

Formula (12) and the algebraic identity are used frequently when analyzing properties of Brownian Motion, in particular stochastic integrals.

3.3.2 Differentiability

Fix a time $t$ and consider the random ratio

$$\frac{B_{t+\epsilon} - B_t}{\epsilon}.$$  \hfill (13)

Its mean

$$E\left[\frac{B_{t+\epsilon} - B_t}{\epsilon}\right] = 0,$$  \hfill (14)

since $B_{t+\epsilon} - B_t \sim N(0, \epsilon)$. Its variance

$$Var[B_{t+\epsilon} - B_t]/\epsilon^2 = 1/\epsilon,$$  \hfill (15)

which goes to infinity as $\epsilon \to 0$. Since the limit of this ratio as $\epsilon \to 0$ would ideally represent the derivative of $B$ at $t$, it would appear, at least heuristically, that the limit does not exist.

A fundamental theorem of Brownian Motion is that each of its sample paths is nowhere differentiable (almost surely)! Essentially, each sample path “wiggles too much”.

5
3.3.3 Variation

Variation is a measure of how much a function “wiggles”. With respect to one measure of variation the sample paths of Brownian Motion wiggle far too much; with respect to a different measure of variation, the sample paths of Brownian Motion are just fine. First, some definitions.

Let \( P_t = \{ t_i \}_{i=0}^n \) denote a finite collection of time points for which \( 0 = t_0 < t_1 < \cdots < t_n = t \). Given a function \( f : [0, \infty) \to \mathbb{R} \), for each \( P_t \) define

\[
v(f; P_t) := \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|
\]

and define the total variation of \( f \) over \([0, t]\) as

\[
v_t(f) := \sup_{P_t} v(f; P_t)
\]

The function \( f(\cdot) \) is said to be of bounded variation if \( v_t(f) < \infty \) for all \( t > 0 \). Each nondecreasing function \( f \) has bounded variation, since \( v_t(f) = f(t) - f(0) \). A fundamental property is that a continuous function has bounded variation if and only if it may be represented as the difference of two nondecreasing functions. This fundamental property makes it possible to define the Riemann-Stieltjes Integral \( \int h \, df \) as the limit, in a certain sense, of \( \sum h(t_i)[f(t_i) - f(t_{i-1})] \).

If we replace the function \( f(\cdot) \) with \( B(\omega, \cdot) \), it turns out that almost all sample paths of Brownian Motion have infinite variation. Heuristically, at least, this is not too surprising, since the sample paths are infinitely jagged. Define the quadratic variation of \( f \) over \([0, t]\) as

\[
q_t(f) := \sup_{P_t} q(f; P_t),
\]

where

\[
q(f; P_t) := \sum_{i=1}^{n} (f(t_i) - f(t_{i-1}))^2.
\]

A fundamental property of Brownian Motion is that almost all of its sample paths have the same quadratic variation given by \( \sigma^2 t \). This property contains the essence of the famous Itô-Doeblin’s formula, which is the key tool for analysis of stochastic processes related to Brownian Motion.

**Remark 3** A heuristic explanation begins by noting that \( E[q(X; P_t)] = \sigma^2 t + \mu^2 \sum (t_i - t_{i-1})^2 \). When \( P_t \) is chosen so that \( t_i := (i/n)t \), then \( E[q(X; P_t)] \to \sigma^2 t \) as \( n \to \infty \). A bit more work shows that \( \text{Var}[q(X; P_t)] \to 0 \) as \( n \to \infty \).
3.3.4 Hitting time distribution

Let \( X \) represent Brownian Motion such that \( X(0) = 0 \). For each \( t \) let \( M_t \) denote the maximum value of \( X \) on the interval \([0, t]\). Its distribution is given by

\[
P\{M_t < y\} = \Phi\left(\frac{y - \mu t}{\sigma \sqrt{t}}\right) - e^{2\mu y/\sigma^2} \Phi\left(\frac{y - \mu t}{\sigma \sqrt{t}}\right).
\]  
(20)

For \( y > 0 \) let \( T(y) \) denote the first time \( t \) at which \( X_t = y \). It is called a one-sided hitting or passage time. It should be clear that

\[
P(T(y) > t) = P(M_t < y),
\]  
(21)

from which the one-sided hitting time distribution may be calculated via (20). The results when \( X(0) = x \) are obtained by considering the process \( X(t) - x \). (See Section 7 for an example.)

3.3.5 Ruin probabilities

Let \( X \) represent Brownian Motion such that the drift \( \mu \neq 0 \) and \( X(0) = x \). Imagine an individual who purchases a unit of a good whose price is governed by the \( X \) process. The individual wishes to hold the good until its price reaches \( b > x \). To limit his losses, the individual will sell the good should its price decline to \( a < x \).

Let \( T_{ab} \) denote the first time the process \( X \) reaches either \( a \) or \( b \). It records the time when the individual will sell the good and either reap the gain \( b - x \) if \( X(T_{ab}) = b \) or take a loss of \( x - a \) if \( X(T_{ab}) = a \). The chance the individual will go home a winner is

\[
P\{X(T_{ab}) = b \mid X(0) = x\} = \frac{e^{-2\mu x/\sigma^2} - e^{-2\mu a/\sigma^2}}{e^{-2\mu b/\sigma^2} - e^{-2\mu a/\sigma^2}},
\]  
(22)

and the chance the individual will go home “ruined” is of course \( 1 - P\{X(T_{ab}) = b \mid X(0) = x\} \).

The formula (22) is valid so long as \( a < x < b \), even if these numbers are negative.

Formula (22) is not valid when \( \mu = 0 \); however, when \( \mu \approx 0 \), we may substitute \( e^z \approx 1 + z \) in (22) to obtain that

\[
P\{X(T_{ab}) = b \mid X(0) = x\} = \frac{x - a}{b - a} \quad \text{when } \mu = 0.
\]  
(23)

(See Section 7 for an example.)

3.3.6 Versions of standard Brownian Motion

It is possible for a function of standard Brownian Motion to be another version of standard Brownian Motion. Examples include:
• $B_1(t) := cB(t/c^2)$, for fixed $c > 0$.
• $B_2(t) := tB(1/t)$ if $t > 0$ and 0 otherwise.
• $B_3(t) := B(t + h) - B(h)$, for fixed $h > 0$.

For each of these examples, it should be clear that every increment $B_i(t + s) - B_i(t)$, $i = 1, 2, 3$, is normally distributed with zero mean, and that the increments over disjoint time intervals determine independent random variables. To complete the verification, one must show the sample paths are continuous, and that variance of each $E[B_i(t + s) - B_i(t)]^2$ is indeed $s$. The verification of the variance is left as an exercise.

4 Random Walks

A cornerstone of the theory of stochastic processes is the random walk. It provides a simple model of a wealth process of an individual who repeatedly tosses a possibly biased coin and places bets. The most common description of the price of a (non-dividend) paying stock will turn out to be a limit of a suitable, simple transformation of a random walk.

4.1 Definition

We imagine an individual who sequentially tosses a fair coin and who will receive one dollar if the coin comes up heads and will lose one dollar if the coin comes up tails. Let \{$X_i : 1 \leq i < \infty$\} denote an infinite sequence of independent and identically distributed (i.i.d.) random variables with common distribution given by

$$P(X_i = 1) = P(X_i = -1) = 1/2.$$  \hfill (24)

The value

$$S_n := \sum_{i=1}^{n} X_i$$  \hfill (25)

is the individual’s wealth after $n$ tosses of the coin. The process \{$S_n : 0 \leq n < \infty$\} is called a simple random walk. The expected winnings on each toss is, of course, zero, and so

$$E[S_n] = 0.$$  \hfill (26)

Moreover,

$$Var[S_n] = n,$$  \hfill (27)

since the $X_i$ are independent and the variance of each $X_i$ is one. (The second moment is always one and the mean is zero.)
Next, we allow the coin to be biased so that the probability the coin will come up heads is given by the parameter $p \in (0, 1)$. In this setting, the $X_i$ have the same distribution given by

$$P(X_i = 1) = p, \quad P(X_i = -1) = 1 - p.$$  \hfill (28)

Of course, the expected winnings after $n$ tosses is no longer zero, and its variance changes, too. In particular,

$$E[S_n] = nE[X_i] = np(1) + (1-p)(-1) = n(2p-1)$$  \hfill (29)

$$Var[S_n] = nVar[X_i] = n[1-(2p-1)^2] = 4np(1-p).$$  \hfill (30)

Finally, we not only allow the coin to be biased, but we now scale the amount won or loss by the parameter $\sigma > 0$. In this setting, the $X_i$ now have common distribution given by

$$P(X_i = \sigma) = p, \quad P(X_i = -\sigma) = 1 - p,$$  \hfill (35)

and

$$E[S_n] = nE[X_i] = n[p(\sigma) + (1-p)(-\sigma)] = n\sigma(2p-1)$$  \hfill (36)

$$Var[S_n] = nVar[X_i] = n[\sigma^2 - \sigma(2p-1)^2] = 4n\sigma^2p(1-p).$$  \hfill (37)

**Remark 4** Of course, when $\sigma = 1$ the mean (38) and variance (41) coincide with (31) and (34), respectively. If in addition $p = 1/2$, then (38) and (41) coincide with the mean and variance of a simple random walk.

**Definition 3** A **random walk** shall refer to the positions $S_n$ generated by the general process defined by (35). A **biased random walk** shall refer to the special case when $\sigma = 1$, and a **simple random walk** shall refer to the special case when $\sigma = 1$ and $p = 1/2$.

### 4.2 Standardization

We now incorporate a time dimension, as follows. Suppose each $\Delta t > 0$ units of time the individual tosses the coin. The value $S_n$ records the individual’s wealth after $T_n = n\Delta t$ units
of time. If we instead set a time $T > 0$ and insist that the number of tosses, $n$, and the time between each toss, $\Delta t$, are such that the game is always over by time $T$, then $n$ and $\Delta t$ determine one another via the equation

$$T = n\Delta t.$$  \tag{42}

(It is understood that $\Delta t$ is always chosen so that $n$ is an integer.) For example, if $T = 1$ year, we could toss the coin quarterly, monthly, daily, hourly or even by the minute, which would result in 4, 12, 365, 8,760 and 525,600 tosses, respectively.

There are two parameters, $p$ and $\sigma$, required to specify a random walk. If we choose $p$ and $\sigma$ in (35) to be independent of $n$, then for a biased coin (38) and (41) show that

- $E[S_n]$ goes to plus or minus infinity (depending on whether $p > 1/2$ or $p < 1/2$) as $n$ tends to infinity, and
- the coefficient of variation $\sqrt{Var[S_n]} / E[S_n]$ converges to zero as $n$ tends to infinity.

Consequently, the value of wealth at time $T$ will be “unbounded” as $\Delta t \to 0$. Now keeping in mind identity (42), suppose instead we define

$$p_n := \frac{1}{2} \left( 1 + \frac{\nu}{\sigma} \sqrt{\Delta t} \right)$$  \tag{43}

$$\sigma_n := \sigma \sqrt{\Delta t},$$  \tag{44}

and set $p = p_n$ and $\sigma = \sigma_n$ into (35). Then

$$E[S_n] = n\sigma_n (2p_n - 1)$$  \tag{45}

$$= [n\sigma \sqrt{\Delta t}] \left( \frac{\nu}{\sigma} \sqrt{\Delta t} \right)$$  \tag{46}

$$= \nu T,$$  \tag{47}

$$Var[S_n] = 4n\sigma_n^2[p_n(1 - p_n)]$$  \tag{48}

$$= 4n\sigma^2 \Delta t \left[ \frac{1}{4} \left( 1 - \frac{\nu^2}{\sigma^2} \Delta t \right) \right]$$  \tag{49}

$$\to \sigma^2 T,$$  \tag{50}

then the limiting value of wealth at time $T$,

$$S_T := \lim_{n \to \infty} S_n$$  \tag{51}

will exist and have a non-trivial distribution. In particular, its mean is always $\nu T$ and its variance is $\sigma^2 T$. The parameter $\nu$ in (43) is an arbitrary finite number that does not depend on $n$, and it is understood that $\Delta t$ is sufficiently small (or $n$ is sufficiently large) to ensure that $p_n \in (0, 1)$. 

10
Remark 5 We may always express $p_n$ in the form $1/2 \cdot (1 + \alpha_n)$. If we insist that $\alpha_n \to \alpha \in (-1, 1)$, then an examination of (45) and (48) shows that both $\alpha_n$ and $\sigma_n$ must be proportional to $1/\sqrt{n}$ if the mean and variance of $S_n$ are to converge.

4.3 Limiting distribution

What is the distribution of $S_T$? First, we recall the well-known Central Limit Theorem (CLT). Let $Z$ denote the standard normal variable.

**Theorem 1 Central Limit Theorem (CLT).** Let $\{X_k\}$ be an infinite sequence of i.i.d. random variables with finite mean $a$ and finite variance $b^2$, let $Y_n := X_1 + \cdots + X_n$ denote the $n^{th}$ partial sum, and let $Z_n := (Y_n - na)/b\sqrt{n}$ denote the standardized sum, $Y_n$ minus its mean divided by its standard deviation. For every fixed real value $x$

$$\lim_{n \to \infty} P(Z_n < x) = P(Z < x) = \Phi(x).$$

Consider the standardized sum

$$Z_n := \frac{\sum_{i=1}^n X_i - \nu T}{\sigma \sqrt{T} \left[ \sqrt{1 - \frac{\nu^2}{b^2}(T/n)} \right]}$$

where the $X_i$ define a random walk. The numerator of (53) is a sum of random variables minus its mean and the denominator is the standard deviation of that sum. What happens to the standardized sum as $n$ tends to infinity? It looks, at first blush, that we may invoke the CLT. However, in our setting, the distribution of each $X_i$ depends on $n$. Fortunately, for the problem studied here, a general CLT due to Lindeberg-Feller ensures that, indeed, for every fixed real value $x$

$$\lim_{n \to \infty} P(Z_n < x) \to \Phi(x).$$

As a direct consequence, to any desired degree of accuracy, for each fixed real number $s$ we may pick $n$ sufficiently large to ensure that

$$P(S_n < s) = P\left( \sum_{i=1}^n X_i < s \right)$$

$$= P\left\{ \frac{\sum_{i=1}^n X_i - \nu T}{\sigma \sqrt{T}} < \frac{s - \nu T}{\sigma \sqrt{T}} \right\}$$

$$\approx \Phi\left( \frac{s - \nu T}{\sigma \sqrt{T}} \right).$$

In fact, to any desired degree of accuracy we may pick $n$ sufficiently large to ensure that for fixed real numbers $s_1, \ldots, s_m$,

$$P(S_n < s_j) \approx \Phi\left( \frac{s_j - \nu T}{\sigma \sqrt{T}} \right).$$
To all intents and purposes, the distribution of $S_n$ converges to $N(\nu T, \sigma^2 T)$, the distribution of the random variable we call $S_T$.

**Remark 6** Note how the reward/loss scale “directly” determines the variance of the limiting random variable $S_T$. Furthermore, given the reward/loss scale, if the form of $p_n$ follows (43), then each choice of $\nu/\sigma$ not only determines the likelihood of success on each coin toss, but it determines the mean of the limiting random variable $S_T$.

There is more we can say. Fix times

$$0 := t_0 < t_1 < t_2 < \cdots < t_K := T,$$

and for each $\Delta t$ define $n_k$ so that

$$n_k \Delta t = t_k, \ 0 \leq k \leq K.$$  \hspace{1cm} (60)

Fix an arbitrarily large integer $N$, pick $\Delta t$ sufficiently small so that

$$\min_{0 \leq k \leq K} (n_k - n_{k-1}) > N.$$  \hspace{1cm} (61)

Note that each $n_k > kN > N$ for $k \geq 1$. Define a random walk in which $p$ and $\sigma$ are set by the right-hand side of (43) and (44), respectively. (The parameters $p$ and $\sigma$ are now functions of $\Delta t$ and not $n$.) The calculations of (47) and (50) are still valid, from which we conclude that

$$E[S_{n_k}] = \nu t_k$$  \hspace{1cm} (62)

$$\text{Var}[S_{n_k}] \approx \sigma^2 t_k.$$  \hspace{1cm} (63)

Since the $X_i$ are i.i.d. random variables in this setting, and since the $n_k$ are sufficiently large, we may directly apply the well-known CLT to establish that

$$S_{t_k} = S_{n_k} \sim N(\nu t_k, \sigma^2 t_k), \ 1 \leq k \leq K.$$  \hspace{1cm} (64)

Moreover, since

$$S_{n_k} - S_{n_{k-1}} = \sum_{i=n_{k-1}}^{n_k-1} X_i,$$  \hspace{1cm} (65)

we have

$$E[S_{n_k} - S_{n_{k-1}}] = \nu(t_k - t_{k-1})$$  \hspace{1cm} (66)

$$\text{Var}[S_{n_k} - S_{n_{k-1}}] \approx \sigma^2(t_k - t_{k-1}).$$  \hspace{1cm} (67)

Since the $n_k - n_{k-1}$ are sufficiently large, direct application of the CLT also establishes that

$$S_{t_k} - S_{t_{k-1}} = S_{n_k} - S_{n_{k-1}} \sim N(\nu(t_k - t_{k-1}), \sigma^2(t_k - t_{k-1})), \ 1 \leq k \leq K.$$  \hspace{1cm} (68)
We now make the fundamental extrapolation known as Donsker’s theorem. In light of (64) and (68), if we let $\Delta t \to 0$ and $K \to \infty$ such that $\max_k (t_k - t_{k-1}) \to 0$, to the naked eye the random walk $S_n$ looks like a $(\nu, \sigma)$ Brownian Motion stochastic process!

The implication of this result can not be understated: the chances of a (reasonable) event occurring for a Brownian Motion process may be estimated by computing the chance the event would occur for a random walk with sufficiently many steps.

4.4 Functions of a random walk

We first consider a biased random walk (i.e. general $p$, $\sigma = 1$). Let the positive integer $L$ represent the maximum amount of money the individual is prepared (or can) lose, and let the integer $M$ represent the amount of money the individual wishes to win. We suppose the individual will continue to toss the coin until either he reaches his goal of $M$ or is forced to quit by losing $L$.

Let $\tau(M, L)$ denote the first index (integer) for which $S_{\tau(M, L)} = M$ or $-L$. If we identify each toss with a unit of time, then $\tau(M, L)$ would record the time at which the game ends. Note that by definition of $\tau(M, L)$, $S_{\tau(M, L)}$ must equal either $M$ or $-L$.

A fundamental probability of interest is the so-called ruin probability given by

$$\pi(M, L) := P\{S_{\tau(M, L)} \text{ reaches } M \text{ before } -L \}.$$  \hspace{1cm} (69)

Let $q := 1 - p$. It can be shown that

$$\pi(M, L) = \frac{(q/p)^L - 1}{(q/p)^{M+L} - 1},$$  \hspace{1cm} (70)

$$E[\tau(M, L)] = \frac{L}{q - p} - \frac{M + L}{q - p} \pi(M, L).$$  \hspace{1cm} (71)

The formulas do not apply for the simple random walk (since $p = q$). However, by setting $p = 1/2 + \epsilon$ and $1 - p = 1/2 - \epsilon$ into (70) and (71) and letting $\epsilon \to 0$, it may be shown that for the simple random walk

$$\pi(M, L) = \frac{L}{M + L},$$  \hspace{1cm} (72)

$$E[\tau(M, L)] = ML.$$  \hspace{1cm} (73)

**Example 2** For a simple random walk there would be a $2/3$ chance of winning 100 before losing 200. On the other hand, when $p = 0.49$,

$$\pi(M, L) = \frac{(0.51/0.49)^{200} - 1}{(0.51/0.49)^{300} - 1} = 0.0183,$$  \hspace{1cm} (74)
Table 1: Sample benchmarks

<table>
<thead>
<tr>
<th>Prob. $p$</th>
<th>0.5000</th>
<th>0.4950</th>
<th>0.4900</th>
<th>0.4800</th>
<th>0.4700</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prob. of winning</td>
<td>0.5000</td>
<td>0.1191</td>
<td>0.0179</td>
<td>0.0003</td>
<td>0.000006</td>
</tr>
<tr>
<td>Expected duration of game</td>
<td>10,000</td>
<td>7,616</td>
<td>4,820</td>
<td>2,498</td>
<td>1,667</td>
</tr>
</tbody>
</table>

a less than 2% chance of being successful! The cost of an even a small bias is surprisingly high. Table 1 shows the probability of winning 100 before losing 100 as the probability of winning varies.

**Remark 7** We analyzed ruin probabilities in our discussion about Brownian Motion. We have argued that the random walk converges in a certain sense to Brownian Motion. It can be shown that the probabilities shown in the Table can be estimated using those earlier formulas. (It will not be exact as the probabilities will be equal only in the limit.) It turns out that event probabilities concerning (continuous) functions of Brownian Motion may be found by taking the limit of the corresponding event probabilities of the corresponding random walk, and vice-versa. The choice as to which approach to take depends on which analytical techniques will be more suited for the task at hand. This fundamental result is knowns as Donsker’s Invariance Principle.

## 5 Geometric Brownian Motion

### 5.1 Definition

A stochastic process \( \{Y(t) : 0 \leq t < \infty\} \) is a *Geometric Brownian Motion* process if it may be represented in the form

\[
Y(t) = Y(0)e^{X(t)},
\]

where \( X \) is a \((\nu, \sigma)\) Brownian Motion Process and \( Y(0) \) is a positive constant independent of \( X \). Note that each sample path of the \( Y \) process is strictly positive. By taking the natural logarithm of both sides of (75),

\[
\ln \frac{Y(t)}{Y(0)} \sim N(\nu t, \sigma^2 t).
\]

A random variable whose natural logarithm has a normal distribution is called a *lognormal* random variable and its distribution is called a *lognormal* distribution. Thus, each \( Y(t) \) is *lognormally distributed.*
For a Geometric Brownian Motion process $Y$, it is the percentage change

$$\frac{Y(t + \Delta t) - Y(t)}{Y(t)} \approx \ln \frac{Y(t + \Delta t)}{Y(t)} \sim N(\nu \Delta t, \sigma^2 \Delta t),$$

(loosely speaking) that is normally distributed. As previously motivated, this property together with the stationary and independent increments are used to justify Geometric Brownian Motion as a reasonable first-cut model of the price process associated with a (non-dividend) paying stock.

5.2 Calculation of probabilities and moments

Probabilities of events on $Y$ are easily calculated via the normal distribution. For example, for $0 < a < b$, $\hat{a} := a/Y_0$, $\hat{b} := b/Y_0$,

$$P(a < Y_t < b) = P(\ln \hat{a} - \nu t < \ln Y_t/Y_0 < \ln \hat{b} - \nu t) \quad (78)$$

$$= \Phi \left( \frac{\ln \hat{b} - \nu t}{\sigma \sqrt{t}} \right) - \Phi \left( \frac{\ln \hat{a} - \nu t}{\sigma \sqrt{t}} \right). \quad (80)$$

Recall the moment generating function for a normal random variable $X \sim N(a, b^2)$ is

$$E(e^{sX}) = e^{sa + s^2b^2/2}. \quad (81)$$

Using this fact and (75), the mean of $Y_t$ is given by

$$E[Y_t] = Y_0 e^{\nu t + \sigma^2 t/2} = Y_0 e^{(\nu + \sigma^2/2)t}, \quad (82)$$

and its variance is given by

$$Var[Y_t] = E[Y_t^2] - (E[Y_t])^2 \quad (83)$$

$$= E[Y_0^2 e^{2X_t}] - (E[Y_t])^2 \quad (84)$$

$$= Y_0^2 [e^{(\nu + \sigma^2)t} - e^{2\nu + 2\sigma^2 t}]. \quad (85)$$

6 Binomial Lattice Revisited

6.1 Convergence

Consider an $N$-period binomial lattice in which $S_k = uS_{k-1}$ with probability $p$ and $S_k = dS_{k-1}$ with probability $1 - p$, with $u > 1$ and $d = 1/u$. By taking the natural logarithm of both sides
of the following identity,

$$\frac{S_n}{S_0} = \frac{S_n}{S_{n-1}} \frac{S_{n-1}}{S_{n-2}} \cdots \frac{S_1}{S_0},$$  \hspace{1cm} (86)$$

it follows immediately that

$$S_n := \ln \frac{S_n}{S_0} := \sum_{i=1}^{n} X_i, \hspace{1cm} (87)$$

where

$$X_i := \ln \frac{S_i}{S_{i-1}} \hspace{1cm} (88)$$

for each $1 \leq i \leq N$. For the binomial lattice the $X_i$ are i.i.d. random variables with common distribution

$$P(X_i = \ln u) = p, \; P(X_i = -\ln u) = 1 - p. \hspace{1cm} (89)$$

Now compare (89) to (35). It should be clear that the $S_n$ define a random walk in which $\ln u$ plays the role of $\sigma$. Accordingly, for a given $T > 0$, we know if we set

$$u := e^{\sigma \sqrt{\Delta t}}, \hspace{1cm} (90)$$

$$d := e^{-\sigma \sqrt{\Delta t}}, \hspace{1cm} (91)$$

$$p := \frac{1}{2} \left(1 + \frac{\nu}{\sigma} \sqrt{\Delta t}\right), \hspace{1cm} (92)$$

then as $\Delta t \to 0$

(i) the distribution of $\ln \frac{S_t}{S_0}$ converges to a normal distribution with mean $\nu t$ and variance $\sigma^2 t$, and

(ii) the distribution of $\ln \frac{S_{t+s}}{S_t}$ will be normal with mean $\nu s$ and variance $\sigma^2 s$.

Thus, the natural logarithm of the normalized price process (normalized by $S_0$) is a Brownian Motion, and so the price process $\{S_t : 0 \leq t < \infty\}$ is a Geometric Brownian Motion process.

6.2 Interpretation of the parameter $\nu$

The parameter $\nu$ determines the mean of $\ln S_t/S_0$. By definition, the random continuously compounded rate of return over the interval $[0, t]$, call it $a$, satisfies the equation $S_t = S_0 e^{at}$. (It is random since it depends on the value of $S_t$.) Since

$$a = [\ln S_t/S_0]/t = Y_t/t, \hspace{1cm} (93)$$

its expectation

$$E[a] = E[Y_t]/t = \nu. \hspace{1cm} (94)$$

Consequently, the parameter $\nu$ is the expected continuously compounded rate of return on the stock.
6.3 Interpretation of the parameter $\mu$

From (82)

$$E[S_t] = S_0 e^{(\nu + \sigma^2/2)t} := S_0 e^{\mu t}. \tag{95}$$

By definition, the (random) growth rate or proportional rate of return of the stock over a short interval $[0, \Delta t]$ is given by

$$(S_{\Delta t}/S_0 - 1)/\Delta t, \tag{96}$$

whose expected value from (95) is

$$e^{\mu \Delta t} - 1 \approx \frac{(1 + \mu \Delta t) - 1}{\Delta t} = \mu \tag{97}$$

when $\Delta t$ is close to 0. (Recall that $e^x \approx 1 + x$ when $x \approx 0$.) Consequently, the parameter

$$\mu = \nu + \sigma^2/2 \tag{98}$$

is the expected growth rate or proportional rate of return of the stock.

6.4 Relationship between $\nu$ and $\mu$

For the stochastic price process described here the expected continuously compounded rate of return and the expected proportional rate of return are linked via the identity (98) and are most definitely not equal. If the price process $S_t$ were deterministic with $S_t = S_0 e^{\mu t}$, then clearly the continuously compounded rate of return would be $\mu$ and $\mu$ would equal $\nu$.

In the stochastic setting, an adjustment of $\sigma^2/2$ must be made when “moving” from the logarithm of (normalized) prices to the (normalized) prices themselves. (We will understand this phenomenon better when we delve into stochastic calculus.) Often, the parameter $\mu$ is given (as well as $\sigma$), from which one calculates $\nu$. The next section gives some example calculations.

7 Examples

We follow common convention and measure time in years so that $t = 1$ corresponds to 1 year from now. A blue chip stock like Citibank would have its annual standard deviation $\sigma$ of say 28% and a stock like Microsoft would have its $\sigma$ around 40%. Keep in mind the variance of the logarithm of the price process is proportional to time, and so the standard deviation is proportional to the square-root of time.

Suppose the expected proportional rate of return on Citibank or Microsoft stock is $\mu = 12\%$ per year or $1\%$ per month. For $t = 1/12$ (1 month)
• \( \ln (S_t/S_0) = X_t \sim N([0.12 - .28^2/2] \ast (1/12), 0.0808^2). \)

• A positive 3σ event corresponds to \( \ln (S_t/S_0) = 0.00673 + 3(0.0808) \) or \( S_t = 1.283S_0. \)

• A negative 3σ event corresponds to \( \ln (S_t/S_0) = 0.00673 - 3(0.0808) \) or \( S_t = 0.790S_0. \)

When \( t = 3/12 \) (3 months)

• \( \ln (S_t/S_0) = X_t \sim N([0.12 - .28^2/2] \ast (3/12), 0.14^2). \)

• A positive 3σ event corresponds to \( \ln (S_t/S_0) = 0.0202 + 3(0.14) \) or \( S_t = 1.553S_0. \)

• A negative 3σ event corresponds to \( \ln (S_t/S_0) = 0.0202 - 3(0.14) \) or \( S_t = 0.670S_0. \)

For Microsoft the corresponding numbers are \( S_t = 1.419S_0 \) and \( S_t = 0.710S_0 \) for \( t = 1/12 \) and \( S_t = 1.847S_0 \) and \( S_t = 0.556S_0 \) for \( t = 3/12. \)

Suppose the current price for Microsoft Stock is \( S_0 = 100, \) and assume \( \mu = 0.12 \) and \( \sigma = 0.40. \) Here, \( \nu = \mu - \sigma^2/2 = 0.12 - (0.40)^2/2 = 0.04. \) Let \( t = 1/12, \) which corresponds to the end of this month. We shall use the fact that the stochastic process \( X_t := \ln S_t/S_0 \) is \((\nu, \sigma)\) Brownian motion with \( \nu = 0.04. \)

• What is the probability that the stock price will exceed 120 by time \( t? \) We seek

\[
P(S_t > 120) = P\left( \frac{\ln S_t/S_0 - \nu t}{\sigma \sqrt{t}} > \frac{\ln(1.2) - 0.04/12}{0.4\sqrt{1/12}} \right)
= 1 - \Phi(1.55) = 0.0657.
\]

• Suppose we hold one share of Microsoft stock and want to know the chance that over the next month the stock price will never exceed 120? The event that \( S_t \leq 120 \) for each \( \tau \in [0, t] \) is equivalent to the event that \( X_\tau = \ln S_\tau/S_0 \leq y \) for each \( \tau \in [0, t], \) where \( y = \ln 1.2. \) As a result, we can directly apply (20) with the \( \mu \) parameter there set to 0.04 to obtain that

\[
P\left\{ \max_{0 \leq \tau \leq t} S_\tau \leq 120 \right\} = P(M_t \leq y)
= \Phi(1.55) - (1.0954)\Phi(-1.61)
= 0.9343 - (1.0954)(0.0537) = 0.8806,
\]

and so there is an 88% chance the stock price will never exceed 120 over the next month. In addition, we know the chance that it will take at least one month for the stock price to exceed 120 is also 88%.  

18
Suppose we buy one share of Microsoft stock and decide to hold it until it either rises to 110 (our cash out point) or it falls to 80 (our stop/loss point). What is the probability we will cash out with a profit of 10? We can directly apply (22) with \( b = \ln 1.2, x = 0 \) and \( a = \ln 0.80 \) to obtain that

\[
P(S(T_{ab}) = 120) = P(X(T_{ab}) = y) = \frac{1 - e^{-[2(0.04)/0.40^2]a}}{e^{-[2(0.04)/0.40^2]b} - e^{-[2(0.04)/0.40^2]a}}
\]

\[
= \frac{-0.118034}{-0.205163} = 0.5753.
\]

8 Derivation of the Black-Scholes Call Option Pricing Formula

We calculate the value of a call option on a non-dividend paying stock by modeling the price process via an \( N \)-step binomial lattice, and then computing the discounted expected value of the option’s payoff at time \( T \) using the risk-neutral probability.

For the binomial lattice we set the parameter \( p \) to be the risk-neutral probability. The risk-neutral probability is the unique value for \( p \) that ensures the discounted expectation of next-period’s stock prices using the risk-free rate of return \( r \Delta t \) always equals the current price, namely,

\[
S_{k-1} = \frac{p(uS_{k-1}) + (1 - p)(dS_{k-1})}{1 + r\Delta t}.
\]

(107)

The unique value for \( p \) is

\[
p = \frac{(1 + r\Delta t)S_{k-1} - dS_{k-1}}{uS_{k-1} - dS_{k-1}} = \frac{(1 + r\Delta t) - d}{u - d}.
\]

(108)

Since \( u = e^{\sigma \sqrt{\Delta t}} \), \( d = e^{-\sigma \sqrt{\Delta t}} \) and \( \Delta t \approx 0 \), we may use the approximation \( 1 + x + x^2/2 \) for \( e^x \) in (108) to obtain

\[
p \approx \frac{(1 + r\Delta t) - (1 - \sigma \sqrt{\Delta t} + \sigma^2 \Delta t/2)}{(1 + \sigma \sqrt{\Delta t} + \sigma^2 \Delta t/2) - (1 - \sigma \sqrt{\Delta t} + \sigma^2 \Delta t/2)}
\]

(109)

\[
= \frac{(r - \sigma^2/2)\Delta t + \sigma \sqrt{\Delta t}}{2\sigma \sqrt{\Delta t}}
\]

(110)

\[
\approx \frac{1}{2} \left(1 + \frac{\nu \sqrt{\Delta t}}{\sigma}\right),
\]

(111)

where

\[
\nu := r - \sigma^2/2.
\]

(112)

In light of previous developments, it should be clear that when there are many steps to the binomial lattice the distribution of \( S(T)/S(0) \) is lognormal with parameters \( \nu = (r - \sigma^2/2) \) and...
Since the expected growth rate of the stock is \( \nu + \sigma^2/2 \), we see that under the risk-neutral probability the expected growth rate of the stock is the risk-free rate!

In the limit, as the number of steps of the binomial lattice goes to infinity, the value of a European call option on a non-dividend paying stock is given by

\[
E[e^{-rT}\max(S_T - K, 0)]
\]

(113)

when the underlying distribution is

\[
\ln \frac{S(T)}{S(0)} \sim N((r - \sigma^2/2)T, \sigma^2 T).
\]

(114)

It is possible to compute this expectation and derive a closed-form analytical solution for a call option value, known as the famous Black-Scholes formula, to which we now turn.

To simplify the derivation to follow we express \( S(T) = S(0)e^X \) where \( X \sim N(a, b^2) \). Upon substitution, (113) may be expressed via the following equalities:

\[
e^{-rT}\int_{-\infty}^{+\infty} \max(S_0e^x - K, 0) \frac{1}{\sqrt{2\pi}b} e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} dx
\]

(115)

\[
e^{-rT}\int_{\ln K/S_0}^{+\infty} (S_0e^x - K) \frac{1}{\sqrt{2\pi}b} e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} dx
\]

(116)

\[
= S_0e^{-rT}\int_{\ln K/S_0}^{+\infty} \frac{1}{\sqrt{2\pi}b} e^{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2} dx - Ke^{-rT}\int_{\ln K/S_0}^{+\infty} \frac{1}{\sqrt{2\pi}b} e^{\frac{1}{2}\left(\frac{x-a}{b}\right)^2} dx.
\]

(117)

By completing the square in the exponent of the exponential, the first integral on the right-hand side of (117) is equivalent to

\[
S_0e^{-rT}e^{a+b^2/2}\int_{\ln K/S_0}^{+\infty} \frac{1}{\sqrt{2\pi}b} e^{-\frac{1}{2}\left(\frac{y-\frac{a+b^2}{b}}{b}\right)^2} dy,
\]

(118)

which may be further simplified via the following chain of equalities:

\[
= S_0e^{-rT}e^{a+b^2/2}\int_{\ln \frac{K}{S_0}-(a+b^2)}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy
\]

(119)

\[
= S_0e^{-rT}e^{a+b^2/2} \left[ 1 - \Phi\left( \frac{\ln K/S_0 - (a + b^2)}{b} \right) \right]
\]

(120)

\[
= S_0\Phi\left[ \frac{\ln S_0/K + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \right]
\]

(121)

\[
:= S_0\Phi(d_1).
\]

(122)

after substituting the definitions for \( a = (r - \sigma^2/2)T \) and \( b^2 = \sigma^2T \).

The second integral on the right-hand side of (117) is equivalent to

\[
Ke^{-rT}\int_{\ln K/S_0-a}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy.
\]

(123)
which may be further simplified via the following chain of equalities:

\[
Ke^{-rT} [1 - \Phi\left(\frac{\ln K - a}{b}\right)] = Ke^{-rT} [\Phi\left(\frac{\ln S_0 - a}{b}\right)] = Ke^{-rT} \left[\Phi\left(\frac{\ln S_0 - (r - \sigma^2/2)T}{\sigma\sqrt{T}}\right)\right] := Ke^{-rT} \Phi(d_2),
\]

after substituting once again for the definitions for \(a\) and \(b\).

**Remark 8** Note that

\[
d_2 = d_1 - \sigma\sqrt{T}.
\]

**Remark 9** Let \(C/S\) denote the value of the call option as a percentage of the current stock price, let

\[
\kappa_1 := S/Ke^{-rT},
\]

\[
\kappa_2 := \sigma\sqrt{T}.
\]

Note that

\[
\frac{C}{S} = \Phi\left[\frac{\ln \kappa_1 + \kappa_2}{2}\right] - \frac{1}{\kappa_1} \Phi\left[\frac{\ln \kappa_1 - \kappa_2}{2}\right],
\]

which shows that one only needs the values for \(\kappa_1\) and \(\kappa_2\) to compute the value of the call option.

### 9 Call Option Pricing Examples

To summarize the development in the previous section the famous Black-Scholes European call option formula is given by

\[
C(K, T) = S\Phi(d_1) - Ke^{-rT}\Phi(d_2)
\]

where recall that \(\Phi(\cdot)\) denotes the cumulative standard normal density, and

\[
d_1 = \frac{\ln (S/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}},
\]

\[
d_2 = \frac{\ln (S/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}.
\]
Notice that the parameter \( \mu \) is nowhere to be found in this formula!

Here are some examples. Consider a call option on Citibank stock. Suppose current value of its stock is \( S_0 = 100 \). The annual standard deviation of log-volatility is 28% and the annual risk-free rate will be set to 5%. As a function of the \( K \) and \( T \) we have that

\[
d_1(K,T) = \frac{\ln \left( \frac{100}{K} \right) + 0.0892T}{0.28\sqrt{T}},
\]
\[
d_2(K,T) = d_1(K,T) - 0.28\sqrt{T},
\]
\[
C(K,T) = 100\Phi(d_1(K,T)) - Ke^{-0.05T}\Phi(d_2(K,T)).
\]

Now consider the following options.

- Cost of a 1-month call option with strike price = 100? Here \( d_1 = 0.0920, d_2 = 0.0111, \Phi(d_1) = 0.5367, \Phi(d_2) = 0.5044 \) and \( C(100, 1/12) = 100(0.5367) - 99.5842(0.5044) = 3.44. \)
- Cost of a 3-month call option with strike price = 100? Here \( d_1 = 0.1593, d_2 = 0.0193, \Phi(d_1) = 0.5632, \Phi(d_2) = 0.5076 \) and \( C(100, 3/12) = 100(0.5632) - 98.75778(0.5076) = 6.19. \)
- Cost of a 1-month call option with strike price = 105? Here \( d_1 = -0.5117, d_2 = -0.5925, \Phi(d_1) = 0.3043, \Phi(d_2) = 0.2768 \) and \( C(105, 1/12) = 100(0.3043) - 104.5634(0.2768) = 1.49. \)
- Cost of a 3-month call option with strike price = 105? Here \( d_1 = -0.1892, d_2 = -0.3292, \Phi(d_1) = 0.4250, \Phi(d_2) = 0.3710 \) and \( C(105, 3/12) = 100(0.4250) - 103.6957(0.3710) = 4.03. \)

Now consider what happens when \( \sigma = 40\%: \)

- Cost of a 1-month call option with strike price = 100? Here \( d_1 = 0.0938, d_2 = -0.0217, \Phi(d_1) = 0.5375, \Phi(d_2) = 0.4916 \) and \( C(100, 1/12) = 100(0.5375) - 99.5842(0.4916) = 4.79. \)
- Cost of a 3-month call option with strike price = 100? Here \( d_1 = 0.1625, d_2 = -0.0375, \Phi(d_1) = 0.5646, \Phi(d_2) = 0.4850 \) and \( C(100, 3/12) = 100(0.5646) - 98.75778(0.4850) = 8.56. \)
- Cost of a 1-month call option with strike price = 105? Here \( d_1 = -0.3287, d_2 = -0.4441, \Phi(d_1) = 0.3712, \Phi(d_2) = 0.3284 \) and \( C(105, 1/12) = 100(0.3712) - 104.5634(0.3284) = 2.78. \)
- Cost of a 3-month call option with strike price = 105? Here \( d_1 = -0.0815, d_2 = -0.2815, \Phi(d_1) = 0.4625, \Phi(d_2) = 0.3891 \) and \( C(105, 3/12) = 100(0.4625) - 103.6957(0.3891) = 5.90. \)

Notice that as the time to expiration increases the call option value increases, and that as the strike price increases the call option value decreases. Both observations are true in general.
10 One-Dimensional Ito Processes

A stochastic process that encompasses Brownian Motion and Geometric Brownian Motion is the Ito Process (also known as a stochastic integral). We shall describe a subset of Ito processes that possess the Markov property.

10.1 Definition and interpretation

In differential form $X$ is a (Markov) Ito process if it may be represented as

$$dX_t = \mu(X_t, t)dt + \sigma(X_t, t)dB_t$$

(135)

for appropriate choices for the functions $\mu$ and $\sigma$. Equations of type (135) are known as stochastic differential equations. One loosely interprets (135) to mean that for sufficiently small $\Delta t$,

$$X(t + \Delta t) - X(t) \approx \mu(X_t, t)\Delta t + \sigma(X_t, t)[B(t + \Delta t) - B(t)].$$

(136)

Keep in mind that the both sides of (136) are random variables, and that

$$B(t + \Delta t) - B(t) \sim N(0, \Delta t).$$

(137)

One may use (136) to simulate the $X$ process, as follows. Fix a small period length $\Delta t > 0$, and define $t_i = i\Delta t$ for integer $i \geq 1$. For each $i$ let $Z_i$ be i.i.d. random variables with common distribution given by $N(0, \Delta t)$. Beginning with the known constant $X(0)$, simulate $Z_1, Z_2, \ldots$ and sequentially set

$$X(t_1) := X(0) + \mu(X(0), 0)\Delta t + \sigma(X(0), 0)Z_1$$

(138)

$$X(t_2) := X(t_1) + \mu(X(t_1), t_1)\Delta t + \sigma(X(t_1), t_1)Z_2$$

(139)

$$\ldots$$

$$X(t_{k+1}) := X(t_k) + \mu(X(t_k), t_k)\Delta t + \sigma(X(t_k), t_k)Z_{k+1} \ldots$$

(140)

Example 3 The simplest example is when $\mu(X_t, t) = \mu$ and $\sigma(X_t, t) = \sigma > 0$, in which case

$$dX_t = \mu dt + \sigma dB_t.$$ 

(142)

Although we have not formally discussed what it means to integrate both sides of (142), any reasonable definition of integration would suggest that

$$X_t - X_0 = \mu t + \sigma B_t,$$

(143)

which is precisely Brownian Motion.

---

3General Ito processes permit the functions $\mu$ and $\sigma$ to depend on the whole history of the $X$ process up to time $t$. 

Example 4 Here, let $\mu(X_t, t) = \mu X_t$ and $\sigma(X_t, t) = \sigma X_t > 0$, in which case

$$dX_t = \mu X_t dt + \sigma X_t dB_t. \quad (144)$$

Since $X_t$ appears on both sides of (144) we cannot directly integrate both sides to solve for $X_t$. By dividing both sides by $X_t$,

$$\frac{dX_t}{X_t} = \mu dt + \sigma dB_t. \quad (145)$$

Since we interpret $dX_t/X_t$ to mean $X(t + \Delta t)/X(t) - X(t)/X(t)$ for small $\Delta t$, and since

$$\frac{X(t + \Delta t) - X(t)}{X(t)} \approx \ln \frac{X(t + \Delta t)}{X(t)} = \ln (X(t + \Delta t) - \ln X(t),$$

it would appear that (145) is equivalent to

$$d \ln X(t) = \mu dt + \sigma dB_t. \quad (146)$$

Since $X_t$ no longer appears on the right-hand side of (146), we can integrate both sides as before to obtain

$$\ln X(t) - \ln X(0) = \mu t + \sigma B_t. \quad (147)$$

The conclusion is that here $X$ is a Geometric Brownian Motion process. Actually, this heuristic derivation is correct up to a point, but the “calculus” performed is not correct. The drift parameter does not equal $\mu$, as suggested in (147), but must be adjusted to $\mu - \sigma^2/2$. This is not too much of a surprise given our discussion relating to (98).

Example 5 For the mean-reversion model

$$\mu(X_t, t) = a(b - X_t), \quad \sigma(X_t, t) = \sigma X_t.$$ 

Here, the parameter $b$ denotes the long-run average value of the process $X$, and the parameter $a > 0$ determines the speed with which the $X$ process adjusts towards $b$. Note that when $X_t < b$ the instantaneous drift is positive, and there will be a tendency for $X_t$ to increase; when $X_t > b$, the instantaneous drift is negative, and there will be a tendency for $X_t$ to decrease.

Mean-reversion is often used to model prices of natural resources. As the price increases above the long-term average, producers increase supply to take advantage of the higher prices, which then has a natural tendency to bring down the prices. The opposite would hold true when the price decreases below the long-term average.

10.2 Functions of Ito processes and Ito-Doeblin’s lemma

Let $X_t$ be an Ito process as in (135). Let $g(t, x)$ be a twice continuously differentiable function defined on $[0, \infty) \times R$. What can be said about the process $Y_t := g(t, X_t)$?
The heuristic development proceeds, as follows. Using a second-order Taylor series expansion about the point \((t, X_t)\), for \(\Delta t\) sufficiently small

\[
Y(t + \Delta t) - Y(t) \approx \frac{\partial g}{\partial t} \Delta t + \frac{\partial g}{\partial x} \Delta X_t + \frac{1}{2} \frac{\partial^2 g}{\partial t \partial x} (\Delta t)^2 + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} (\Delta X_t)^2,
\]

where we let

\[
\Delta X_t := X(t + \Delta t) - X(t).
\]

(148)

The idea from here on is to eliminate those terms on the right-hand side of (148) that involve powers of \(\Delta t\) greater than 1. As \(\Delta t \to 0\) such terms will become increasingly inconsequential. The term involving \(\Delta t\) must obviously be kept, and so must the term involving \(\Delta X_t\), since its standard deviation goes as \(\sqrt{\Delta t}\), which becomes large relative to \(\Delta t\) when \(\Delta t\) is small.

Obviously, the \((\Delta t)^2\) may be eliminated. The term involving \(\Delta t \Delta X_t\) has (conditional) mean of zero and variance of \((\Delta t)^3\), and so this random variable converges to zero fast enough as \(\Delta t \to 0\), and so it may be eliminated, too.

We are left with the \((\Delta X_t)^2\). For notational convenience let

\[
\Delta B_t := B(t + \Delta t) - B(t), \; \mu_t := \mu(X_t, t), \; \sigma_t := \sigma(X_t, t).
\]

(150)

Using (136),

\[
(\Delta X_t)^2 = \mu_t^2 (\Delta t)^2 + 2 \mu_t \sigma_t \Delta B_t + \sigma_t^2 (\Delta B_t)^2.
\]

(151)

By similar reasoning as before, the first and second terms on the right-hand side of (151) may be eliminated, but the third term must be kept. In particular, it “converges” to \(\sigma_t^2 \Delta t\).

Putting it all together, as \(\Delta t \to 0\)

\[
dY_t := Y(t + \Delta t) - Y(t)
\]

\[
= \frac{\partial g}{\partial t} \Delta t + \frac{\partial g}{\partial x} (\Delta X_t) + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} (\Delta X_t)^2
\]

(152)

\[
= \frac{\partial g}{\partial t} \Delta t + \frac{\partial g}{\partial x} (\mu_t \Delta t + \sigma_t \Delta B_t) + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \sigma_t^2 \Delta t
\]

(153)

\[
= \left(\frac{\partial g}{\partial t} + \frac{\partial g}{\partial x} \mu_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \sigma_t^2 \right) \Delta t + \left(\frac{\partial g}{\partial x} \sigma_t \right) \Delta B_t.
\]

(154)

(155)

We see that \(Y_t\) is also an Itô process given by\(^4\)

\[
dY_t = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} (dX_t) + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} (dX_t)^2
\]

(156)

\[
= \left(\frac{\partial g}{\partial t} + \frac{\partial g}{\partial x} \mu_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \sigma_t^2 \right) dt + \left(\frac{\partial g}{\partial x} \sigma_t \right) dB_t.
\]

(157)

Equation (156) is known as *Itô-Doeblin’s formula* proved by Itô and Doeblin in what is now known as *Itô-Doeblin’s Lemma*. Equation (157) is the specific form for the special class of Itô processes considered here.

\(^4\) We assume we know the value of \(X_t\) at time \(t\) from which we may compute the value of \(Y_t\).
10.3 Examples of Ito Calculus

Ito-Doeblin’s formula is a powerful tool for analyzing functions of Brownian Motion. We illustrate with a number of examples.

**Example 6** Consider the Ito process described in (144). Let 
\[ Y_t = g(t, X_t) = \ln X_t. \]
Then
\[
\begin{align*}
dY_t &= \left( \frac{1}{X_t}(\mu X_t) + 1/2 \frac{1}{X_t^2}(\sigma^2 X_t^2) \right) dt + \left( \frac{1}{X_t}(\sigma X_t) \right) dB_t \\
&= (\mu - \sigma^2/2)dt + \sigma dB_t. \quad (159)
\end{align*}
\]
We may now integrate both sides of (159) to obtain that
\[
\ln X_t = \ln X_0 + \nu t + \sigma B_t \quad \text{where } \nu := \mu - \sigma^2/2. \quad (160)
\]
Equivalently, 
\[ X_t = X_0 e^{Z_t} \text{ with } Z_t \text{ being } (\nu, \sigma) \text{ Brownian Motion}. \]
Now since
\[ X_t = f(Z_t) := X_0 e^{Z_t}, \quad (161)\]
another application of Ito-Doeblin’s Lemma yields
\[
\begin{align*}
dx_t &= f'(Z_t) dZ_t + 1/2 f''(Z_t)(dZ_t)^2 \\
&= X_t[\nu dt + \sigma dB_t] + 1/2 X_t \sigma^2 dt \\
&= (\nu + \sigma^2/2)X_t dt + \sigma X_t dB_t \\
&= \mu X_t dt + \sigma X_t dB_t, \quad (165)
\end{align*}
\]
which is identical to (144), as it should.

We may use Ito-Doeblin’s Formula to evaluate stochastic integrals, as the following example illustrates.

**Example 7** Although we have as yet given no meaning to the expression
\[
\int_0^t B_s dB_s, \quad (166)
\]
we can evaluate it. From classical calculus one would suspect the term 1/2 \( B_t^2 \) would crop up somewhere. So, let \( g(t, B_t) = 1/2 B_t^2 \). Here, \( X_t \) is simply \( B_t \) and \( \mu_t = 0 \) and \( \sigma_t \) is identically one. We have
\[
dY_t = 1/2 \ dt + B_t dB_t, \quad (167)
\]
which when “integrated” gives
\[
1/2 B_t^2 = 1/2 t + \int_0^t B_s dB_s, \quad (168)
\]
In other words,

\[ \int_0^t B_s dB_s = \frac{1}{2} (B_t^2 - t). \]  

(169)

Working in reverse, if we define \( f(B_t, t) := \frac{1}{2} (B_t^2 - t) \), then another application of Ito-Doeblin’s Lemma will yield that \( df_t = B_t dB_t \), as it should. (Verify this.)

11 Black-Scholes Revisited

11.1 Black-Scholes differential equation

Let \( C(S_t, t) \) denote the value of a European call option on a non-dividend paying stock, whose stock price \( S_t \) follows the stochastic differential equation (144). Recall that under the risk-neutral probability the growth rate of the stock is the risk free rate \( r \) and so \( \mu = r \). By Ito-Doeblin’s Formula,

\[ dC = \left( C_t + rSC_S + \frac{1}{2} \sigma^2 S^2 C_{SS} \right) dt + (\sigma SC_S) dB_t, \]  

(170)

where for convenience we let \( C_t = \partial C/\partial t \), \( C_S = \partial C/\partial S \) and \( C_{SS} = \partial^2 C/\partial S^2 \), and we suppressed the \( t \) from \( S_t \).

Let \( \Pi = -hS + C \) denote a portfolio of \( -h \) units of \( S \) and 1 unit of \( C \). The instantaneous change in the portfolio \( \Pi \) is given by

\[ d\Pi = -hdS + dC = -h(rSdt + \sigma SdB_t) + dC. \]  

(171)

In light of (170), if we set \( h = C_S \), then the stochastic term involving \( B_t \) vanishes! In particular, for this choice of \( h \) we have

\[ d\Pi = (C_t + 1/2 \sigma^2 S^2 C_{SS}) dt. \]  

(172)

Since the portfolio is instantaneously risk-free, it must earn the risk-free rate, namely,

\[ d\Pi = r\Pi dt = r(-C_S S + C) dt. \]  

(173)

Putting (172) and (173) together, it follows that the Call Option value must satisfy the partial differential equation given by

\[ rC = C_t + rSC_S + \frac{1}{2} \sigma^2 S^2 C_{SS}. \]  

(174)

Equation (174) is known as the Black-Scholes differential equation. One may verify that the previously given closed-form expression for the Black-Scholes formula does indeed satisfy (174).
11.2 The “Greeks”

The instantaneous units of stock to hold, \( h \), is called the hedge ratio. It is so named since
\[
\frac{\partial C}{\partial S} = \frac{\Delta C}{\Delta S}.
\]
It is the limit of the hedge ratio obtained in the replicating portfolio obtained in our Binomial Lattice models. Instantaneously, \( \Pi \) acts like a riskless “bond”, and the replicating portfolio is given by \( hS + \Pi \).

Let

- \( \Delta := C_S \) denote the sensitivity of the call option value to the price \( S \);
- \( \Theta := C_t \) denote the sensitivity of the call option value to the remaining time \( t \); and
- \( \Gamma := C_{SS} \) denote the sensitivity of \( \Delta \) to \( S \).

From the Black-Scholes differential equation we see that
\[
rc = \Theta + rS\Delta + \frac{1}{2} \sigma^2 S^2 \Gamma. \tag{175}
\]
It is not too difficult to show that
\[
\Delta = \Phi(d_1), \tag{176}
\]
from which it follows that
\[
\Gamma = \frac{\phi(d_1)}{S\sigma\sqrt{T}}, \tag{177}
\]
where recall that \( \phi(\cdot) \) denotes the density function of a standard normal random variable. We know that
\[
C = S\Phi(d_1) - Ke^{-rT}\Phi(d_2).
\]
Consequently, after a little algebra, the value of \( \Theta \) is determined as
\[
\Theta = -\frac{S\phi(D_1)\sigma}{2\sqrt{T}} - rKe^{-rT}\Phi(d_2). \tag{178}
\]

12 Application to Real Options Valuation

Let the value of a project be governed by the stochastic differential equation
\[
dV_t = \mu V_t dt + \sigma V_t dB_t. \tag{179}
\]
Let \( \rho \) denote the appropriate discount rate (cost of capital) for projects of this type. We assume \( \rho > \mu > 0 \).
Let the fixed constant $I$ denote the required investment when the project is undertaken. The only question is one of timing, namely, when should the project be undertaken? Clearly, the optimal timing decision should be based on the current value of the project, $V_t$. Since the problem parameters are independent of time, this is the only piece of information required, i.e., it is the single state variable.

Let $F(V_t)$ denote the value of the project with this embedded *delay option*. At the current point in time, only two actions are possible: (i) invest now in the project or (ii) continue to wait. At the instant the investment takes place, the value of the project is $V_t - I$. If, on the other hand, it is not optimal at this time to invest, then the current value must equal the discounted expectation of the future value. For the next $\Delta t$ units of time, this discounted expectation is given by

$$
1 + \rho \Delta t E_t[F(V_{t+\Delta t})],
$$

where the expectation is taken with respect to the information known at time $t$. By Bellman’s Principle of Optimality,

$$
F(V_t) = \max \left\{ V_t - I, \frac{1}{1 + \rho \Delta t} E_t[F(V_{t+\Delta t})] \right\}.
$$

Suppose the value of $V_t$ is such that we are in the “continuation region”. Then,

$$
F(V_t) = \frac{1}{1 + \rho \Delta t} E_t[F(V_{t+\Delta t})].
$$

Multiply both sides of (182) by $1 + \rho \Delta t$ and then subtract $F(V_t)$ to obtain:

$$
\rho F(V_t) \Delta t = E_t[F(V_{t+\Delta t}) - F(V_t)].
$$

(183)

(As before, the subscripts on $F$ indicate partial derivatives.) It then follows that

$$
E_t[F(V_{t+\Delta t}) - F(V_t)] = E_t[dF] = (\mu V F_t + 1/2 \sigma^2 V^2 F_{VV}) \Delta t.
$$

(Keep in mind that $E_t[\Delta B_t] = 0$.) We conclude then that in the continuation region, the value function $F$ satisfies the following differential equation:

$$
0 = \mu V F_t + 1/2 \sigma^2 V^2 F_{VV} - \rho F.
$$

(187)

It may be readily checked that $F(V) = AV^\beta$ is a solution to (187). Substituting into (187), we have

$$
0 = \mu \beta + 1/2 \sigma^2 \beta (\beta - 1) - \rho := q(\beta).
$$

(188)
Let quadratic form $q(\beta)$ is convex, its value at 0 is negative, its value at 1 is negative, too, since $\rho > \mu$ by assumption, and it is unbounded as $\beta$ or $-\beta$ tend to infinity. Consequently, $q$ has two real roots $\beta_1 < 0$ and $\beta_2 > 1$. The function

$$A_1 V^{\beta_1} + A_2 V^{\beta_2} \tag{189}$$

solves (187). Clearly, if $V = 0$, then $F(V) = 0$, too, which rules out the negative root. Thus,

$$F(V) = A_2 V^{\beta_2}. \tag{190}$$

It remains to determine the value of $A_2$ and to determine the optimal control policy. The optimal control policy is a threshold policy: when $V$ reaches a critical threshold $V_*$, then it is time to implement. At implementation, it must be the case that

$$F(V_*) = V_* - I; \tag{191}$$

the left-hand side can never be less than the right-hand side, and if it were greater, then implementation should have commenced a few moments ago. Equation (191) is one of two necessary equations to pin down the two values $V_*$ and $A_2$. It turns out that at implementation

$$\frac{d}{dV} F(V) \big|_{V_*} = \frac{d}{dV} (V - I) \big|_{V_*}, \tag{192}$$

which is known as the smooth pasting condition. Thus,

$$A_2 \beta_2 V_*^{\beta_2 - 1} = 1. \tag{193}$$

With this second equation in hand, both (191) and (193) pin down both $A_2$ and $V_*$. Their solutions are, respectively,

$$V_* = \frac{\beta_2}{\beta_2 - 1} I, \tag{194}$$

$$A_2 = \frac{V_* - I}{V_*^{\beta_2}} = \frac{(\beta_2 - 1)^{\beta_2 - 1}}{\beta_2^{\beta_2 - 1} I^{\beta_2 - 1}}. \tag{195}$$

We conclude that

$$F(V) = \begin{cases} A_2 V^{\beta_2}, & \text{if } V \leq V_*, \\ V - I, & \text{if } V \geq V_*. \end{cases} \tag{196}$$