Imperial College London

Department of Mathematics

Year 2 Mathematics Group Project

From random walk to the Black-Scholes model: A brief tour of stochastic processes and their applications

---

Supervisor: Dr Andreas Sojmark

Department: Department of Mathematics

Course: M2R - Mathematics Group Project

Group Number: PA4

Students: Timothy Minn Kang, Arthur Limoge, Munazza Sarwar, Zhuohan Yang, Yongda Zhu

CIDs: 01501998, 01579725, 01494078, 01523004, 01492047

Date: November 17, 2020

Contents

Introduction

Our aim, along this project, will be to study the properties of stochastic processes -which are simply collections of random variables indexed by time- in order to model the evolution of random events over time. We will try to start building up from simple processes, and see how their study can lead to a better understanding of real-world models, like financial markets.

Therefore, we will start with one of the simplest stochastic models: the random walk. We will then try and see what happens when we make this discrete model continuous: defining Brownian motion along the way. From this, we will seek to derive the heat equation and see how it can be used to model the dynamics of financial markets: under the *Black-Scholes* model.

We will then explore whether this is a reasonable model for financial markets, by testing our models against real data. We will finish by discussing the motivation for further variants of stochastic calculus, in particular Itō calculus.

1 Theory of random walks: from discrete to continuous models

1.1 Simple random walks in a discrete model

1.1.1 Study of simple random walks

Consider a walker placed at 0 on the integer line. At each time unit, the walker has a 1/2 chance of either moving forward or backward 1 unit length. Let S_n denote the position of the walker at time n, then:

$$S_n = \sum_{i=1}^{n} X_i$$

where

$$X_i = \begin{cases} 1, & \text{if the walker moved forward at time i} \\ -1, & \text{if the walker moved backward at time i} \end{cases}$$

In order to understand the properties of the walk, we will ask ourselves:

1. On average, how far is the walker from the starting point?
2. More generally, what is the probability distribution for the position of the walker?
3. Does the random walker keep returning to the origin or do they eventually leave forever?

Question 1: Average distance from the origin

Since at each time unit, the walker has an equal chance of moving forward or backward, the expectation is $\mathbb{E}(S_n) = 0$. The average distance, however, can be expressed as the standard deviation: $\sigma_n = \sqrt{\mathbb{E}(S_n^2) - \mathbb{E}(S_n)^2} = \sqrt{\mathbb{E}(S_n^2)}$, since $\mathbb{E}(S_n) = 0$.

$$\begin{aligned} \mathbb{E}(S_n^2) &= \mathbb{E}\Big[\Big(\sum_{i=1}^{n} X_i\Big)^2\Big] \\ &= \mathbb{E}\Big[\sum_{i=1}^{n}\sum_{j=1}^{n} X_i X_j\Big] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}(X_i X_j) = n, \text{ since } \mathbb{E}(X_i X_i) = \delta_{ij} \end{aligned}$$

Therefore $\sigma_n = \sqrt{n}$. This is a very important property, which will come in useful when we start testing data to check a Random Walk Hypothesis.

Question 2: Probability distribution of the position

Whenever the walker returns to origin, they must have moved forward and backward equally often. Thus the (discrete) index of time has to be even. And one can find the probability as:

$$\forall n \in \mathbb{N} : \mathbb{P}(S_{2n} = 0) = 2^{-2n} \cdot \binom{2n}{n}$$

More generally, if the walker is to be at $2j$ for $j \in \mathbb{N}$, there must be $(n+j)$ steps of moving forward and $(n-j)$ steps of moving backward. So: $\mathbb{P}(S_{2n} = 2j) = 2^{-2n} \cdot \binom{2n}{n+j}$. We can now use Stirling's formula: $n! \sim \sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}}$ asymptotically. So for $n \sim \infty$:

$$\mathbb{P}(S_{2n} = 2j) \sim \frac{n^{2n+\frac{1}{2}}}{\sqrt{\pi}(n+j)^{n+j+\frac{1}{2}}(n-j)^{n-j+\frac{1}{2}}}$$
$$\mathbb{P}(S_{2n} = 0) \sim \frac{n^{2n+\frac{1}{2}}}{\sqrt{\pi} n^{2n+1}} = \frac{1}{\sqrt{n\pi}}$$

Now, if we assume that $n >> j$, then $\frac{j}{n} \sim 0$. So one can consider the quantity $\ln \mathbb{P}(S_{2n} = j)$, and Taylor-expand it with respect to $\frac{j}{n}$.[1] In the limit, this gives $\mathbb{P}(S_{2n} = 2j) = \sqrt{\frac{1}{\pi n}} \exp(-\frac{j^2}{n})$. Using that we can derive the cumulative distribution function

$$\lim_{n\to\infty} \mathbb{P}\left\{a\sqrt{2n} \leq S_{2n} \leq b\sqrt{2n}\right\} \propto \lim_{n\to\infty} \sum_{n\geq 1} \frac{\sqrt{2}}{\sqrt{n}} e^{-j^2/n}$$

Question 3: Does the random walker ever escape to infinity?

To answer this, consider V be to be the number of times that the walker returns at the origin. The following argument is adapted from Lawler's book on random walks [1]

$$\mathbb{E}(V) = \sum_{n=0}^{\infty} P(S_{2n} = 0) = \pi^{-1/2} \cdot \sum_{n=0}^{\infty} n^{-1/2}$$

This well-known infinite series diverges to ∞. Now, let q be the probability that the walker ever returns to the origin, and compute the sum above differently:

$$\mathbb{E}(V) = 1 + q \cdot \mathbb{E}(V) \implies \mathbb{E}(V) = \frac{1}{1-q}$$

Since $E(V)$ diverges to ∞, $q = 1$. So the walker will return to the origin with probability one. But all that happens after can be considered as a new random walk; which will hence also have probability one of returning to the origin. This implies that the walker will return to the origin infinitely often, and hence almost surely cannot escape to infinity.

[1] Such an "informal" proof gives a (historical) intuition as to why a binomial distribution converges (in distribution) to a normal. This is the Moivre-Laplace theorem, a special case of the Central Limit Theorem. This proof was inspired by the physics lectures of François Vanderbrouck (PC*2 2017, Louis-Le-Grand, France.)

Generalization of the previous results to d dimensions

Now, let's generalize to higher dimension. Consider the integer lattice $\mathbb{Z}^d$, in discrete time. At each time unit, a walker has a $1/n$ chance of moving on a specific axis and a $1/2$ chance of either moving forward or backward on that axis. Let S_n denotes the walker's position at time n. If we assume that the walker starts at origin then

$$S_k = \begin{pmatrix} \sum_{i=1}^{n} X_{1i} \\ \sum_{i=1}^{n} X_{2i} \\ \vdots \\ \sum_{i=1}^{n} X_{ni} \end{pmatrix}$$

where

$$X_{ji} = \begin{cases} 1, & \text{if the walker moved forward at time i on j-axis} \\ -1, & \text{if the walker moved backward at time i on j-axis} \\ 0, & \text{if the walker didn't move at time i on j-axis} \end{cases}$$

Similarly to the 1-dimension case, $\mathbb{E}(S_n) = 0$ and $\sigma_n = \sqrt{n}$.

Furthermore, if the walker returns to origin, then they must have moved an even number of times on each axis - which happens with probability 2^{-d+1}. Generalizing from one dimension, we get:

$$\mathbb{P}(S_{2k} = 0) \sim 2^{-d+1} \cdot (\frac{1}{\sqrt{\frac{k}{d}\pi}})^d$$

Thus,

$$\mathbb{E}(V) = \sum_{k=0}^{\infty} \mathbb{P}(S_{2k} = 0) = 2^{-d+1} \cdot \sum_{k=0}^{\infty} (\frac{1}{\sqrt{\frac{k}{d}\pi}})^d \propto \sum_{r=0}^{\infty} r^{-d/2}$$

This well-known series converges when $n \geq 3$, and diverges to infinity otherwise. So, building up from our one-dimension reasoning, the walker will return to the origin infinitely many times if and only if $d \leq 2$.

1.1.2 Study of discrete boundary problems

Now, let us consider a volume A in space, and assume that heat particles follow a random walk in space (with infinitesimal time and space increments). Then the density of particles at some point in A gives the temperature. The differential equation governing the change in temperature with respect to time and space will be derived in the next part. To introduce it, let us consider some technical aspects such as: *what happens when the particle reaches the boundary of A?*

The following reasoning is inspired from that of [2]. In some simplified, for instance, we assume that particles of heat are "killed" at the boundary - if the latter has zero temperature. Such a definition naturally makes us wonder how long a particle usually takes to reach it (still in a discrete setting). As an example, let's first consider the integer lattice to be $(-1, +\infty)$. The walker will return to the origin infinitely many times. And every time after reaching origin, they have a $1/2$ chance of avoiding the boundary. So the probability that the walker reaches the boundary eventually can be expressed as $\lim_{n\to\infty} 1 - 2^{-n} = 1$.

Now, let's define a finite one-dimensional lattice - without loss of generality $\{0, 1, ..., N\}$. Again, set S_n to be the position of the walker (or particle) at time n, and let T be the number of steps until the walker reaches the boundary.[2] Let $E_x = \mathbb{E}(T|S_0 = x)$ be the expected time to reach a boundary given the walker starts at x. Naturally $E_x = 0$ if $x = 0 \vee N$, and since a particle can only move forward or backward:

$$E_x = \mathbb{P}(S_1 = x-1) \cdot E_{x-1} + \mathbb{P}(S_1 = x+1) \cdot E_{x+1} + 1 = \frac{1}{2}(E_{x-1} + E_{x+1}) + 1$$

Now, let's try to solve this recurrence equation to have a clearer picture of E_x.

$$\text{Set } E_x = A_x + B_x \text{ where } \begin{cases} A_x &= \frac{1}{2}(A_{x-1} + A_{x+1}) \qquad (\dagger) \\ B_x &= \frac{1}{2}(B_{x-1} + B_{x+1}) + 1 \quad (\star) \end{cases}$$

As described in the footnote, the solution to $(\dagger)$ can be found to be an affine function of x. So for some $k, h \in \mathbb{R} : A_x = kx + h$. Now, as we have seen in **1.1.1**, the walker's average deviation from the origin is proportional to the square root of time. From this property, we can guess that E_x will have a term varying in x^2. Hence, we make the Ansatz $B_x = ax^2$, for some $a \in \mathbb{R}$, and see:

$$\begin{aligned} &\text{Assume } B_x = ax^2 \\ &\qquad \implies ax^2 = a(x^2+1) + 1 \implies a = -1 \\ &\qquad \implies E_x = -x^2 + kx + h = -x^2 + Nx \text{ (by plugging in } E_0 = E_N = 0) \end{aligned}$$

This is hence the average number of steps a walker/particle will take to reach the boundary when starting at x. Notice that $0 \le x \le N \implies E_x \ge 0$, which is consistent with E_x being a measure of time.

This gives us a good understanding of boundary problems related to particle diffusion (and in particular, heat diffusion), and such reasoning can be used when studying the heat equation in a discrete model for example. We will not go into so much details, because we will generalize this to a continuous setting a bit later in section **1.2**. For the moment, we will try and see what happens to a random walk when the time and space increments are taken to go to 0.

[2] An interesting boundary problem would be to study the probability $F(x)$ that a particle starting at x reaches N before it reaches 0 - for example, if x is the money of a gambler, this is the probability of them not going ruined. Then, arguments of functional analysis give a unique and affine solution for F, extending to d dimensions. The arguments for this can be found in Lawler's book and give a great insight into discrete functional analysis.

1.2 Generalization to a continuous model

Brownian motion is the random movement of particles as a result of collisions in a fluid; named from the botanist Robert Brown, who observed pollen grains moving randomly in water. The mathematical model of Brownian motion can be roughly regarded as the one-dimensional scaled symmetric random walk by taking the time and space increments infinitely small.

As we will show in this section, Brownian motion is considered as a Gaussian process with special properties and continuous paths over continuous time. In the following, Brownian motion is constructed in an analytic way and some useful properties of Brownian motion will be given with proof. The construction of Brownian motion (BM) is a two-step process, we first define it on a countable, dense set of times -for simplicity- and then extend to all real by continuity.[3]

1.2.1 Construction of Brownian motion

We first construct a countable dense set in the following way: for $n \geq 0$, let D_n= $\{\frac{k}{2^n} : k = 0, 1, 2...\}$ The Dyadic rationals $\mathcal{D}$ are defined as: $\mathcal{D} = \bigcup\limits_{n\in\mathbb{N}} D_n$.

A standard one-dimensional Brownian motion on the dyadic rationals is a collection of random variables $\{W_t : t \in D\}$ satisfying:

(1)$W_0 = 0$

(2) For each $n \in \mathbb{N}$, the random variables $W_{\frac{k}{2^n}} - W_{\frac{k-1}{2^n}}$, k=1,2,3... are independent normal random variables with mean zero and variance 2^{-n}

(3) If $s, t \in \mathcal{D}$, then $W_t - W_s$ is independent of $\{W_r : r \leqslant s\}$and has a normal distribution with mean zero and variance t - s.

We write $J(k, n) = 2^{\frac{n}{2}}$ $[W_{\frac{k}{2^n}} - W_{\frac{k-1}{2^n}}]$. To show that the process defined above exists, we proceed as following:

Lemma 1.1

> Suppose X and Y are independent normal random variables with mean 0 and variance 1. If Z = $\frac{1}{\sqrt{2}}$X - $\frac{1}{\sqrt{2}}$Y and $\overline{Z}$ = $\frac{1}{\sqrt{2}}$X - $\frac{1}{\sqrt{2}}$Y, then Z and $\overline{Z}$ are independent normal random variables with mean 0 and variance 1.

The lemma above can be verified by M1S and M2S materials. We denote $J(k, 0) = Z_k$, $k \in \mathcal{D}$ and we define $J(k, n)$ recursively by using only Z_k. Note that Z_k is a normal random variable with mean 0 and variance 1 by Central Limit Theorem (another proof showing Z_k is normally distributed involves taking the limit of scaled symmetric random walk and using moment generating function).

[3]This idea of first defining BM on a dense subset of $\mathbb{R}$ before extending the definition (as well as the following arguments) also comes from an argument given by Lawler in [1].

Then : $J(2k-1, n+1) = \frac{1}{\sqrt{2}}J(k,n) + \frac{1}{\sqrt{2}}\ Z_{\frac{2k+1}{2^n+1}}$, and $J(2k, n+1) = \frac{1}{\sqrt{2}}J(k,n)$ - $\frac{1}{\sqrt{2}}\ Z_{\frac{2k+1}{2^n+1}}$
We proceed recursively, then we can see $J(k,n)$, are normal random variables with mean 0 and variance 1 for $k \geq 1$. Thus, we have verified the existence of Brownian motion over dyadic rationals other than intuitively taking the limit of one-dimensional scaled symmetric random walk.
To extend domain of BM from dyadic rationals to $\mathbb{R}$, we need to show W_t is a continuous function, it suffices to show that W_t is uniformly continuous on each compact as $\mathcal{D}$ is dense in $\mathbb{R}$. Now we prove the uniform continuity.

Proof: We do the proof on interval [0,1], other intervals can be shown in a similar way.

Let $\bar{K_n} = \sup\{|W_s - W_t| : 0 \leqslant s,t \leqslant 1, |s-t| \leqslant 2^{-n}, s,t \in D\}$

To show W_t is uniformly continuous, it suffices to show $\bar{K_n} \to 0$ as $n \to \infty$.

Consider $\mathrm{K}_n = \max\ sup\{|\mathrm{W}_q - W_{\frac{k-1}{2^n}}|$: q$\in \mathcal{D}$, $\frac{k-1}{2^n} \leqslant q \leqslant \frac{k}{2^n}, k = 1,2,3...\}$

We now use triangular inequality to deduce that $\bar{K_n} \leqslant 3\mathrm{K}_n$

Take s $\in [\frac{k-1}{2^n}, \frac{k}{2^n}], t \in [\frac{k}{2^n}, \frac{k+1}{2^n}]$, then $|s-t| \leqslant 2^{-n}$.

$\bar{K_n} \leqslant$ sup $(|W_s - W_{\frac{k-1}{2^n}}| + |W_{\frac{k}{2^n}} - W_{\frac{k-1}{2^n}}| + |W_t - W_{\frac{k}{2^n}}|) \leqslant 3K_n$

Then, it suffices to show $K_n \to 0$ as $n \to \infty$.

Now we can show $\sum \mathbb{P}[K_n \geqslant 2^{-n/2}] < \infty$

By Borel- Cantelli lemma and implication of a lemma stated later, we can show $K_n < 2\sqrt{n}2^{-n}$ with probability one. ie: $K_n \to 0$ as $n \to \infty$. ■

1.2.2 Properties of Brownian Motion

Non-differentiability

For each $t \geqslant 0$, W_t is not differentiable at t with probability 1. It is unusual to find a continuous function that is nowhere differentiable, but this function amazingly satisfies such properties. This is important because normal calculus doesn't work anymore (we will delve into more details in section **3.2**.), we need a different method for analyzing this situation. The following proof is inspired from Choongbum Lee's lectures on stochastic processes at MIT in 2013 [3].

To prove this proposition, we need a lemma first: Define M(t) $= \max_{s \leq t} W_s$. Then:

Lemma 1.2

$\mathbb{P}(M(t) > a) = 2\mathbb{P}(W_t > a)\ \forall t > 0$ and $\forall a > 0$

Proof: Indeed, define $T = \min\{t : W_t = a\}$

$\mathbb{P}(W_t - W_T > 0 \,|\, T < t) = \mathbb{P}(W_t - W_T < 0 \,|\, T < t)$ as, after T, the Brownian motion has equal probability of moving is both directions. So:

$$\mathbb{P}(M(t) > a) = \mathbb{P}(T < a) = \mathbb{P}(W_t - W_T > 0 \cap T < t) + \mathbb{P}(W_t - W_T < 0 \cap T < t)$$
$$= 2\mathbb{P}(W_t > a \cap T < t) = 2\mathbb{P}(W_t > a), \text{ as stated.}$$

We can now proceed with proving that W_t is nowhere differentiable:

Proof: Assume it were differentiable at some $t_0 \in R$, then $\exists A \in \mathbb{R}$ such that $\frac{dW_{t_0}}{dt} = A$

Then $\forall \delta > 0 : |W_{t_0+\delta} - W_{t_0}| = A\delta + o(\delta)$, from the definition of derivatives.

Hence, $W_{t_0+\delta} - W_{t_0} \leq A\delta$. And note that from definition, $W_{t_0+\delta} - W_{t_0} \sim \mathcal{N}(0, \delta)$; and this difference is also a Brownian motion. Set $M(\delta) := \max\limits_{s \leq t}(W_{t_0+\delta} - W_{t_0})$.

In particular, as $\delta \longrightarrow 0$: $\mathbb{P}(M(\delta) > A\delta) = 2\mathbb{P}(W_{t_0+\delta} > A\delta) \longrightarrow 1$ as $\delta \longrightarrow 0$.

So in the limit $M(\delta) > A\delta$ with probability 1. However, we saw earlier that $M(\delta) \leq A\delta$. This is a contradiction. Hence, W_t is nowhere differentiable. ■

Let us now study one last very important property of Brownian motion. Let W_t be a Brownian motion defined on $[0, T]$ $(T \in \mathbb{R})$.

Quadratic Variation Theorem

$$\lim_{n \to \infty} \sum_{i=1}^{n} (W_{\frac{iT}{n}} - W_{\frac{(i-1)T}{n}})^2 = T$$

It is a unique property for Brownian motion since this quantity is always zero for a smooth function, as we will see in section **3.2**. It is often stated as $(\mathrm{d}W_t)^2 = dT$, where $\mathrm{d}W_t$ represents the infinitesimal movement of W_t - although it is not differentiable. We will come back to this a bit later in the project.

Proof: Denote $t_i = \frac{iT}{n}$

$$\sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2$$

$$= \sum_{i=0}^{n-1} (X_i)^2 \ , \ \text{ where } \mathrm{X}_i \sim \mathcal{N}(0, \tfrac{T}{n}) \ (E(X_i)^2 = \tfrac{T}{n})$$

$$= \sum_{i=0}^{n-1} Y_i, \ \text{ where } \ Y_i = X_i^2$$

$$= n\overline{Y_n} \longrightarrow n \times \frac{T}{n} = T \ \text{ as } \ n \longrightarrow +\infty$$

Where $\overline{Y_n} := \frac{1}{n}\sum_{i=0}^{n-1} Y_i$, and the last equality is obtained from the strong law of large numbers.

1.2.3 Derivation of the continuous heat equation

Let us now apply this knowledge on random walks and Brownian motion to derive an important equation of particle diffusion, which we will later apply to financial models. We will now derive the heat equation $\dfrac{\partial u}{\partial t} = \dfrac{1}{2}\dfrac{\partial^2 u}{\partial x^2}$.

Let $W_t, t \geq 0$ be a standard Brownian motion (Wiener process) such that $W_0 = 0$.

Let $f : \mathbb{R} \to \mathbb{R}$ be $\mathcal{C}^2$ (i.e. f is twice differentiable and its second derivative is continuous).

Let $u : \mathbb{R} \times [0, \infty) \to \mathbb{R} : (x, t) \mapsto u(x, t) = \mathbb{E}(f(W_t + x))$.

Note that $\forall x \in \mathbb{R}, u(x, 0) = \mathbb{E}(f(W_0 + x)) = \mathbb{E}(f(x)) = f(x)$.

Also, u is continuous at $(x,0), \forall x \in \mathbb{R}$.

Claim: $\dfrac{\partial u}{\partial t} = \dfrac{1}{2}\dfrac{\partial^2 u}{\partial x^2}$

Proof: Fix $t \in \mathbb{R}_+, s > 0, x \in \mathbb{R}$. Consider:

$$\begin{aligned}\frac{u(x,t+s)-u(x,t)}{s} &= \frac{\mathbb{E}(f(W_{t+s}+x)) - E(f(W_t+x))}{s} \\ &= \frac{\mathbb{E}(f(W_{t+s}+x) - f(W_t+x))}{s}\end{aligned}$$

Now, $f(W_{t+s}+x) - f(W_t+x) = f((W_{t+s}-W_t)+W_t+x) - f(W_t+x)$.

Let $A := W_{t+s} - W_t$. Then $f(W_{t+s}+x) - f(W_t+x) = f((W_t+x)+A) - f(W_t+x)$. Taking a Taylor expansion of f about (W_t+x) (note that we can do this up to a term x^2 as f is C^2),

$$f(W_t+x+A) = f(W_t+x) + Af'(W_t+x) + \frac{f''(W_t+x)}{2!}A^2 + O(A^3)$$

$$\Rightarrow \frac{u(x,t+s)-u(x,t)}{s} = \frac{1}{s}\mathbb{E}\left(\left[f(W_t+x) + Af'(W_t+x) + \frac{f''(W_t+x)}{2!}A^2 + O(A^3)\right]\right)$$

where $O(A^3)$ denotes terms in the Taylor expansion which have order of at least 3. We note that $A = (W_{t+s} - W_t) \sim \mathcal{N}(0,s)$ (using Lemma 1.1). Consider

$$\begin{aligned}\mathbb{E}\left(Af'(W_t+x)\right) &= \mathbb{E}\left((W_{t+s}-W_t)f'(W_t+x)\right) \\ &= \mathbb{E}\left(\mathbb{E}(W_{t+s}-W_t \mid W_t)f'(W_t+x)\right)\end{aligned}$$

where the second equality follows due to the fact that $f'(W_t+x)$ is known, given (conditional on) W_t. Now, we use the property that $W_{t+s} - W_t$ is independent of W_t, and the fact that $W_{t+s} - W_t = A \sim \mathcal{N}(0,s)$. We deduce that $\mathbb{E}(W_{t+s} - W_t \mid W_t) = 0$, so $\mathbb{E}\left(Af'(W_t + x \mid W_0 = 0\right) = \mathbb{E}\left(0(f'(W_t+x) \mid W_0 = 0\right) = 0$. Similarly,

$$\begin{aligned}\mathbb{E}\left(\frac{A^2}{2!}f''(W_t+x)\right) &= \mathbb{E}\left(\mathbb{E}\left(\frac{A^2}{2!}f''(W_t+x) \mid W_t\right)\right) \\ &= \mathbb{E}\left(\frac{1}{2}\mathbb{E}\left(A^2 \mid W_t\right)f''(W_t+x)\right)\end{aligned}$$

where again the last equality follows due to the fact that $f''(W_t+x)$ is known, given (conditional on) W_t. Since A is independent of W_t, $\mathbb{E}(A^2) = (\mathbb{E}(A))^2 + \mathrm{Var}(A)$. Because $\mathrm{Var}(A) = s$, we have $\mathbb{E}\left(\frac{A^2}{2!}f''(W_t+x)\right) = \frac{1}{2}s\mathbb{E}\left(f''(W_t+x)\right)$.

Similarly, we find, using knowledge about finding the moments of normally distributed random variables from M1S and M2S, that for $r \in \mathbb{Z}_{\geq 3}$:

$$\mathbb{E}(A^r) = \begin{cases} 0 & \text{if } r \text{ is odd} \\ s^{\frac{r}{2}}(r-1)...(1) & \text{if } r \text{ is even}\end{cases}$$

We find:

$$\frac{u(x,t+s)-u(x,t)}{s} = \frac{1}{s}\left[\frac{1}{2}s\mathbb{E}\left(f''(W_t+x)\right) + O(s^2)\right] \quad = \frac{1}{2}\mathbb{E}\left(f''(W_t+x)\right) + O(s)$$

So, $\lim_{s\to 0}\left(\frac{u(x,t+s)-u(x,t)}{s}\right)$ exists and is equal to $\frac{1}{2}\mathbb{E}\left(f''(W_t+x)\right)$, since $O(s)\sim 0$.

However, $\frac{\partial u}{\partial x}=\lim_{s\to 0}\left(\frac{u(x,t+s)-u(x,t)}{s}\right)$.

Moreover, $\frac{1}{2}\mathbb{E}\left(f''(W_t+x\mid W_0=0)\right)=\frac{1}{2}\frac{\partial^2\left[\mathbb{E}\left(f(W_t+x)\mid W_0=0\right)\right]}{\partial x^2}=\frac{1}{2}\frac{\partial^2 u}{\partial x^2}$

$$\therefore\ \boxed{\frac{\partial u}{\partial t}=\frac{1}{2}\frac{\partial^2 u}{\partial x^2}} \qquad \blacksquare$$

We now derive an explicit formula for u.

We have that $\forall x\in\mathbb{R},\forall t\in\mathbb{R}_+,u(x,t)=\mathbb{E}\left(f(W_t+x)\right)$.

$\Rightarrow \mathbb{E}\left(f(W_t+x)\right)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi t}}f(w+x)e^{-w^2/2t}\,\mathrm{d}w$, by definition of $W_t\sim N(0,t)$.

For slight convenience, using the substitution $y=w+x$, we find that

$$\mathbb{E}\left(f(W_t+x)\right)=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi t}}f(w+x)e^{-w^2/2t}\,\mathrm{d}w=\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi t}}f(y)e^{-(y-x)^2/2t}\,\mathrm{d}y \tag{1}$$

1.2.4 Motivation for Geometric Brownian Motion

Since Brownian motion can take negative values, when applying this to model stock prices, it becomes questionable. Thus, people often use the exponential function to transform Brownian motion into Geometric Brownian motion (GBM). This non-negative property hence makes it a really useful tool in financial analysis and modelling. Stock prices at time t(often denoted by S_t) can hence be described as $S_t=S_0e^{W_t}$.[4] We will get into more details about this in part **2**.

[4] As Choongbum Lee said in a comment of his lecture, this does not exactly correspond to taking the pointwise exponential of the function W_t. But a lot of theory is required to formalize it, and it is safe to consider it that way in the scope of our project

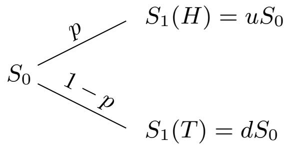
2 Applications to finance

2.1 Introduction and derivation of Black-Scholes

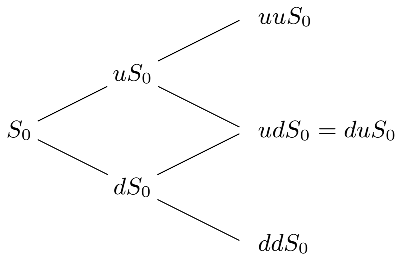
We have seen how Brownian motion can be constructed by taking the length of a time step of a random walk to 0. In this section we will make another construction to derive the Black-Scholes equation using a binomial model, where we assume stock prices to be modelled as Geometric Brownian Motion. We will then solve the Black-Scholes differential equation by converting it into the heat equation.

2.1.1 Introduction to the Binomial model

Consider at time $t = 0$ the price of a stock is S_0 and flip a coin that has a probability p of getting heads. If the coin lands on heads, the price of the stock at $t = 1$ is $S_1 = uS_0$ and if the coin lands on tails then the stock at time $t = 1$ is $S_1 = dS_0$, where $0 < d < u$. The diagram below summarises this:



This diagram represents a 1-step binomial model [4]. Now consider repeating this N times, the tree diagram will start like the following:



Hence after N time steps, $S_t = S_0 u^{H_n} d^{T_n}$, where H_n is the number of heads and T_n is the number of tails and $H_n + T_n = N$ (where we recall the coin flipping analogy from above).

2.1.2 Arbitrage

The concept of arbitrage is very important in finance. In [4] it is defined as "a trading strategy that begins with no money, has zero probability of losing money, and has a positive probability of making money", essentially you can make money out of nothing.

Let us introduce an interest rate $r > 0$, into the 1-step Binomial model. It can be shown that in order to get an arbitrage free model, we need

$$0 < d < 1 + r < u \tag{2}$$

Assume not, then we can see it leaves us with two cases: $u > d \geq 1 + r$ or $1 + r \leq u < d$. In the first case, an investor can borrow some money, say X, and put it into the stocks market, and after one time step, can guarantee a return of at least $(1+r)X$, after returning X and a bit, the investor has made money out of nothing. In the second case, the owner could sell the stock and guarantee after a time step the stocks would be worth less so that he could buy back the stocks at a lower price. Both cases lead to arbitrage.

There is one more condition we need to assume for an arbitrage free model, it is that the probabilities in the binomial model are governed by $\tilde{p} = \dfrac{1 + r - d}{u - d}$ and $\tilde{q} = \dfrac{u - 1 - r}{u - d}$ [5]. Note that due to the equation above $\tilde{p}$, $\tilde{q} < 1$ and their sum is 1.

2.1.3 Derivative Securities

So far we have seen two types of securities:

1. Riskless bond B, with an interest rate $r > 0$

$$B_{t+1} = (1 + r)B_t$$

   for $t = 0, 1, \ldots.$

2. Risky asset (stocks) S,

$$S_{t+1} = \begin{cases} uS_t, & \text{with probability } p \\ dS_t, & \text{with probability } 1 - p \end{cases}$$

   for $t = 0, 1, \ldots.$

Let us introduce some terminology. A European *option* is a financial tool that gives someone the right (but not the obligation) to buy or sell an asset at a specific date and price.[5] A *call* option gives the right to buy, whereas a *put* option gives the right to sell and a *strike price* is the agreed exchange price between the buyer and seller [6].

Hence, a European Call Option with a strike price K has a payoff given by $[S_t - K]^+ = \max\{S_t - K, 0\}$ and a European Put Option with strike price K has a payoff given by the formula $[K - S_t]^+ = \max\{K - S_t, 0\}$.

2.1.4 Black-Scholes equations

There are a number of assumptions that we need to make in order to work with the Black-Scholes model [4]: shares of stock can be subdivided for sale or purchase (always satisfied in reality), interest rate for investing is the same as for borrowing (mostly satisfied in reality), and is **risk-free**, purchase price of stock is the same as the selling price (not satisfied in reality) and stock price

[5] An American option is slightly different in that it allows its holder to buy (or sell) the option at any time before the expiration date, instead of only at this specific date.

follows a geometric brownian motion (unlikely to be exactly satisfied in reality).

Now we want to build an appropriate binomial model that converges to the Black-Scholes formula. For N steps between time 0 and τ, where a time step is given by $\dfrac{\tau}{N}$, given an interest rate r, we take $u = 1 + \dfrac{r\tau}{N} + \dfrac{\sigma\sqrt{\tau}}{\sqrt{N}}$ and $d = 1 + \dfrac{r\tau}{N} - \dfrac{\sigma\sqrt{\tau}}{\sqrt{N}}$. Note this gives $\tilde{p} + \tilde{q} = 1$.

We want to compute the limit as $N \to \infty$ of the put price in a binomial model which is given by

$$P_0 = \frac{1}{\left(1 + \frac{r\tau}{N}\right)^N}[(K - S_N)^+]$$

Claim: $\lim\limits_{N\to+\infty} P_0$ gives the Black-Scholes price of a put with strike price K and expiration time τ on a stock with volatility σ. [5]

Sketch of the proof:

$$\begin{aligned}
\text{Consider } \log(S_N) &= \log(S_0) + H_n \log u + T_n \log d \\
&= \log(S_0) + H_n \log\left(1 + \frac{r\tau}{N} + \frac{\sigma\sqrt{\tau}}{\sqrt{N}}\right) + T_n \log\left(1 + \frac{r\tau}{N} - \frac{\sigma\sqrt{\tau}}{\sqrt{N}}\right) \\
&(\text{Since } \log(1+x) = x - \frac{1}{2}x^2 + O(x^3)) \\
&= \log(S_0) + H_n\left(\frac{r\tau}{N} + \frac{\sigma\sqrt{\tau}}{\sqrt{N}} + \frac{\sigma^2\tau}{2N}\right) + T_n\left(\frac{r\tau}{N} - \frac{\sigma\sqrt{\tau}}{\sqrt{N}} - \frac{\sigma^2\tau}{2N}\right) + O\left(\frac{1}{N\sqrt{N}}\right) \\
&= \log(S_0) + \sigma\sqrt{\tau}\frac{(H_n - T_n)}{\sqrt{N}} + \left(r - \frac{1}{2}\sigma^2\right)\tau\left(\frac{H_n + T_n}{N}\right) + O\left(\frac{1}{N\sqrt{N}}\right)
\end{aligned}$$

If we associate X_i with scoring $+1$ if we get a H and -1 if we get a T on the i^{th} toss.

Then $H_n - T_n = \displaystyle\sum_{i=1}^{N} X_i$

Hence using the Generalised Central Limit Theorem, with $\gamma = \sigma\sqrt{\tau}$ and

$Y_N = \log(S_0) + \left(r - \frac{1}{2}\sigma^2\right)\tau + O\left(\frac{1}{N\sqrt{N}}\right)$ we get:

$$\lim_{N\to+\infty} P_0 = \frac{e^{-r\tau}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}\left(K - \exp\left\{x\sigma\sqrt{\tau} + \log(S_0) + \left(r - \frac{1}{2}\sigma^2\right)\tau\right\}\right)^+ e^{\frac{-x^2}{2}}\mathrm{d}x$$

Now we want to find out the values of x for which the integral is non-zero:

$$K - \exp\left\{x\sigma\sqrt{\tau} + \log(S_0) + \left(r - \frac{1}{2}\sigma^2\right)\tau\right\} > 0$$

$$\implies x < -\frac{1}{\sigma\sqrt{\tau}}\left[\log\frac{S_0}{K} + \left(r - \frac{1}{2}\sigma^2\right)\tau\right] = -d_2$$

Hence we get:

$$\lim_{N\to+\infty} P_0 = e^{-r\tau}K\int_{-\infty}^{-d_2}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}\mathrm{d}x - \frac{e^{-r\tau}}{\sqrt{2\pi}}\int_{-\infty}^{-d_2}\exp\left\{x\sigma\sqrt{\tau} + \log(S_0) + \left(r - \frac{1}{2}\sigma^2\right)\tau\right\}e^{\frac{-x^2}{2}}\mathrm{d}x$$

$$= e^{-r\tau}K\int_{-\infty}^{-d_2}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}\mathrm{d}x - S_0\int_{-\infty}^{-d_2}\frac{1}{\sqrt{2\pi}}e^{\frac{-(x-\sigma\sqrt{\tau})^2}{2}}\mathrm{d}x$$

Making a substitution $y = x - \sigma\sqrt{\tau}$ gives

$$d_1 = -d_2 - \sigma\sqrt{\tau} = \frac{1}{\sigma\sqrt{\tau}}\left[\log\frac{S_0}{K} + \left(r + \frac{1}{2}\sigma^2\right)\tau\right]$$

Hence we get:

$$\lim_{N\to+\infty} P_0 = e^{-r\tau}K\int_{-\infty}^{-d_2}\frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}\mathrm{d}x - S_0\int_{-\infty}^{-d_1}\frac{1}{\sqrt{2\pi}}e^{\frac{-y^2}{2}}\mathrm{d}y$$

$$= e^{-r\tau}K\Phi(-d_2) - S_0\Phi(-d_1)$$

where Φ is the cumulative distribution function (cdf) of the normal distribution.

In addition, there is a relationship between P_0 and C_0, called the put-call parity [6]. It states

$$S_0 + P_0 - C_0 = Ke^{r(T-t)}$$

where we have taken $\tau = T - t$. Using this we can show that $\lim_{N\to+\infty} C_0$ gives the Black-Scholes price of a call with strike price K, expiration time τ on a stock with volatility σ

$$\begin{aligned}\lim_{N\to+\infty} C_0 &= \lim_{N\to+\infty}(P_0 + S_0 - Ke^{r\tau})\\ &= e^{r\tau}K\Phi(-d_2) - S_0\Phi(-d_1) + S_0 - Ke^{-r\tau}\\ &= S_0(1-\Phi(-d_1)) - e^{r\tau}K(1-\Phi(-d_2))\\ &= S_0(\Phi(d_1)) - e^{r\tau}K(\Phi(d_2))\end{aligned}$$

This is often known as the Black-Scholes formula. And we will later see that it can also be found as a solution to the Black-Scholes equation.

2.1.5 Derivation of the Black-Scholes equation

Firstly, we need a portfolio[6] (ϕ_t, ψ_t) and, as shown in [8], the value of the portfolio is given by $V_t = \phi_t S_t + \psi_t rP\mathrm{d}t$, where ϕ_t and ψ_t are the amounts of share and cash at time t, respectively.

In addition to this, we will need to use a general version Ito's Lemma which is stated here but will be introduced in more detail later:

General Itô's Lemma

$$\boxed{\mathrm{df} = \left(\frac{\partial f}{\partial t} + \mu S_t\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2}\right)\mathrm{d}t + \sigma S_t\frac{\partial f}{\partial x}\mathrm{d}W_t}$$

This derivation is similar to the one in [8]. Let $f = V_t$ and $\mathrm{d}x = \mathrm{d}S_t$,

$$\frac{\partial f}{\partial t} = \phi_t rP\mathrm{d}t,\ \frac{\partial f}{\partial x} = \phi_t,\ \frac{\partial^2 f}{\partial x^2} = 0$$

[6]A portfolio is defined as a collection of assets [7]

Hence, putting this into Itô's Lemma we get:

$$\mathrm{d}V_t = (\psi_t rP + \mu S_t \phi_t)\,\mathrm{d}t + \sigma S_t \phi_t \mathrm{d}W_t$$

Comparing the coefficients of $\mathrm{d}W_t$ of Itô's Lemma and the equation above, we get: $\frac{\partial f}{\partial x} = \phi_t$ and using the definition of V_t, where $P = rPdt$ and substituting in ϕ_t gives $\psi_t = \frac{1}{P}\left[f - \frac{\partial f}{\partial x}S_t\right]$. Now compare the coefficients of dt of the equation above and Itô's Lemma and substitute in ϕ_t and ψ_t we get

$$\frac{\partial f}{\partial t} + \mu S_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} = \frac{1}{P}\left[f - \frac{\partial f}{\partial x}S_t\right] rP + \mu S_t \frac{\partial f}{\partial x}$$

Simplifying this gives the
Black-Scholes differential equation

$$\boxed{\frac{\partial f}{\partial t} + rS_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2} = rf}$$

2.2 Solving the Black-Scholes Equation

We can solve the Black-Scholes equation for certain conditions. We will solve the equation assuming we have stocks with a constant volatility (σ) and a constant (disregarding risk) interest rate (r), where we have the Black-Scholes equation $\dfrac{\partial C}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 C}{\partial t^2} + rS\dfrac{\partial C}{\partial S} - rC = 0$, with $C : \mathbb{R}_+{}^7 \times [0,T] \to \mathbb{R}_+$. Furthermore we have the initial condition $C(0,t) = 0, \forall t \in [0,T]$, the condition $C(S,T) = \max(S - E, 0)$ and the boundary condition $\lim\limits_{S\to\infty}((C(S,t)) = S$.
We will solve the equation by 'reducing' it to the heat equation and then using methods known to us from our prior studies in mathematics. The complete derivation can be found in 'An Introduction to Financial Derivatives: A Student Introduction' [9], but we will highlight the crucial steps in this derivation.
NB: *There are several other ways to solve the Black-Scholes equation with the aforementioned conditions, such as through the method of 'separation of variables'.*

2.2.1 Transformation to the heat equation

1. We shall commence by implementing the following changes of variables, which we do in order to remove 'products of powers' of the partial derivatives of C in the Black-Scholes differential equation.

   Let

$$S = Ee^x, \qquad x \in \mathbb{R};$$

$$t = T - \frac{2\tau}{\sigma^2}, \qquad t \in [0,T];$$

$$C(S,t) = Ev(x,\tau)$$

[7] $\mathbb{R}_+ = [0,\infty)$

(where $C : \mathbb{R}_+ \times [0,T] \to \mathbb{R}_+, \quad V : \mathbb{R} \times [0,T] \to \mathbb{R}_+$).

Then, consider the boundary condition $C(S,T) = \max(S-E,0)$, [8] $\forall S \in \mathbb{R}_+$.

When $t = T$, we have that $T = T - \frac{2\tau}{\sigma^2}$ (substituting T into our expression for t)

$\Rightarrow \tau = 0$

Using the expression for C above, and for S, we now have the initial condition:

$$Ev(x,0) = \max(Ee^x - E, 0) \Rightarrow v(x,0) = \max(e^x - 1, 0) \tag{3}$$

In addition, we obtain:

$$\frac{\partial C}{\partial t} = -\frac{E\sigma^2}{2}\frac{\partial v}{\partial \tau};$$

$$\frac{\partial C}{\partial S} = \frac{E}{S}\frac{\partial v}{\partial x};$$

$$\frac{\partial^2 C}{\partial S^2} = \frac{E}{s^2}(\frac{\partial^2 v}{\partial x^2} - \frac{\partial v}{\partial x})$$

.

(Note that we do not substitute for 'S' and 'S^2' in the expressions above as these will 'cancel' out in the next few lines of working.)

Substituting the expressions for $\frac{\partial C}{\partial t}$, $\frac{\partial C}{\partial S}$ and $\frac{\partial^2 C}{\partial S^2}$ obtained above, we obtain the expression:

$-\frac{E\sigma^2}{2}\frac{\partial v}{\partial \tau} + \frac{E\sigma^2}{2}(\frac{\partial^2 v}{\partial x^2} - \frac{\partial v}{\partial x}) + rE\frac{\partial v}{\partial x} - rEv = 0$

$\Rightarrow \frac{\partial^2 v}{\partial x^2} - \frac{\partial v}{\partial x} + \frac{r}{\frac{\sigma^2}{2}}\frac{\partial v}{\partial x} - \frac{r}{\frac{\sigma^2}{2}}v = \frac{\partial v}{\partial \tau}$.

For convenience, let us denote $m := \frac{r}{\sigma^2/2}$.

We can thus rewrite the equation as $\frac{\partial^2 v}{\partial x^2} + (m-1)\frac{\partial v}{\partial x} - mv = \frac{\partial v}{\partial \tau}$

2. We seek to 'eliminate' any terms on the left-hand side of the equation which are not second derivatives. We can therefore make a substitution for v.

   Ansatz: Let $v(x,\tau) = e^{\alpha x + \beta\tau} g(x,\tau)$, where $g : \mathbb{R} \times [0,T] \to \mathbb{R}_+; \alpha, \beta \in \mathbb{R}$.

   Then, the initial condition (3) implies $v(x,0) = e^{\alpha x} g(x,0) = \max(e^x - 1, 0)$

$$\Rightarrow g(x,0) = \max(e^{(1-\alpha)x} - e^{-\alpha x}, 0)$$

   We also have:

$$\frac{\partial^2 g}{\partial x^2} + 2\alpha\frac{\partial g}{\partial x} + \alpha^2 g + (m-1)(\alpha g + \frac{\partial g}{\partial x}) - mg = \beta g + \frac{\partial g}{\partial \tau}$$

$$\Rightarrow \frac{\partial^2 g}{\partial x^2} + (2\alpha + m - 1)\frac{\partial g}{\partial x} + (\alpha^2 + \alpha(m-1) - m - \beta)g = \frac{\partial g}{\partial \tau}$$

[8] Here, $\max(a,b)$ denotes the maximum of a and b.

We set:

$$2\alpha + m - 1 = 0 \tag{4}$$

$$\alpha^2 + \alpha(m-1) - m - \beta = 0 \tag{5}$$

so that the terms containing $\dfrac{\partial g}{\partial x}$ and g are eliminated.
The condition (4) implies $\alpha = \dfrac{1-m}{2}$, and we find that:

$$\beta = \frac{1-m}{2}^2 + \frac{1-m}{2}(m-1) - m \Rightarrow \beta = -\frac{(k+1)^2}{4}$$

.

We re-express the initial condition for g (when $\tau = 0$) as:

$$g(x,0) = \max(e^{\frac{m+1}{2}x} - e^{\frac{m-1}{2}}x, 0)$$

We have the equation $\dfrac{\partial^2 g}{\partial x^2} = \dfrac{\partial g}{\partial \tau}$.
We make one final substitution $u(x,t) = g(x, \frac{1}{2}t)$ and find that $\dfrac{\partial^2 u}{\partial x^2} = \dfrac{\partial^2 g}{\partial x^2}$, $\dfrac{\partial g}{\partial t} = 2\dfrac{\partial u}{\partial \tau}$, and $u(x,0) = g(x,0) = \max(e^{\frac{m+1}{2}x} - e^{\frac{m-1}{2}}x, 0)$

$\therefore$ We find that

$$\frac{1}{2}\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial \tau}$$

and have thus reduced the Black-Scholes equation to the heat equation from Subsection 1.2.1.

2.2.2 Solving Black-Scholes

We now have the partial differential equation

$$\frac{1}{2}\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial \tau}$$

with the initial condition $u(x,0) = f(x)$, where

$$f(x) = \begin{cases} e^{\frac{(m+1)x}{2}} - e^{\frac{(m-1)x}{2}}, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

.

We then use (1) (from Subsection **1.2.1**) to find that [9]

$$u(x,\tau) = \int_{-\infty}^{\infty} f(y)\frac{1}{\sqrt{2\pi\tau}}e^{-\frac{(y-x)^2}{2\tau}}\,\mathrm{d}y$$

$$\Rightarrow g(x,\tau) = u(x,2\tau) = \int_{-\infty}^{\infty} f(y)\frac{1}{2\sqrt{\pi\tau}}e^{-\frac{(y-x)^2}{4\tau}}\,\mathrm{d}y$$

[9] What we obtain in this line is a convolution of two functions, and this result would also arise if we had sought to solve the partial differential equation for u and hence g using Fourier transforms (which we have learnt about in our studies of mathematics).

Then:

$$g(x,\tau)=\int_0^\infty (e^{\frac{(m+1)y}{2}}-e^{\frac{(m-1)y}{2}})\frac{1}{2\sqrt{\pi\tau}}e^{-\frac{(y-x)^2}{4\tau}}\,\mathrm{d}y$$

For ease of manipulation, let $\gamma_1=\dfrac{m+1}{2}$ and $\gamma_2=\dfrac{m-1}{2}$. Then we have that:

$$\begin{aligned}\gamma_1 y-\frac{1}{4\tau}(y-x)^2 &= -\frac{1}{4\tau}(y^2-2xy-4\tau\gamma_1 y+x^2)\\ &= -\frac{1}{4\tau}((y-(x+2\gamma_1\tau))^2-4\tau(\gamma_1 x+\gamma_1^2\tau))\\ &= -\frac{1}{4\tau}(y-(x+2\gamma_1\tau))^2+(\gamma_1 x+\gamma_1^2\tau)\end{aligned}$$

Similarly,

$$\gamma_2 y-\frac{1}{4\tau}(y-x)^2=-\frac{1}{4\tau}(y-(x+2\gamma_2\tau))^2+(\gamma_2 x+\gamma_2^2\tau)$$

We obtain the following:

$$g(x,\tau)=e^{\gamma_1 x+\gamma_1^2\tau}\int_0^\infty \frac{1}{\sqrt{2\pi}\sqrt{2\tau}}e^{-\frac{1}{2(2\tau)}(y-(x+2\gamma_1\tau))^2}\,\mathrm{d}y-e^{\gamma_2 x+\gamma_2^2\tau}\int_0^\infty \frac{1}{\sqrt{2\pi}\sqrt{2\tau}}e^{-\frac{1}{2(2\tau)}(y-(x+2\gamma_2\tau))^2}\,\mathrm{d}y$$

Now, the integrand of the first integral is the probability density function for a normally distributed random variable Y_1 with expectation $(x+2\gamma_1\tau)$ and variance 2τ.
In the same manner, the integrand of the second integral is the probability density function for a normally distributed random variable Y_2 with expectation $(x+2\gamma_2\tau)$ and variance 2τ. Therefore,

$$g(x,\tau)=e^{\gamma_1 x+\gamma_1^2\tau}\mathbb{P}(Y_1\geq 0)-e^{\gamma_2 x+\gamma_2^2\tau}\mathbb{P}(Y_2\geq 0)$$

We know from first and second year mathematics that

$$\mathbb{P}(Y_1\geq 0)=\mathbb{P}\left(Z\geq\frac{-(x+2\gamma_1\tau)}{\sqrt{2\tau}}\right)=\Phi\left(\frac{x+2\gamma_1\tau}{\sqrt{2\tau}}\right),$$

where $Z\sim N(0,1)$ denotes the standard normal distribution and Φ is its cumulative distribution function.
We also have that

$$\mathbb{P}(Y_2\geq 0)=\Phi\left(\frac{x+2\gamma_2\tau}{\sqrt{2\tau}}\right).$$

It thus follows that $g(x,\tau)=e^{\gamma_1 x+\gamma_1^2\tau}\Phi\left(\frac{x+2\gamma_1\tau}{\sqrt{2\tau}}\right)-e^{\gamma_2 x+\gamma_2^2\tau}\Phi\left(\frac{x+2\gamma_2\tau}{\sqrt{2\tau}}\right)$.

Now, $v(x,\tau)=e^{\alpha x+\beta\tau}g(x,\tau)$, so

$$v(x,\tau)=e^{\alpha x+\beta\tau}\left(e^{\gamma_1 x+\gamma_1^2\tau}\Phi\left(\frac{x+2\gamma_1\tau}{\sqrt{2\tau}}\right)-e^{\gamma_2 x+\gamma_2^2\tau}\Phi\left(\frac{x+2\gamma_2\tau}{\sqrt{2\tau}}\right)\right).$$

$$\Rightarrow v(x,\tau)=e^x\Phi(d_1)-e^{-m\tau}\Phi(d_2),$$

where we have used the identities $\alpha=\frac{1-k}{2},\beta=-\frac{(1+k)^2}{4},\gamma_1=\frac{k+1}{2},\gamma_2=\frac{k-1}{2}$ and carried out some algebraic manipulation.
Recalling the substitution $S=Ee^x$ we implemented, $m=\frac{2r}{\sigma^2}$, $\tau=\frac{T-t}{2}\sigma^2$, and setting $d_1=$

$\dfrac{x+2\gamma_1\tau}{\sqrt{2\tau}}$ and $d_2 = \dfrac{x+2\gamma_2\tau}{\sqrt{2\tau}}$, we have that: $d_1 = \dfrac{ln(\frac{S}{E}) + 2(\frac{m+1}{2})\frac{T-t}{2}\sigma^2}{\sqrt{2\frac{(T-t)}{2}\sigma^2}}$

$\Rightarrow d_1 = \dfrac{ln(\frac{S}{E}) + r(T-t) + \frac{(T-t)}{2}\sigma^2}{\sigma\sqrt{T-t}} = \dfrac{ln(\frac{S}{E}) + (T-t)(r+\frac{\sigma^2}{2})}{\sigma\sqrt{T-t}}$.

Similarly, $d_2 = \dfrac{ln(\frac{S}{E}) + (T-t)(r-\frac{\sigma^2}{2})}{\sigma\sqrt{T-t}}$.

Moreover,

$$C(S,t) = Ee^x\Phi(d_1) - Ee^{-(\frac{2r}{\sigma^2})\frac{(T-t)}{2}\sigma^2}\Phi(d_2)$$

$$\underline{\therefore C(S,t) = S\Phi(d_1) - Ee^{-r(T-t)}\Phi(d_2)}$$

.

This is hence consistent with the Black-Scholes formula derived in part **2.1**.

2.2.3 Remarks on solving the Black-Scholes equation

1. Whilst we have been able to solve the Black-Scholes equation assuming that the volatility and interest rate of the stock in question are constant, these assumptions are not valid in practice. In fact, the volatility of a stock can change very frequently. The Black-Scholes equation cannot generally be solved for C (as a function of S and t) explicitly if the volatility and interest rate of the stock in question are not taken to be constant. One must thus use numerical methods to find a solution, such as the Crank-Nicholson scheme.

2. Unsurprisingly, the prices of call options predicted by the equation for C we have obtained from solving the Black-Scholes equation are usually different from the actual prices call options would take. This is due to several strong assumptions we have made, including the assumption that the volatility of the stock and interest rate are constant. This will be illustrated in the next section of this report.

3 To go further

The code and the datasets for each of the simulations can be found on the group's Github repository:

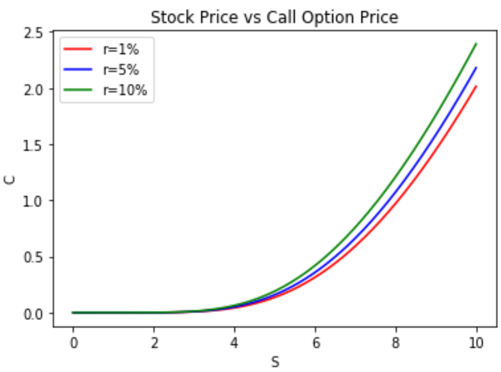
https://github.com/ArthurLimoge/M2R-Project

3.1 Simulations: testing the accuracy of our models

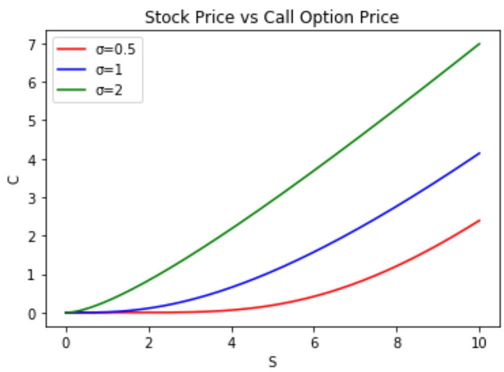
3.1.1 Solution curves for the Black-Scholes equation

We will now plot some curves for call option prices which are predicted by the Black-Scholes equation (with constant volatility and interest rate), whilst varying the volatility and interest rate. The plots in this subsection have been created using Python and the code used can be found on the group's GitHub repository, a link to which can be found at the start of this section.

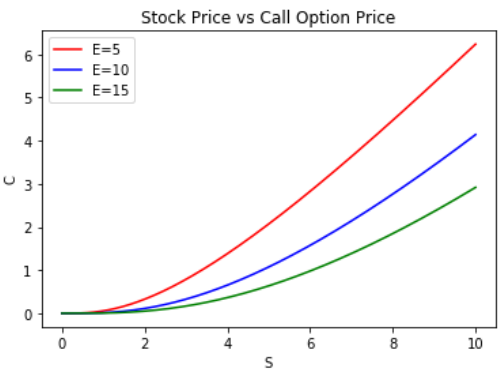
We first plot graphs (Figure 1) with the underlying stock price (S) of a give stock plotted against the value of the call option for this stock (C), whilst keeping all other parameters constant.



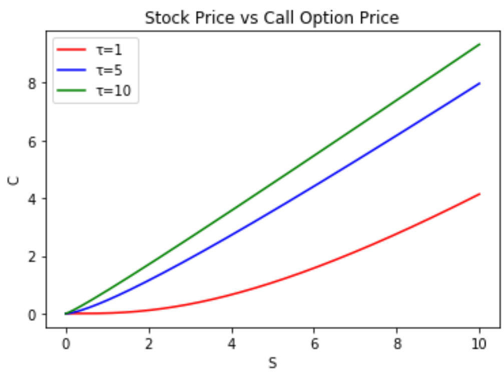
(a) Different values of r; $(T-t)=1, E=S=10, \sigma=0.5$



(b) Different values of σ; $T-t=1, E=S=10, r=0.1$



(c) Different values of E; $T-t=1, r=0.1, \sigma=1, S=10$



(d) Different values of $T-t$; $S=E=10, r=0.1, \sigma=1$

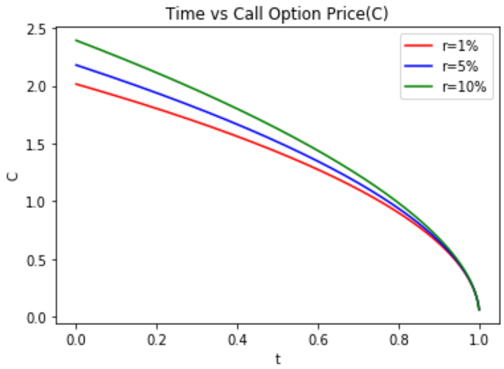
Figure 1: Underlying stock price against value of the call option, for different values of different parameters

The graphs all suggest that the price of the call option (C) increases as the value of the underlying stock of the call option (S) increases. Moreover, as S increases in size, the relationship between C and S is suggested to be 'approximately' linear. This is to be expected; indeed, the Black-Scholes equation has the boundary condition $\lim_{S\to\infty}(C(S,t)=S$.
We see from Figure 1 that as r,σ and $T-t$ increase, for any given stock and its underlying value, S, C (the value of the call option associated with the stock) increases. Considering the formula

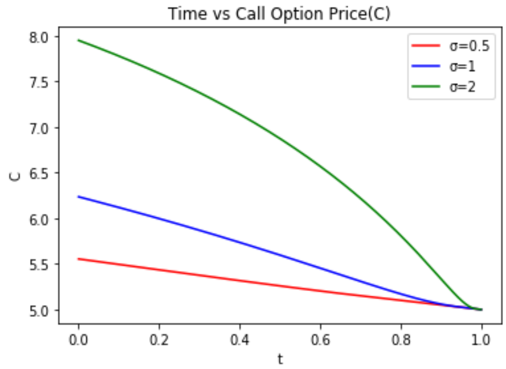
$C(S,t) = S\Phi(d_1) - Ee^{-r(T-t)}\Phi(d_2)$, we see that as each of r, σ and τ increases, $Ee^{-r(T-t)} \to 0$, whilst $0 \leq \Phi(d_1) \leq 1$ and $0 \leq \Phi(d_2) \leq 1$ (as Φ is a cdf). Therefore, $C(S,t)$ must increase as each of r, σ and $T-t$ increases.

On the other hand, as E increases, for any given S, C decreases, which is expected as $Ee^{-r(T-t)}$ increases, whilst $0 \leq \Phi(d_1) \leq 1$ and $0 \leq \Phi(d_2) \leq 1$.

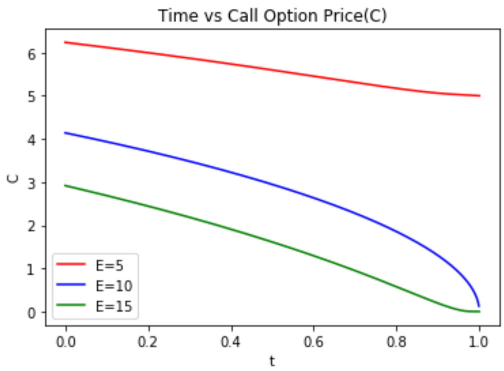
We now plot graphs (Figure 2) with time (t) plotted against the value of a call option C, whilst keeping all other parameters constant.



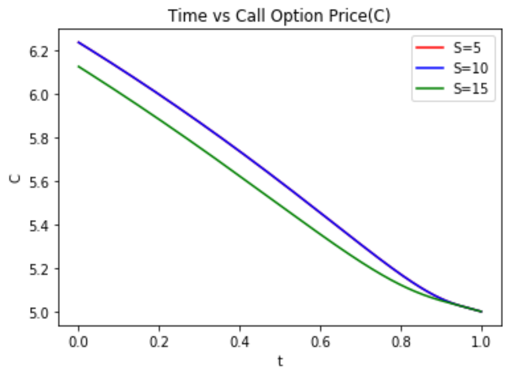
(a) Different values of r; $T = 1, E = S = 10, \sigma = 0.5$



(b) Different values of σ; $T = 1, E = S = 10, r = 0.1$



(c) Different values of E; $T = 1, r = 0.1, \sigma = 1, S = 10$



(d) Different values of S; $T = 5, E = 10, r = 0.1, \sigma = 1$

Figure 2: Time against value of the call option, for different values of different parameters

This time, the graphs suggest that the price of the call option decreases as time increases and goes to $C(S,1) = max(S-E,0)$. This is exactly what the boundary condition $C(S,T) = max(S-E,0)$ states.

3.1.2 Error Analysis

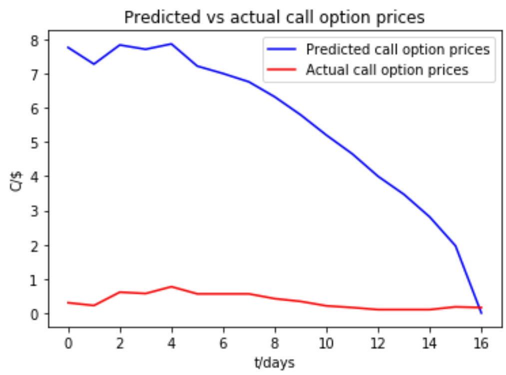
Whilst the Black-Scholes equation provides us with a model for predicting call option prices, the actual call option prices are usually different from the prices predicted by the relationship between call option price and stock price.

There are many reasons for this, one of which (as implied in Subsection **2.2.4**) is that it is a very strong assumption to take the volatility and interest rate to be constant. Indeed, volatilities of stocks are almost never constant, and are even considered difficult to measure accurately.

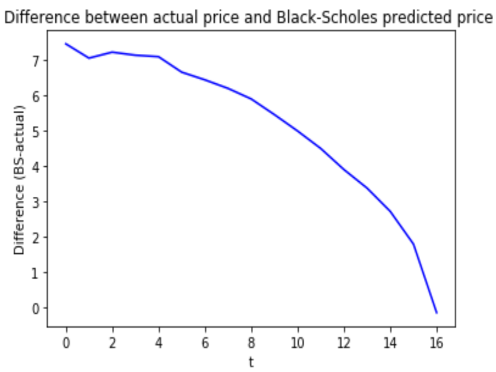
Furthermore, there any many other variables which will affect stock prices which the Black-Scholes model (as we have discussed) does not take into account.

Example: The following is taken from a randomly chosen call option, DB#F1920E950000 (named on the financial website 'MarketWatch' [10] as such), whose maturity date was 12/06/2020, from the investment bank Deutsche Bank[11]. We set the interest rate r to the current UK interest rate, 0.001 (1%), which was found from the website of the Bank of England [12] on Saturday 13^{th} June 2020. We set the volatility to the average implied volatility (measured over 20 day periods, and as of 12/06/2020) for calls from Deutsche Bank, 0.6838 (68.38%), obtained from the website AlphaQuery [13] on Saturday 13^{th} June 2020.
Implied volatility denotes an 'estimate' for the volatility of a call, especially for the future; this is deduced from the price history of the call option. We use this figure as the volatility of a call is generally very difficult to determine. [9, p. 52-53] We have used stock price data for Deutsche Bank found on the webpage relating to Deustche Bank on Yahoo Finance as of Friday 12^{th} June 2020[11], which can be found in the group's GitHub repository also. We have obtained data for the price of the call option DB#F1920E950000 and the strike price of this option, \$9.50, from the website 'MarketWatch' [10]; this was done on Saturday 13^{th} June 2020.

We plot the predicted progression of the call option value with time (from 26/05/2020 until 12/06/2020) and the actual progression of the call option value with time (this has been done in Python and the code can again be found on the group's GitHub repository). We observe that the



(a) Predicted call values and actual call values



(b) Error between predicted and actual call option values

Figure 3: An illustration of the different behaviour between predicted and actual prices over time for a randomly chosen call of Deutsche Bank

behaviour of the predicted call option values and actual call values are in fact very different.
We can attribute at least part of this disparity between the two to the values for the other parameters governing the value of the call option, which are likely to have limited accuracy.
However, the disparity between the predicted values and the actual values of the call option is largely because of the previously stated fact that the stock price depends on several variables which the Black-Scholes equation that *we* are considering (there are other variants where r and σ are not taken to be constant) does not consider.

3.1.3 Variance Ratio Test for Market Efficiency

A common property of random walk theory and Brownian motion is that the average radius of diffusion of a particle is proportional to $\sqrt{t}$, where t is time. In this part, we decided to use this property as a test: to verify whether the price returns in the market follow a Random Walk Hypothesis (RWH), based on whether or not they satisfy this particular property.

The **Variance Ratio Test** [14][10] (developed by Lo and MacKinley in 1988) gives a method to check this property in a financial setting. The Lo-MacKinley Variance Ratio is defined as $VR := \frac{\hat{\sigma}(k)^2}{S^2}$ where $\hat{\sigma}(k)$ is the k-period return variance, and S^2 is the usual sample variance (or $\hat{\sigma}(1)^2$). Whenever the data follows a RWH, this ratio should follow a normal distribution.

We decided to verify this by importing high-frequency data from various stocks of the S&P 500 index, and using the package *vrtest* in R (all code is available on the project's Github). We then tried to verify that the sample of variance ratios follows (approximately) a normal distribution:

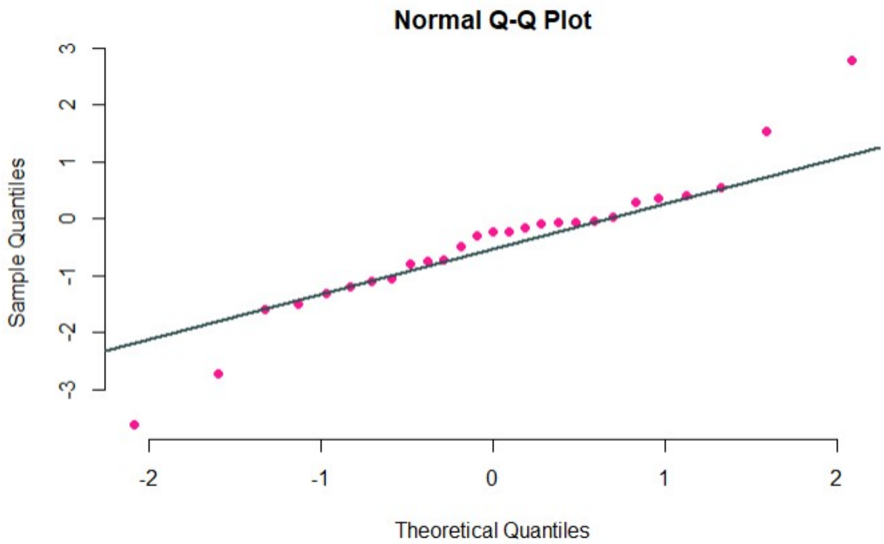


Figure 4: Quantile-quantile plot of the empirical variance ratios from the S&P 500 stock data (17/11/2015).

The plot is almost linear, which confirms our hypothesis: random walk can sometimes be a sensible model for stock markets. In particular, this outlines the hypothesis of *weak-form market efficiency*: the theory that the current price of a security reflects all its price history.

However, this is far from a perfect model. Some outliers in our simulations ended up having very large p-values (≈ 0.7), which led to completely rejecting the normal hypothesis. Indeed, in real-world conditions, a stock price can be affected by various outside conditions (war, pandemic, tweet,...), so it cannot be fully modelled by a purely random walk - which by definition depends **only** on the price history of said stock.

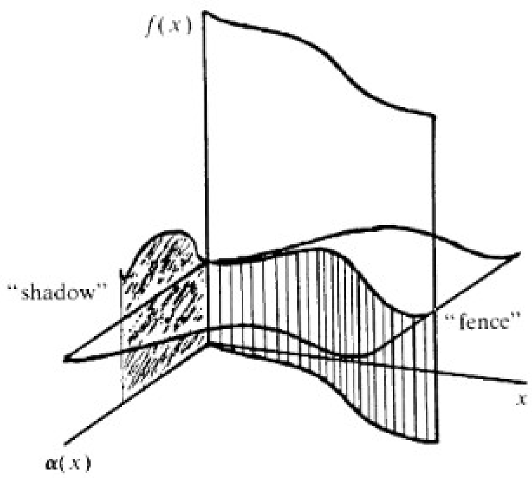
[10]The idea to use this test follows from discussions with Dr. Andreas Sojmark and on the forum *Quant Stack Exchange*.

3.2 Introduction to Itô Calculus

3.2.1 Motivation for Itô Calculus

Along this report, we have been using discrete models to introduce Brownian motion, and then tried to see how this stochastic process could be used to model financial markets. One of the earliest known developments of stochastic calculus was made in 1900 by Louis Bachelier, under the supervision of Henri Poincaré. His thesis introduced Brownian motion and actually also used it as a tool to model the (Paris) stock market; and his work was later recognized by pioneers of the theory of probability such as Kolmogorov.

Now, we have also seen that a Brownian Motion is almost surely continuous but nowhere differentiable. This makes it impossible to use the "classical" rules of calculus, and one needs a new theory of integration and differentiation. A first (but incorrect) guess -to go around this technicality- would be to use Riemann-Stielje calculus.



This generalization of Riemann calculus allows us in particular to integrate a function $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$ with respect to another real-valued function α. Here, the graph of f has been extended horizontally to a 2D surface, and the graph of α has been extended vertically to a 2D surface - for visual reasons. The Riemann-Stielje integral is the area of the fence, as defined on the picture.

Figure 5: Graph by Gregory L. Bullock [15]

The area of the fence is computed in the same way as the usual Riemann integral: by taking a partition of the interval, and taking the limit as the mesh of the partition goes to 0. So if f is Stielje-integrable w.r.t α:

$$\int_a^b f(x)\,\mathrm{d}\alpha(x) = \lim_{||\mathcal{P}||\to 0} \sum_{t_i \in \mathcal{P}} f(a_i)\,(\alpha(t_{i+1}) - \alpha(t_i))$$

where $\mathcal{P}$ ranges over the set of partitions of $[a, b]$, and $a_i \in [t_i, t_{i+1}]$.

However, this is still not good enough, because the Riemann-Stielje integral only allows integration with respect to particularly well-behaved functions: bounded variation functions. That is, functions whose "vertical" path length is finite.

The total variation of a function f (with $\mathcal{P}$ ranging over partitions of $[a,b]$) is defined as $V_1(f) := \sup_{\mathcal{P}} \sum_i |f(t_{i+1}) - f(t_i)|$. For smooth f, this quantity is finite. Indeed, take $f : [a,b] \to \mathbb{R}$ to be a $\mathcal{C}^1$ function. Then, it can be shown that $V_1(f) \leq \int_a^b |f'|$ by expanding the Riemann-Stielje sum and using the triangle inequality.[11]

Now, the Brownian motion does **not** have bounded total variation (or finite path length). Indeed, it has been proven in part **1.2.2** that the quadratic variation $[W]$ of W is linear in time. More precisely:

$$[W]_T := \lim_{||\mathcal{P}|| \to 0} \sum_i \left(W_{t_{i+1}} - W_{t_i}\right)^2 = T \text{ almost surely (a.s.)}$$

where $\mathcal{P}$ ranges over the set of partitions $\{t_i\}$ of $[0,T]$. Now, assume that W had finite total variation (*ie* $V_1(W) < \infty$), then:

$$\begin{aligned}
T &= \lim_{||\mathcal{P}|| \to 0} \sum_i \left(W_{t_{i+1}} - W_{t_i}\right)^2 \\
&\leq \lim_{||\mathcal{P}|| \to 0} \left((\sup_i |W_{t_{i+1}} - W_{t_i}|) \sum_i |W_{t_{i+1}} - W_{t_i}| \right) \\
&\leq \lim_{||\mathcal{P}|| \to 0} \sup_i |W_{t_{i+1}} - W_{t_i}| \cdot V_1(W) = 0 \text{ a.s.}
\end{aligned}$$

because W is continuous, hence uniformly continuous on $[0,T]$, implying that the left-hand limit is 0. Note that we also used $\lim_{||\mathcal{P}|| \to 0} \sum_i \left|W_{t_{i+1}} - W_{t_i}\right| \leq \sup_{\mathcal{P}} \sum_i \left|W_{t_{i+1}} - W_{t_i}\right| = V_1(W)$. We will go into more details about the different types of variation we can define a bit later.

This shows that Riemann-Stielje calculus is still not a good model when treating with stochastic processes, and that a new theory is required. This is where Itô Calculus comes in.

[11] Actually, this is an equality. The converse can be shown by choosing a partition of $[a,b]$ whose points are the zeros of f' - in the case where f' is constant on an interval $[c,d] \subseteq [a,b]$ (wlog this is the largest such interval), we only add c and d to the partition.

3.2.2 The Itô Integral, and Itô's lemma

The usual theory of calculus can be extended to stochastic processes in two main ways: Itô Calculus, and Malliavin Calculus. We will here focus on the former, and try to start by giving a sensible definition of integration.

As we saw earlier, a smooth function has bounded variation, which means the tag of the partition (*ie* the set of points we pick in each subinterval of a partition) we choose is irrelevant. This is not the case when we want to integrate with respect to W_t: we need to make a fixed choice for the tag and stick to it.

There are various ways to do it: one can choose the leftmost point of each subinterval (Itô integral), the midpoint (Stratonovich integral) or yet the rightmost point, among others. Each has its uses in different areas of applied maths, and one can move from one theory to another under suitable assumptions[12]; but we will here focus on the Itô integral, which is the most used in quantitative finance.

It consists in taking the leftmost point in each interval, so the integral of a function f with respect to W_t can be written as:

$$\int_a^b f(t)\,\mathrm{d}W_t = \lim_{||\mathcal{P}||\to 0} \sum_i f(t_i)\left(W_{t_{i+1}} - W_{t_i}\right)$$

where $\mathcal{P}$ ranges over the set of partitions $\{t_i\}$ of $[a, b]$.

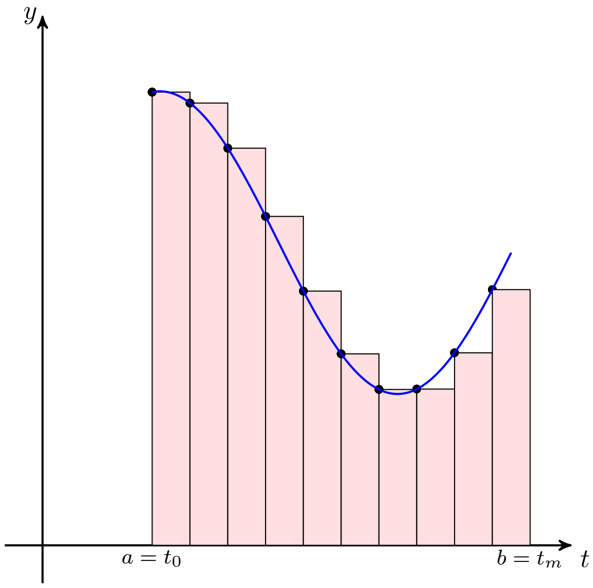


Figure 6: Visual representation of the Itô integral, where the tags are taken to be the leftmost points.

[12]One can, for instance, derive the Black-Scholes-Merton model from Stratonovich calculus instead of the usual Itô Calculus (which was done in section 2.1), as shown in [16].

Now, while that decision of choosing the leftmost point every time might seem a little arbitrary, there is a practical reason for it. In a financial setting, the different subdivisions represent very small periods of the trading day. Choosing any other point in the subdivision would be equivalent to "predicting" the future. A trading decision is always made at the start of the interval. So, even though Stratonovich calculus is equivalent, it is rarely used in quantitative finance.

Now, to introduce a cornerstone of Itô Calculus -Itô's lemma-, let us generalize the notion of total variation defined in **3.2.1**. Let $f : [a, b] \to \mathbb{R}$ be a function. Its $p-$variation is defined as:

$$V_p(f) := \sup_{\mathcal{P}} \sum_i |f(t_{i+1}) - f(t_i)|^p$$

where $\mathcal{P}$ ranges over the set of partitions $\{t_i\}$ of $[a, b]$. [13]

We will also define a second type of variation, which is now taken to be the limit *as the partition gets finer*, instead of the supremum over all partitions:

$$\Pi_p(f) := \lim_{||\mathcal{P}|| \to 0} \sum_i |f(t_{i+1}) - f(t_i)|^p$$

Note that the quadratic variation Π_2 is not the same as the 2-variation V_2. From definition, we can see that $\Pi_p(f) \leq V_p(f)$ for any $p \in \mathbb{N}\backslash\{0\}$, and for any function f. Now, let us prove an important result of real analysis: $\boxed{f \in \mathcal{C}^1 \implies \Pi_1(f) = V_1(f)}$.

Proof that $V_1(f) \leq \Pi_1(f)$:[14] $V_1(f)$ being the supremum of the sum over the set of partitions, for any $\varepsilon > 0$, one can find a partition $\mathcal{Q}$ such that $V_1(f) - \sum_{t_i \in \mathcal{Q}} |f(t_{i+1}) - f(t_i)| < \varepsilon$. This will hence hold for any refinement $\mathcal{Q}'$ of $\mathcal{Q}$. So in particular, if we take a sequence of refinements $\mathcal{Q}_n$ of $\mathcal{Q}$ whose mesh goes to 0, then:

$$\forall n \in \mathbb{N}\backslash\{0\} : V_1(f) - \sum_{\mathcal{Q}_n} |f(t_{i+1}) - f(t_i)| < \varepsilon$$

$$\implies V_1(f) < \lim_{n \to +\infty} \sum_{\mathcal{Q}_n} |f(t_{i+1}) - f(t_i)| + \varepsilon$$

$$\implies V_1(f) \leq \lim_{n \to +\infty} \sum_{\mathcal{Q}_n} |f(t_{i+1}) - f(t_i)| = \lim_{||\mathcal{P}|| \to 0} \sum_{\mathcal{Q}_n} |f(t_{i+1}) - f(t_i)| = \Pi_1(f) \qquad \square$$

Where the last two sums are equal because they're both Riemann sums of f, which is taken to be continuous - hence Riemann-integrable.

[13] Some authors consider the p-th root of the sum, but for practical reasons, we will not. In the context, it doesn't make a difference which definition we choose.

[14] The following proof was given by user *h3fr43nd* on StackExchange in order to simplify our original attempt (which can be found on [17]). Note that one needs to be careful with the definition of limit as the partition gets finer, this is also specified in one of the StackExchange answers.

Now, applying the mean value theorem on f (defined as above) implies $\Pi_2(f) \leq f'(\xi) \lim_{||\mathcal{P}|| \to 0} \sum_i ||\mathcal{P}||^2$ for some $\xi \in [a, b]$. This implies $\Pi_2(f) = 0$ and actually, one can show that $\Pi_p(f) = 0$ for $p > 2$:

$$\Pi_p(f) = \lim_{||\mathcal{P}|| \to 0} \sum_i |f(t_{i+1}) - f(t_i)|^p = \lim_{||\mathcal{P}|| \to 0} \sup_i |f(t_{i+1}) - f(t_i)|^{p-2} \cdot \Pi_2(f) = 0$$

Another way to see this is that, for smooth f, the infinitesimal variations of order 2 or more can be neglected. This is why, when defining differentiation in classical analysis, we take the derivative $Df(p)$ of f to be the linear map such that: $f(x) = f(p) + f(x)Df(p)[x - p] + o(||x - p||)$.

However, as we saw earlier with Brownian motion: infinitesimal variations of order 2 **cannot** be neglected, because on $[a, b] : \Pi_2(W) = b - a \neq 0$ if $a \neq b$. Now, pick $p > 2$

$$\Pi_p(W) = \lim_{||\mathcal{P}|| \to 0} \sup_i |W_{t_{i+1}} - W_{t_i}|^{p-2} \sum_i |W_{t_{i+1}} - W_{t_i}|^2 = 0$$

This comes from the fact that, since W is continuous on $[a, b]$ compact, it is uniformly continuous. Hence, $\lim_{||\mathcal{P}|| \to 0} \sup_i |W_{t_{i+1}} - W_{t_i}|^{p-2} = 0$. This means that terms of order higher than two can be neglected. Hence, take a smooth function $f = f(t, x)$; and take x to be the Brownian motion, *ie* $x = W_t$. To Taylor expand f, we need to consider $\mathrm{d}t$ terms up to order 1 (classic rules of calculus apply), and $\mathrm{d}W_t$ terms **up to order 2**. This gives:

$$f(t + \mathrm{d}t, W_{t+\mathrm{d}t}) = f(t, W_t) + \frac{\partial f}{\partial t}\mathrm{d}t + \frac{\partial f}{\partial x}\mathrm{d}W_t + \frac{1}{2}\frac{\partial^2 f}{\partial W \partial t}\mathrm{d}W_t\mathrm{d}t + \frac{1}{2}\frac{\partial^2 f}{\partial W^2}\mathrm{d}W_t^2 + o(\mathrm{d}t) + o(\mathrm{d}W_t^2)$$

Finally, we saw in part **1.2.2** that the quadratic variation property of W can be expressed in the form $\mathrm{d}W_t^2 = \mathrm{d}t$. This, in particular, implies that we can neglect the term in $\mathrm{d}W_t\,\mathrm{d}t$ (because, informally, we have $\mathrm{d}W_t \sim \mathrm{d}t^{1/2} \implies \mathrm{d}W_t\,\mathrm{d}t \sim \mathrm{d}t^{3/2} \sim 0$). This finally gives us:

Itô's Lemma

$$\boxed{\text{If } f \in \mathcal{C}^2(\mathbb{R}, \mathbb{R}) : \mathrm{d}f(t, W_t) = \left(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right)\mathrm{d}t + \frac{\partial f}{\partial x}\mathrm{d}W_t + o(\mathrm{d}W_t)}$$

This result is the foundation of Itô Calculus.[15] And the intuition behind this "variational argument" follows from discussions with our supervisor, Dr. Andreas Sojmark, which have been enriched by results from the Math StackExchange forum, and with Choongbum Lee's quantitative finance lectures at MIT [18].

The lemma is often stated in a slightly more general way by considering an arbitrary drift process X_t defined by $\mathrm{d}X_t = \mu_t\mathrm{d}t + \sigma_t\mathrm{d}W_t$. μ represents the drift of the process, and σ its volatility. The general form is given in section **2.1**, and is used to derive the Black-Scholes equation.

[15] This derivation was inspired by Choongbum Lee's lectures on Itô Calculus at MIT in the fall of 2013 [18].

The lemma can also be reformulated in integral form, for practical reasons. Assume that the function f does not depend on t, then:

$$f(W_t) = f(0) + \int_0^T f'(W_t)\,\mathrm{d}W_t + \frac{1}{2}\int_0^T f''(W_t)\,\mathrm{d}t$$

This rewriting gives a way to evaluate integrals without having to resort to the sum-definition. For example, plugging in $f : x \mapsto x^2/2$ gives $\displaystyle\int_0^T W_t\,\mathrm{d}W_t = \frac{1}{2}W_t^2 - \frac{1}{2}T$.

3.2.3 Conclusion

We have only seen a few of the many ways in which Itô's lemma appears as meaningful and significant. It links to martingale theory, a branch of probability; but also to the Black-Scholes-Merton model in finance, as we saw in part **2**. It has, in general, been useful in all areas of applied mathematics where stochastic calculus comes in; this includes, for example, the study of noise-disturbed models in physics, or algorithmic models like the Monte Carlo methods.

However, one should not forget that this model also has its limits. The Black-Scholes equation is only valid under specific assumptions (*no transaction costs, a known risk-free interest rate, no limit on short-selling, etc...*). A lot of experts -including the mathematician Benoit Mandelbröt- have been criticizing the abuse of the equation by financial institutions without even making sure those assumptions held. And many actually assign part of the blame for the 2007 subprimes crash to the abuse of this very model.

> The longer-term the option, the sillier the results generated by the Black-Scholes option pricing model, and the greater the opportunity for people who didn't use it.
>
> -Michael Lewis, *The Big Short: Inside the Doomsday Machine*

Along this project, we have been able to gain a deeper understanding of stochastic processes, in particular simple random walks and Brownian motions. We've been able to see that, despite their random nature, they could be modelled by deterministic PDEs (*partial differential equations*), in particular the heat equation. This understanding has then helped us understand direct applications to real-world models: financial markets. We have been trying to see how they behave, how well this behaviour could be approximated, and in particular why their dynamics could be linked to random walks models - through the heat equation.

Finally, all these concepts guided us towards a new form of Calculus, underlying the whole theory of stochastic processes - and thus giving us access to a wider set of tools to study them. This project helped us develop more "team-oriented" skills, by requiring us to organize carefully our work and our schedule as a group, instead of individually. We owe to the resourceful help and advice of Dr. Andreas Sojmark, for guiding us through the maze of stochastic calculus and making sure we didn't get lost.

References

[1] Gregory F Lawler. *Random Walk and the Heat Equation.* American Mathematical Society, 2010. URL https://www.math.uchicago.edu/~lawler/reu.pdf.

[2] Ronitt Rubinfeld Tom Leighton. *Mathematics for Computer Science, Random Walks.* URL https://web.mit.edu/neboat/Public/6.042/randomwalks.pdf. [Accessed: 09/06/2020].

[3] MIT OpenCourseWare. *17. Stochastic Processes II.* [Video], 2015. URL https://www.youtube.com/watch?v=92WaNz9mPeY. [Accessed: 30/05/2020].

[4] Steven E. Shreve. *Stochastic Calculus for Finance I, The Binomial Asset Pricing Model.* Springer Science & Business Media, 2005.

[5] Steve E. Shreve. *BS.dvi.* URL http://www.math.cmu.edu/~handron/21_370/BS.pdf. [Accessed: 11/06/2020].

[6] Rüdiger Kiesel N H Bingham. *Risk-neutral Valuation: Pricing and Hedging of Financial Derivatives, Second Edition.* Springer Science & Business Media, 2004.

[7] Investopedia. *Portfolio Definition.* URL https://www.investopedia.com/terms/p/portfolio.asp. [Accessed:13/06/2020].

[8] Younggeun Yoo. *STOCHASTIC CALCULUS AND BLACK-SCHOLES MODEL.* URL http://math.uchicago.edu/~may/REU2017/REUPapers/Yoo.pdf. [Accessed: 12/06/2020].

[9] Jeff Dewynne Paul Wilmott, Sam Howison. *The Mathematics of Financial Derivatives: A Student Introduction.* Cambridge University Press, 1995.

[10] MarketWatch. *Big Charts from MarketWatch.* URL https://bigcharts.marketwatch.com/quickchart/quickchart.asp?symb=DB%23F1920E950000&insttype=&freq=1&show=&time=4. [Accessed:13/06/2020].

[11] Yahoo Finance. *Deutsche Bank AG/Historical Data.* URL https://finance.yahoo.com/quote/DB. [Accessed: 13/06/2020].

[12] Bank of England. *Official Bank Rate History.* URL https://www.bankofengland.co.uk/boeapps/database/Bank-Rate.asp. [Accessed: 12/06/2020].

[13] AlphaQuery. *Deutsche Bank Aktiengesellschaft (DB.* URL https://www.alphaquery.com/stock/DB/volatility-option-statistics/20-day/iv-call. [Accessed:13/06/2020].

[14] Jeng-Hong-Chen. Variance ratio tests of random walk hypothesis of the euro exchange rate. *International Business & Economics Research Journal (IBER)*, 7(12), 2008.

[15] Gregory L. Bullock. A geometric interpretation of the riemann-stieltjes integral. *The American Mathematical Monthly*, 95(5):448–455, 1988.

[16] M. Montero J. Masoliver J. Perello, J. M. Porra. Black-scholes option pricing within ito and stratonovich conventions. *Physica A: Statistical Mechanics and its Applications*, 278(1-2): 260–274, 2000.

[17] h3fr43nd. *StackExchange*. URL `https://math.stackexchange.com/questions/3715636/proof-check-alternative-formula-for-the-total-variation/3715691#3715691`. [Accessed: 11/06/2020].

[18] MIT OpenCourseWare. *18. Itō Calculus*. [Video], 2015. URL `https://www.youtube.com/watch?v=Z5yRMMVUC5w`. [Accessed: 30/05/2020].