BROWNIAN MOTION

EMMA CARDWELL, JOSHIKA CHAKRAVERTY, KATHERINE TAYLOR

Abstract. We introduce Brownian motion as a physical process, and go over the definition and basic properties of Brownian motion as a mathematical process, including nondifferentiability. We explore the connections Brownian motion has to Markov chains, martingales, and random walks, and conclude with an explanation of stochastic calculus and an application of geometric Brownian motion to the Black-Scholes model for stock option pricing.

1. Introduction to Brownian Motion

Brownian motion is the random movement of particles in a liquid or gas. This phenomenon was first observed by botanist Robert Brown in a sample of pollen particles suspended in water. Later on, Albert Einstein modelled this motion with individual water molecules moving these particles. The basic idea is that, due to random interactions with other particles, a particle in Brownian motion will take a series of steps, each in a random direction independent from any of its past steps. This makes it essentially a random walk. In the mathematical abstraction of this physical process, we will take the limit of the random walk as the step size and the time interval between steps go to zero. (Of course, the ratio between step size and time interval also matter; we shall see that in order to get Brownian motion we must have step size vary with the square root of the time interval.) The resulting process of mathematical Brownian motion has many of the properties of a random walk, but it is also a continuous function of time, lending it many other interesting properties.

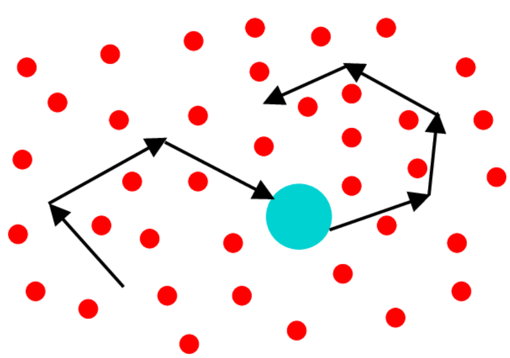

Figure 1. The motion of the blue particle follows a Brownian motion. Image source: [DBS16].

Date: November 22, 2020.

2. Modeling Brownian Motion

Before we introduce the mathematical definition of Brownian motion, we shall introduce the notions of probability space and stochastic processes. Probability spaces are used to model the outcomes of random processes.

Definition 2.1 (Probability Space)**.** A probability space is a triple $(\Omega, \mathcal{F}, P)$, where Ω denotes the sample space, $\mathcal{F}$ denotes the event space, and P denotes the probability function.

More formally, the sample space, Ω, can be any non-empty set. The event space, $\mathcal{F}$, is a set consisting of some (or all) subsets of Ω. The event space is modelled as a σ-algebra[1] of the power set[2] of Ω, meaning it satisfies the following properties:

(1) $\Omega \in \mathcal{F}$.
(2) $\mathcal{F}$ is closed under set complements, so for all $A \in \mathcal{F}$, we have $\Omega \backslash A \in \mathcal{F}$.
(3) $\mathcal{F}$ is closed under countable unions, so for all $A_1, A_2, \ldots \in \mathcal{F}$, we have $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

In other words, for any events A and B in $\mathcal{F}$, the events A^c and (A or B) must also be in $\mathcal{F}$. And the probability function, P, is a function $P : \mathcal{F} \to [0,1]$ such that the following properties are satisfied:

(1) For all $A \in \mathcal{F}$, $P(A) \geq 0$.
(2) $P(\Omega) = 1$.
(3) P is countably additive, so if $A_1, A_2, \ldots \subseteq \mathcal{F}$ is a countable collection of pairwise disjoint sets, then $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right)$.

Definition 2.2 (Real-Valued Stochastic Process)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space, and the index set, $T \subset \mathbb{R}$, be of infinite cardinality. Let $X(t, \omega) : T \times \Omega \to \mathbb{R}$ be a random variable defined on $(\Omega, \mathcal{F}, P)$. The function $X(t, \omega)$ is a stochastic process with indexing set T.

$X(t, \omega)$ can also be interpreted as a collection of random variables $\{X(t) : t \in T\}$ on Ω indexed by T. (Here, given some $t_i \in T$, we let $X(t_i)$ be the random variable taking values in the set $\{X(t_i, \omega) : \omega \in \Omega\}$ according to P). Note that what we have introduced is a "real-valued stochastic process" because each random variable X_t takes values in $\mathbb{R}$ rather than some other set (in which case it would be called simply a "stochastic process"). In this paper, we will only use real-valued stochastic processes, and we'll generally just call them stochastic processes. The indexing set $T \subset \mathbb{R}$ is often considered to be time, which will be especially relevant for Brownian Motion.

Definition 2.3 (Brownian Motion (Wiener Process))**.** Brownian motion is a stochastic process $B = \{B(t) : t \in \mathbb{R}_{\geq 0}\}$ on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with the following properties [Dun16]:

[1]This is also known as a σ-field.
[2]The power set of Ω is the set $\mathcal{P}(\Omega)$ containing all possible subsets of Ω, including the empty set and the set Ω itself.

(1) The random variable defined by $B(t)-B(s)$ over an interval of length $t-s$ is normally distributed with mean 0 and variance[3] $t-s$, that is

$$B(t)-B(s) \sim N(0, t-s)$$

(2) For every pair of disjoint time intervals $[t_1, t_2]$ and $[t_3, t_4]$, with $t_1 < t_2 \leq t_3 < t_4$, the random variables $B(t_4)-B(t_3)$ and $B(t_2)-B(t_1)$ are independent, and similarly for n disjoint time intervals where n is an arbitrary positive integer.

(3) $B(t)$ is continuous for all t.

Definition 2.4 (Standard Brownian Motion (Standard Wiener Process))**.** A Standard Brownian Motion is a Brownian motion $B(t)$ where $B(0)=0$.

Somewhat confusingly, "Brownian motion" is used to refer both to the physical process of random motion that we described in the introduction and to the mathematical process we've just defined. Those who care about the distinction will often call the mathematical process the Wiener process instead, but we'll just call it Brownian motion. We can think of the first property of Brownian motion in terms of statistics. It says that $\mathbb{E}[W(t)]=0$, and that the expectations of all other sums and differences of random variables will also be 0. Also, $\mathrm{Var}(W(t))=t$ for all t. Intuitively, the second property says that the movement of Brownian motion over one interval is not related to its movement over any past interval. And of course, the third property of continuity is what makes Brownian motion different from a random walk and gives it some of its special characteristics.

Before we move on, we should note that it takes some nontrivial (more precisely, theorem-level) work to prove that a stochastic process satisfying all the properties of Brownian motion actually exists.

Theorem 2.5 (Wiener's Theorem)**.** *Brownian motion exists.*

The proof of this theorem is rather lengthy and has more to do with analysis than probability, so we omitted it for the sake of brevity. A nice, self-contained proof can be found in [McK09].

3. Basic Properties of Brownian Motion

Here are some interesting properties of Brownian Motion that follow easily from the definition.

Proposition 3.1 (Time translation)**.** *If $B(t)$ is a standard Brownian motion and $h \geq 0$, then $W(t)=B(t+h)-B(h)$ is also a Brownian motion.*

Proof. $W(t)$ is clearly still a continuous stochastic process. The expected value of $W(t)$ is 0. Moreover, for any t_1, t_2, we have $\mathrm{Var}(W(t_2)-W(t_1)) = \mathrm{Var}(B(t_2+h)-B(t_1+h)) = t_2-t_1$, so we see that the variance property still holds. We can likewise show that the independence property holds, and thus $W(t)$ satisfies all the properties of a Brownian motion. Additionally, $W(0)=B(h)-B(h)=0$, so this is a standard Brownian motion. ■

[3] The variance of a random variable X is $\mathrm{Var}(X)=\mathbb{E}\left[(X-\mathbb{E}[X])^2\right]$.

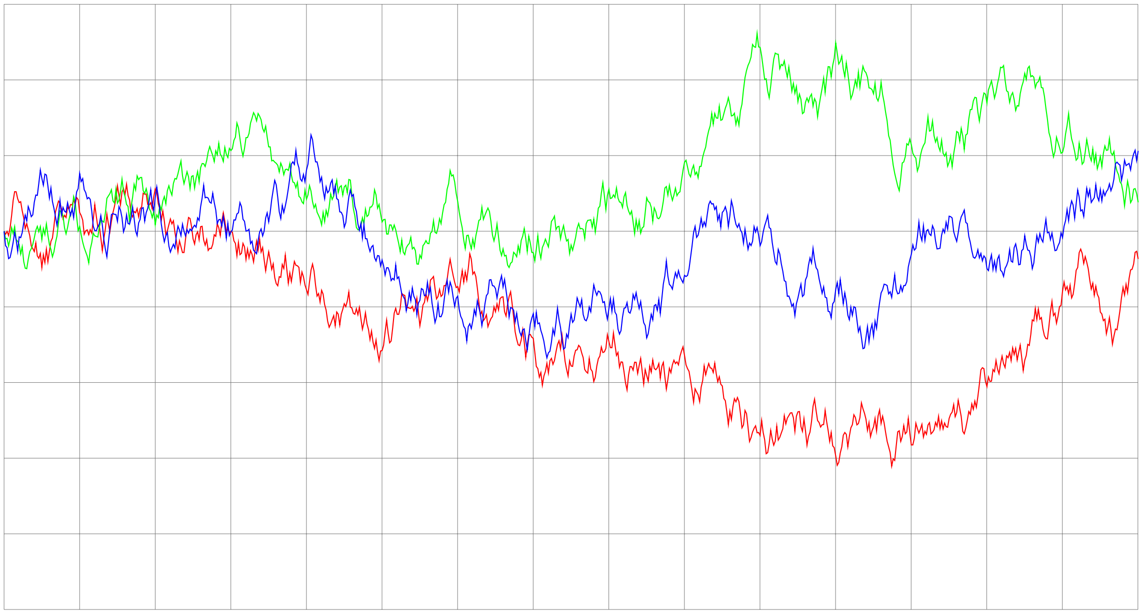

Figure 2. Three standard Brownian motions.

Proposition 3.2 (Scaling invariance)**.** *If $B(t)$ is a standard Brownian motion, then $\frac{1}{a}B(a^2t)$ is also a standard Brownian motion.*

Proof. Suppose $B(t)$ is a standard Brownian motion. We wish to show that $W(t) = \frac{1}{a}B(a^2t)$ also satisfies the properties of a standard Brownian motion. It is clear that $W(t)$ is still a stochastic process. Next, we wish to show that $W(t) - W(s)$ is normally distributed with mean 0 and variance $t - s$, for any t and s. By the definition of Brownian motion and the properties of variance, we have

$$\begin{aligned}
\mathrm{Var}(W(t) - W(s)) &= \mathrm{Var}\left(\frac{1}{a}B(a^2t) - \frac{1}{a}B(a^2s)\right) \\
&= \frac{1}{a^2}\mathrm{Var}(B(a^2t) - B(a^2s)) \\
&= \frac{1}{a^2}(a^2t - a^2s) \\
&= t - s
\end{aligned}$$

as desired. Clearly $W(0) = 0$, and the independence of time intervals and continuity are also not changed by rescaling. Thus, $W(t)$ is a standard Brownian motion. ■

The scaling invariance property tells us that Brownian motion behaves and looks essentially the same at every scale. It is sometimes said that Brownian motion is a fractal process, and this is why. Next, we have one more invariance property.

Proposition 3.3 (Time inversion)**.** *If $B(t)$ is a Brownian motion, then the following process is also a Brownian motion:*

$$V(t) = \begin{cases} 0 & t = 0 \\ tB(1/t) & t > 0 \end{cases}$$

To prove this, we need a lemma, and to explain the lemma, we need a definition:

Definition 3.4 (Covariance)**.** The covariance of two random variables X and Y is

$$\operatorname{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]$$

Note that $\operatorname{Cov}(X,X) = \operatorname{Var}(X)$, the variance of X. Roughly, the covariance captures the strength of the relationship between X and Y, and in fact the correlation coefficient of two variables X and Y is just a normalized version of the covariance.

Lemma 3.5. *Suppose $X(t)$ is a stochastic process which is known to satisfy all properties of Brownian motion except for the property that $\operatorname{Var}(X(t)-X(s)) = t-s$ for all $t \geq s$. Then $X(t)$ is a Brownian motion if and only if $\operatorname{Cov}(B(t+h), B(t)) = t$.*

Proof. We first wish to show that for a Brownian motion $B(t)$ and any $t > 0$, $h \geq t$, we have $\operatorname{Cov}(B(t+h), B(t)) = t$. We have

$$\begin{aligned}\operatorname{Cov}(B(t+h), B(t)) &= \mathbb{E}[(B(t+h)-\mathbb{E}[B(t+h)])(B(t)-\mathbb{E}[B(t)])] \\ &= \mathbb{E}[B(t+h)B(t)]\end{aligned}$$

Next, we substitute $B(t+h) = B(t+h) - B(t) + B(t)$:

$$\begin{aligned}\mathbb{E}[B(t+h)B(t)] &= \mathbb{E}[(B(t+h)-B(t)+B(t))B(t)] \\ &= \mathbb{E}[(B(t+h)-B(t))B(t)] + \mathbb{E}[B(t)^2]\end{aligned}$$

By the definition of Brownian motion, $B(t) = B(t) - B(0)$ is independent from $B(t+h) - B(t)$. Thus,

$$\mathbb{E}[(B(t+h)-B(t))B(t)] + \mathbb{E}[B(t)^2] = \mathbb{E}[B(t+h)-B(t)]\mathbb{E}[B(t)] + \mathbb{E}[B(t)^2]$$

But, also by the definition of Brownian motion, $\mathbb{E}[B(t+h)-B(t)] = 0$ and $\mathbb{E}[B(t)] = 0$. So, we have

$$\operatorname{Cov}(B(t+h), B(t)) = \mathbb{E}[B(t)^2] = \operatorname{Var}(B(t)) = t$$

as desired. Next, suppose $X(t)$ is known to satisfy all properties of Brownian motion except for the property that $\operatorname{Var}(X(t)-X(s)) = t-s$ for all $t \geq s$, and in addition, we know that $\operatorname{Cov}(X(t+h), X(t)) = t$ for all $t, h \geq 0$. We wish to show that $\operatorname{Var}(X(t)-X(s)) = t-s$. We have:

$$\begin{aligned}\operatorname{Var}(X(t)-X(s)) &= \mathbb{E}\left[X(t+h)^2 - 2X(t+h)X(t) + X(t)^2\right] \\ &= \mathbb{E}\left[X(t+h)^2\right] - 2\mathbb{E}\left[X(t+h)X(t)\right] + \mathbb{E}\left[X(t)^2\right]\end{aligned}$$

Recall that $\operatorname{Cov}(X(t+h), X(t)) = \mathbb{E}[X(t+h)X(h)]$ and that $\operatorname{Cov}(X(t+h), X(t)) = t$. Thus we have

$$\mathbb{E}\left[X(t+h)^2\right] - 2\mathbb{E}\left[X(t+h)X(t)\right] + \mathbb{E}\left[X(t)^2\right] = (t+h) - 2(t) + (t) = h$$

So, $\operatorname{Var}(X(t+h)-X(t)) = h$ and the proof is complete. ∎

Now we may prove the time inversion property.

Proof. Suppose $B(t)$ is a Brownian motion, and let $V(t)$ be defined by

$$V(t) = \begin{cases} 0 & t = 0 \\ tB(1/t) & t > 0 \end{cases}$$

We wish to show that $V(t)$ is a Brownian motion. We can check most of the properties fairly quickly. It is clear that $tB(1/t)$ has expected value 0 and that the random variables $V(t_4) - V(t_3)$ and $V(t_2) - V(t_1)$ are independent for $t_1 < t_2 \leq t_3 < t_4$. Continuity is a little trickier to prove because of the case $t = 0$. We will gloss over that issue here, but a rigorous proof that $V(t)$ is continuous at $t = 0$ can be found in the proof of time inversion given in [MP10]. Next, we show that $\text{Cov}(V(t+h), V(t)) = t$. Remembering that t and $t+h$ are scalars in this situation, we have

$$\begin{aligned} \text{Cov}(V(t+h), V(t)) &= \text{Cov}\left((t+h)B(1/(t+h)), tB(1/t)\right) \\ &= (t+h)(t)\,\text{Cov}(B(1/(t+h)), B(1/t)) \\ &= (t+h)(t)\frac{1}{t+h} \\ &= t \end{aligned}$$

As desired. Thus, by Lemma 3.5, $V(t)$ is a Brownian motion. ■

Because of the scaling invariance property, we know that as we zoom in on Brownian motion it stays bumpy instead of becoming smoother. This should make us suspect that Brownian motion is nondifferentiable, and we'll prove that next. To prove this, we need a few outside results and definitions.

Definition 3.6 (Limit superior)**.** The **limit superior** of a sequence (x_n) is

$$\limsup_{n\to\infty} x_n = \lim_{n\to\infty} \left(\sup_{m\geq n} x_n \right)$$

where sup denotes the supremum.

The limsup can be thought of as the limiting upper bound of the sequence.

Definition 3.7 (Limit inferior)**.** The **limit inferior** of a sequence (x_n) is

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} \left(\inf_{m\geq n} x_n \right)$$

where inf denotes the infimum.

This can be thought of as the limiting lower bound of the sequence. Note that the limsup and liminf can also be defined for any partially ordered set, such as a set indexed by $\mathbb{R}$.

Theorem 3.8 (Borel-Cantelli Lemma)**.** *Let $\{E_n : n \in \mathbb{N}\}$ be a countable set of events in a probability space $\mathbb{P}$ such that $\sum_{n=1}^{\infty} P(E_n) < \infty$. Then*

$$P(\limsup_{n\to\infty} E_n) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k\geq n}^{\infty} E_k \right) = 0$$

$\limsup E_n$ is the set of outcomes x in the probability space for which there exist infinitely many E_k such that $x \in E_k$. So, the Borel-Cantelli lemma says that the probability of infinitely many of the E_n's happening is 0.

Proposition 3.9 (Nondifferentiability)**.** *Almost surely, Brownian motion is nowhere differentiable.*

Proof. We shall prove a stronger statement: for any t, at least one of the following holds:

$$\limsup_{h\to 0} \frac{B(t_0+h)-B(t_0)}{h} = \infty$$

$$\liminf_{h\to 0} \frac{B(t_0+h)-B(t_0)}{h} = -\infty$$

Suppose we have some $t_0 \in [0,1]$ such that

$$\limsup_{h\to 0} \frac{|B(t_0+h)-B(t_0)|}{h} = C < \infty$$

Limit h to the range $[0,1]$. Then, since Brownian motion is bounded on $[0,2]$, there exists some constant $M \geq C$ such that

$$\sup_{h\in[0,1]} \frac{|B(t_0+h)-B(t_0)|}{h} \leq M$$

Now, let M be any constant. We wish to show that for this fixed value of M, the probability that there exists a $t_0 \in [0,1]$ satisfying the above property is 0.

Letting M be any constant, suppose that there exists such a t_0. In order to make some conclusions using t_0, we bound it within a specific interval. Let $n > 2$. We can say that $t_0 \in [(k-1)/2^n, k/2^n]$. Next, let $1 \leq j \leq 2^n - k$. By the Triangle Inequality,

$$|B((k+j)/2^n) - B((k+j-1)/2^n)| \leq |B((k+j)/2^n) - B(t_0)| + |B(t_0) - B((k+j-1)/2^n)|$$

Let $h = (k+j)/2^n - t_0$. We know that $j/2^n \leq h \leq (j+1)/2^n$. By the assumption on t_0, we know that

$$\frac{|B((k+j)/2^n) - B(t_0)|}{h} \leq M$$

and so

$$|B((k+j)/2^n) - B(t_0)| \leq M(j+1)/2^n$$

We can show by a similar argument that

$$|B(t_0) - B((k+j-1)/2^n)| \leq Mj/2^n$$

and so, plugging these in to our original inequality, we get

$$|B((k+j)/2^n) - B((k+j-1)/2^n)| \leq M(2j+1)/2^n$$

which will be true whenever t_0 exists. Define the events

$$\Omega_{n,k} = \{|B((k+j)/2^n) - B((k+j-1)/2^n)| \leq M(2j+1)/2^n \text{ for } j = 1,2,3\}$$

Because the increments of Brownian motion are independent, and then by the scaling invariance and time translation properties, we have

$$\begin{aligned}
P(\Omega_{n,k}) &= \prod_{j=1}^{3} P(|B((k+j)/2^n) - B((k+j-1)/2^n)| \le M(2j+1)/2^n) \\
&= \prod_{j=1}^{3} P(|B(k+j) - B(k+j-1)| \le M(2j+1)/\sqrt{2^n}) \\
&= \prod_{j=1}^{3} P(|B(1) - B(0)| \le M(2j+1)/\sqrt{2^n}) \\
&\le P\left(|B(1)| \le 7M/\sqrt{2^n}\right)^3
\end{aligned}$$

It is a fact that for a normally distributed random variable X with mean 0 and variance 1, we have $P(-c \le X \le c) = 2P(0 \le X \le c) \le c$ for all c. (To see this, observe that the slope of the normal cumulative density function is always less than $\frac{1}{2}$.) Thus,

$$P\left(|B(1)| \le 7M/\sqrt{2^n}\right) \le 7M/\sqrt{2^n}$$

and we conclude that $P(\Omega_{n,k}) \le (7M/\sqrt{2^n})^3$. Thus, for a given $n > 2$, we can see that

$$P\left(\bigcup_{k=1}^{2^n-3} \Omega_{n,k}\right) \le 2^n(7M/\sqrt{2^n})^3$$

When we sum over all $n > 2$ on both sides, we get $P(\bigcup_{k=1}^{2^n-3} \Omega_{n,k}$ for every n) on the left and a convergent geometric series on the right. Thus, the sum of the probabilities of all $\bigcup_{k=1}^{2^n-3} \Omega_{n,k}$ is finite, so we can apply the Borel-Cantelli lemma: the probability of infinitely many $\bigcup_{k=1}^{2^n-3} \Omega_{n,k}$'s happening is 0. However, recall that whenever t_0 exists, each $\Omega_{n,k}$ must be true. Thus, the probability that t_0 exists must also be 0, meaning that Brownian motion is differentiable with probability 0. ■

Thus, though Brownian motion is continuous everywhere, it is differentiable nowhere.

4. Markov Property

In this section we will explore higher dimensional Brownian Motion, which requires linear Brownian Motion in every component.

Definition 4.1 (d-dimensional Brownian Motion)**.** If $B_1, B_2, ..., B_d$ are independent linear Brownian motions started at $x_1, x_2, ..., x_d$ respectively, then $B(t) : t \ge 0$ given by

$$B(t) = (B_1(t), B_2(t), ..., B_d(t))^T$$

is **d-dimensional Brownian Motion** started in $(x_1, x_2, ..., x_d)^T$.

The Markov Property on any stochastic process $X(t)$ tells us that if we know the process $X(t) : t \geq 0$ up to some time s, in order to predict the future, we would have the same information if we just knew the value at time s. This idea rings especially true for a time-homogeneous Markov process, where the stochastic process starts over at any time s.

We will need the concepts of filtrations to better explore the markov properties of Brownian Motion.

Definition 4.2 (Filtration)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space. A **filtration** on this probability space is a family $((t) : t \geq 0)$ of σ-algebras such that $\mathcal{F}(s) \subset \mathcal{F}(t) \subset \mathcal{F}$ for all $s < t$. A probability space with a filtration is called a **filtered probability space**.

A filtration $\mathcal{F}(t)$ represents the set of events that are observable at time t but nothing past that time. There are two general types of filtration that can be defined on all spaces: the standard filtration and the natural filtration. The natural filtration works on a process that records and maintains all past information. This is a simple filtration that just records every previous step. A standard filtration is right continuous and complete, so it is a natural filtration that is also continuous as time increases.

We can definite a filtration for any Brownian Motion $B(t)$ by letting $\mathcal{F}(t)$ be the σ-algebra generated by random variables $B(s) : 0 \leq s \leq t$. This Brownian Motion is adapted to the filtration, which contains all the information from the stochastic process until time t. We can think of this as the Markov chain shifting after a certain amount of time. The process $X(s+t) : t, s \geq 0$ has the same distribution as $X(s) : s \geq 0$.

It is important to know whether the state X_t, so we define the concept of adaptation.

Definition 4.3 (Adapted)**.** A stochastic process $X(t) : t \geq 0$ on a filtered probability space is **adapted** if $X(t) \in \mathcal{F}(t)$ for all $t \geq 0$.

Definition 4.4 (Admissible)**.** A filtration $\mathcal{F}(t)$ is **admissible** for Brownian motion $B(t)$ if

(1) $B(t)$ is adapted to $\mathcal{F}(t)$
(2) for every $t \geq 0$, $W_{t+s} - W_{t_{s \geq 0}}$ is independent of the σ-algebras.

The standard filtration is in any admissible filtration.

Proposition 4.5 (Markov Property)**.** *If $B(t)$ is a standard Brownian motion, then the standard filtration $\mathcal{F}(t)$ is admissible.*

Proof. This proof is based on fact that Brownian motion has independent increments. Fix a time $s \geq 0$ and look at two events A, B where

$$A = \cap_{j=1}^{n} B(s_{j-1}) \leq x_j \in \mathcal{F} \text{ and}$$

$$B = \cap_{j=1}^{m} B(t_j + s) - B(t_j + s) - W(t_{j-1} + s) \leq y_j.$$

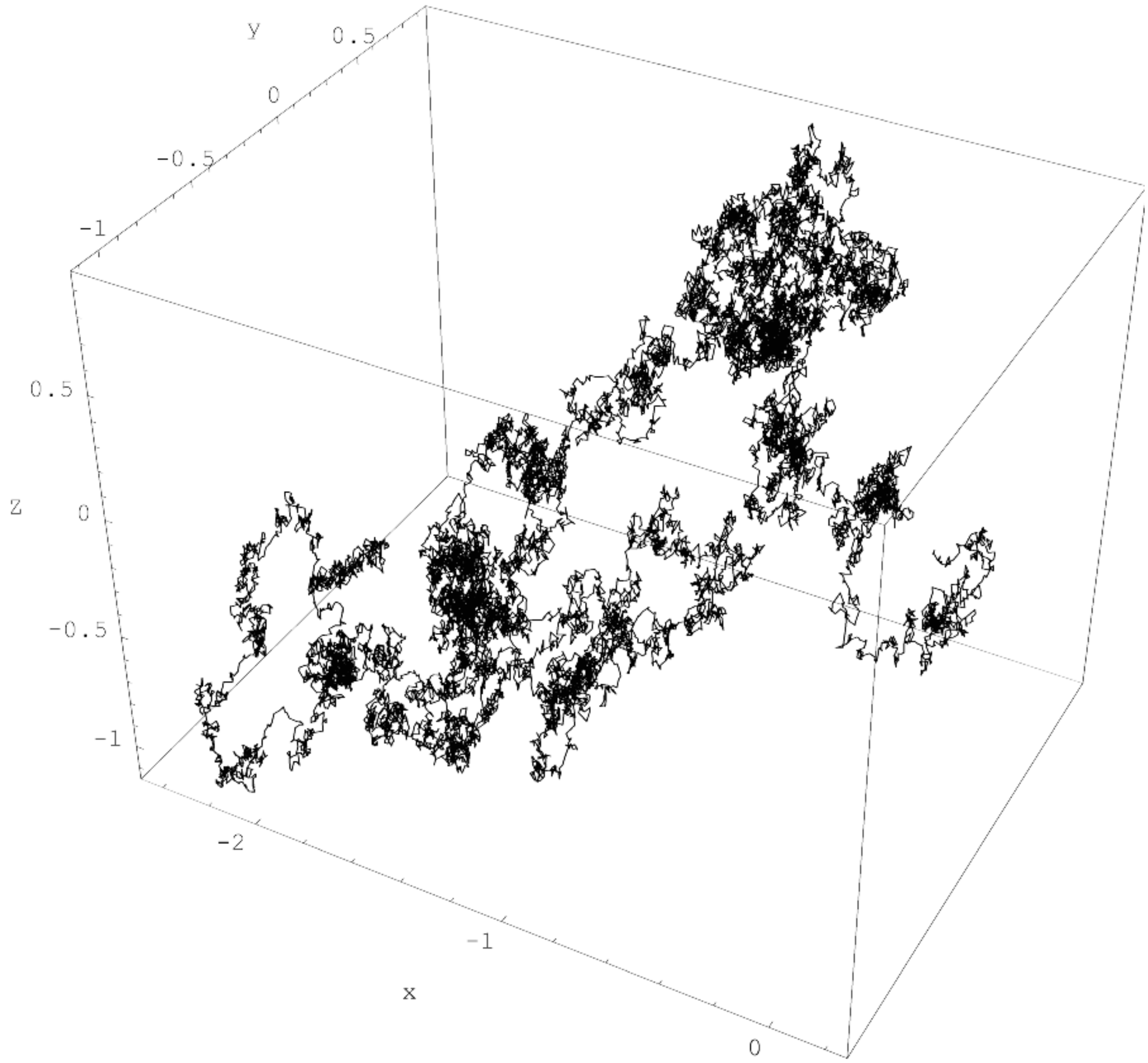

Figure 3. A single path for three-dimensional Brownian Motion. Shows the Markov Property as the future displacements do not depend on past displacements. Image source: [Wik13].

Because of the independent increments, events A and B are independent. A will generate the σ-algebra $\mathcal{F}(s)$ and B will generate the smallest σ-algebra where $B(t+s)-B(s)$ is measurable after time s. Thus, Brownian motion after time s is independent of the σ-algebra $\mathcal{F}(s)$. ■

For general Markov chains, the Strong Markov Property is a generalization of the Markov Property that says that after each stopping time, the future states of the chain are independent of anything else. For Brownian motion, the Strong Markov Property has to do with stopping times of the filtration. For that we need to define another random variable τ.

Definition 4.6 (Stopping Time)**.** A non negative random variable τ, which can go up to $+\infty$ is a **stopping time** with respect to a filtration $\mathcal{F}(t)$ is for every $t \geq 0$, the event $\tau \leq t \in \mathcal{F}(t)$.

A stopping time is **proper** if is is always less that infinity on the state space. Each stopping time has a **stopping field** which is a σ-algebra consisting of all events $B \subset \mathcal{F}(\infty)$ such that $B \cap \tau \leq t \in \mathcal{F}(t)$ for every $t \geq 0$.

For a fixed (F):

(1) Every constant $t \geq 0$ is a stopping time
(2) If τ and υ are stopping times, then $\tau \wedge \upsilon$ and $\tau \vee \upsilon$ are also stopping times
(3) If τ and υ are stopping times and $\tau \leq \vee$ then $\mathcal{F}(\tau) \subset \mathcal{F}(\vee)$

Theorem 4.7 (Strong Markov Property)**.** *Let $B(t)$ be a standard Brownian motion and τ be a stopping time relative to the standard filtration which an associated stopping σ-algebra. For $t \geq 0$, define the post-τ process*

$$B * (t) = B(t + \tau) - B(\tau),$$

and let $\mathcal{F}*$ be a standard filtration of this process. Then

(1) *$B * (t)$ is a standard Brownian motion*
(2) *For each $t > 0$, the σ-algebra $F * (t)$ is independent of $F(t)$*

5. Connection to Martingales

In this section, we will explore the martingale property of Brownian Motion. Intuitively, a martingale is a stochastic process where the expected value at any future time is just equal to the current value, creating a "fair" game. In real life, the gambling events that are modelled using martingales have pretty uniform profit margins. Some examples of martingales are random walks and the money a gambler has after a sequence of fair bets. The martingale property refers to any stochastic process that behaves like a martingale.

Definition 5.1 (Martingales)**.** A **martingale** is a stochastic process $X_t : t \in T$ such that $\mathbb{E}(X_t) < \infty\ \forall t \in T$, and for every $t_1 < t_2 < ... < t_n < t_{n+1}$ we have $\mathbb{E}(X_{t_n+1}|X_{t_n}, ..., X_{t_1}) = X_{t_n}$.

Life isn't always perfect, so there are definitions of martingales to account for the situation where the expected value in the future are larger or smaller than the current value.

Definition 5.2 (Submartingales)**.** A **submartingale** is a stochastic process $X_t : t \in T$ such that $\mathbb{E}(X_t) < \infty\ \forall t \in T$, and for every $t_1 < t_2 < ... < t_n < t_{n+1}$ we have $\mathbb{E}(X_{t_{n+1}}|X_{t_n}, ..., X_{t_1}) \geq X_{t_n}$.

Definition 5.3 (Supermartingales)**.** A **supermartingale** is a stochastic process $X_t : t \in T$ such that $\mathbb{E}(X_t) < \infty\ \forall t \in T$, and for every $t_1 < t_2 < ... < t_n < t_{n+1}$ we have $\mathbb{E}(X_{t_{n+1}}|X_{t_n}, ..., X_{t_1}) \leq X_{t_n}$.

Here are some basic properties of martingales that are good to keep in mind:

(1) For any integrable random variable X. $\{\mathbb{E}(X|\mathcal{F}(t)) : t \geq 0\}$ is a martingale.
(2) If $M(t)$ is a submartingale, then $t \to \mathbb{E}(M(t))$ is nondecreasing.
(3) If $M(t)$ is a martingale and ϕ is a convex function where $\mathbb{E}(|\phi(M(t))|) < \infty$ for all $t \geq 0$, then $\phi(M(t))$ is a submartingale.

The first two properties are easy to check. From now on, we will keep the first property in mind and just write $\mathbb{E}(X)$ instead of $\mathbb{E}(X|\mathcal{F}(t)$.

Proof of the Third Property. For this proof, we will need Jensen's Inequality, which says that for a random variable X where $\mathbb{E} < \infty$ and a convex function ϕ,

$$\phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X)).$$

Basically, Jensen's Inequality let's us the switch the order of the expectation and the convex function with the knowledge that putting the expectation on the outside might have more variability. ■

Wald's Identity or Wald's Lemma looks at a series of random variables $\chi_1, \chi_2, ..., \chi_n$ with the same mean and random variable τ representing the stopping time where

$$\mathbb{E}(\chi_1 + \chi_2 + ... + \chi_\tau) = \mathbb{E}(X_1)\mathbb{E}(\tau) \text{ holds.}$$

Proposition 5.4. *If $B(t)$ is a Brownian motion with an admissible filter $\mathcal{F}(t)$: each of the following are continuous martingale with respective to the admissible filtration:*

(1) $B(t)_{t\geq 0}$
(2) $B(t)^2 - t_{t\leq 0}$

Example. Brownian motion is a martingale for times s and t where $s < t$ for

$$\mathbb{E}_s(X_t) = \mathbb{E}_s(X_s) + \mathbb{E}_s(X_t - X_s) = X_s.$$

Example. Another example of involves one-dimensional Brownian motion . A stochastic process $M(t)$ is a martingale for a Brownian motion $B(t)$ and for times s and t where $s < t$ where

$$M(t) = B(t)^2 - t.$$

Proof. We have $\mathbb{E}_s(M(t)) = \mathbb{E}_s(B(s) + \delta B)^2 - t$ where $\delta B = B(t) - B(s)$. Using linearity of expectation, we can split the expectation up. We can split the expectation up, getting $\mathbb{E}_s(\delta B) + \mathbb{E}_s(\delta B)^2 - t$. $\mathbb{E}_s(\delta B) = 0$ and $\mathbb{E}_s(\delta B)^2 = t - s$. Thus, this whole equation equals $M(s)$, showing that this example is a martingale. ■

There are two important facts about martingales that help understand why gambling doesn't go on forever. There are discrete time equivalents to both of these facts.

First we have the optional stopping theorem, which shows how martingales can be extended from fixed times $0 \le s \le t$ to stopping times $0 \le S \le T$. For the example of a gambler in a fair game, the optional stopping theorem reveals that nothing can be gained from stopping without looking into the future.

For the optional stopping theorem we need the definition of continuous time martingales, which is a stochastic process $Y(t)$ with respect to another stochastic process $X(t)$ where $\mathbb{E}(Y(t)) < \infty$ and $\mathbb{E}(Y(t)|X(\tau), \tau \le s) = Y(s) \forall s \le t$.

Theorem 5.5 (Optional Stopping Theorem). *Let $Y(t)$ is a continuous martingale, and $0 \le S \le T$ are stopping times. If the process $Y(t \wedge T) : t \ge 0$ is dominated by an integrable random variable X where $|Y(t \wedge T)| \le X$ for all $t \ge 0$, then*

$$\mathbb{E}(Y(T)|\mathcal{F}(S)) = X(S)$$

almost surely.

The proof of this stopping theorem comes from the proof of the theorem for the discrete time case and then using dominated convergence for conditional expectations to get the theorem for continuous martingales. This proof can be seen in [MP10].

Next we have Doob's maximal inequality gives a bound on the probability that a stochastic process will be greater than a certain value in a certain interval of time.

Theorem 5.6 (Doob's Maximal Inequality). *If $B(t)$ is a continuous submartingale and $p > 1$, for any $t \ge 0$*

$$\mathbb{E}(\sup_{0 \le s \le t} |X(t)|^p) \le (\tfrac{p}{p-1})^p \mathbb{E}(|X(t)|^p).$$

This theorem will also have to be proved for the discrete case first. Using monotone convergence, we can generalize this inequality to continuous submartingales.

These two theorems and the martingale property can be used to prove Wald's lemmas for Brownian motion.

Lemma 5.7 (Wald's Lemma for Brownian Motion). *[Dur19] Let $\{B(t) : t \ge 0\}$ be a standard linear Brownian motion and τ be a stopping them where either*

$$\mathbb{E}(\tau) < \infty \textit{ or}$$

$$\{B(t \wedge \tau) : t \ge 0\} \textit{ is } L^1\textit{-bounded.}$$

Then $\mathbb{E}(B(\tau)) = 0$.

Wald's lemma can either be proved with an optional stopping argument or with the strong Markov property and the law of large numbers. This proof will use the optional stopping argument.

Proof. We first want to show that there exists a stopping time τ that satisfies both conditions. Suppose $\mathbb{E}(\tau) < \infty$. Define a new stochastic process $M(k)$ where

$$M(k) = \max_{0\le t\le 1} |B(t+k) - B(k)) \text{ and } M = \textstyle\sum_{k=1}^{|\tau|} M(k)$$

for a Brownian motion $B(t)$.

We can take the expectation of this process to check whether the expectation is less than ∞.

$$\begin{aligned}
\mathbb{E}(M) &= \mathbb{E}(\sum_{k=1}^{|\tau|} M(k)) \\
&= \sum_{k=1}^{\infty} \mathbb{E}(1\{\tau \ge k\} M(k) \\
&= \sum_{k=1}^{\infty} ((P)\{\tau \ge k\} \mathbb{E}(M(k)) \\
&= \mathbb{E}(M(0))\mathbb{E}(\tau) \\
&< \infty
\end{aligned}$$

To solve $\mathbb{E}(M(0))$, we can use the fact that we have an upper bound for probabilities using the tail estimate for the Gaussian distribution:

$$\mathbb{P}_0\{M(t) > a\} \le \tfrac{\sqrt{2t}}{a\sqrt{\pi}} e^{-\frac{a^2}{2t}}.$$

We also have Fubini's Theorem to place an upper bound on the probability, getting

$$\begin{aligned}
\mathbb{E}(M(0)) &= \int_0^\infty \{\max_{0\le t\le 1} |B(t)| > x\} dx \\
&\le \int_0^\infty \frac{2\sqrt{2}}{x\sqrt{\pi}} e^{-\frac{x^2}{2t}} \\
&< \infty
\end{aligned}$$

$|B(t\wedge\tau)| \le M$ so part 2 of the theorem is true. Under the part 2, we can apply the optional stopping theorem with $S = 0$, to get $\mathbb{E}(B(T)) = 0$. ■

There is a second, more general lemma about stopping times for Wald.

Lemma 5.8 (Wald's Second Lemma). *Let τ be a stopping time for standard Brownian motion such that $\mathbb{E}(\tau) < \infty$. Then*

$$\mathbb{E}(B(\tau)^2) = \mathbb{E}(\tau).$$

Proof. We will first define stopping times for the martingale $\{B(t)^2 - t : t \geq 0\}$. We have

$$\tau_n = \inf(t \geq 0 : |B(t)| = n).$$

This means that $\{B(t \wedge \tau \wedge \tau_n)^2 - t \wedge \tau \wedge \tau_n : t \geq 0\}$ will always be less than the integrable random variable $n^2 + \tau$. The optional stopping theorem gets

$$\mathbb{E}(B(\tau \wedge \tau_n)^2) = \mathbb{E}(\tau \wedge \tau_n).$$

We know that $\mathbb{E}(B(\tau)^2) \geq \mathbb{E}(B(\tau \wedge \tau_n)^2)$. We can take the limits as the value of the Brownian Motion approaches infinity and use monotone convergence to show that

$$\mathbb{E}(B(\tau)^2) \geq \lim_{n\to\infty} \mathbb{E}(B(\tau \wedge \tau_n)^2) = \lim_{n\to\infty} \mathbb{E}(\tau \wedge \tau_n) = \mathbb{E}(\tau).$$

To prove the inequality on the other side we need Fatou's lemma, which says that for a random variable Y where $\mathbb{E}(Y) < \infty$ then if $Y \leq X_n$ for all n, then

$$\mathbb{E}(\liminf_{n\to\infty} X_n) \leq \liminf_{n\to\infty} \mathbb{E}(X_n).$$

We can take advantage of our construction of the stopping time and apply Fatou's Lemma to get

$$\mathbb{E}(B(\tau)^2) \leq \liminf_{n\to\infty} \mathbb{E}(B(\tau \wedge \tau_n)^2) = \liminf_{n\to\infty} \mathbb{E}(\tau \wedge \tau_n) \leq \mathbb{E}(\tau).$$

Thus, $\mathbb{E}(B(\tau)^2) = \mathbb{E}(\tau)$, completing the proof. ■

6. Brownian Motion from Random Walks

Suppose we have a random walk on $\mathbb{R}$, where for every interval of time of length $1/n$ we take a step of length $\sqrt{1/n}$ either up or down. We'll represent the step taken at time k by the random variable $Z(k)$, which takes the values $\sqrt{1/n}$ and $-\sqrt{1/n}$ with equal probability. Let $X_n(t)$ be our position at time t. Then, since the expected value of $Z(k)$ is always 0, the expected value of $X_n(t)$ is

$$\mathbb{E}\left[\sum_{m=1}^{tn} Z\left(\frac{m}{n}\right)\right] = 0$$

We also have that the variance of $Z(k)$ is always $\frac{1}{n}$, and thus the variance of $X_n(t)$ is

$$\begin{aligned}
\operatorname{Var}\left(\sum_{m=1}^{tn} Z\left(\frac{m}{n}\right)\right) &= \mathbb{E}\left[\left(\sum_{m=1}^{tn} Z\left(\frac{m}{n}\right)\right)^2\right] \\
&= \mathbb{E}\left[\sum_{m=1}^{tn} Z\left(\frac{m}{n}\right)^2 + 2\sum_{j\neq m}^{n-1} Z\left(\frac{j}{n}\right) Z\left(\frac{m}{n}\right)\right] \\
&= \sum_{m=1}^{tn} \mathbb{E}\left[Z\left(\frac{m}{n}\right)^2\right] + 2\sum_{j\neq m}^{tn} \mathbb{E}\left[Z\left(\frac{j}{n}\right) Z\left(\frac{m}{n}\right)\right] \\
&= \sum_{m=1}^{tn} \frac{1}{n} + 2\sum_{j\neq m}^{tn} \mathbb{E}\left[Z\left(\frac{j}{n}\right)\right] \mathbb{E}\left[Z\left(\frac{m}{n}\right)\right] \\
&= t
\end{aligned}$$

This is starting to look a lot like Brownian motion, and in fact $X_n(t)$ satisfies all the properties of Brownian motion except for continuity. Intuitively, if we take the limit of $X_n(t)$ as $n \to \infty$ (i.e., taking smaller and smaller steps in shorter and shorter intervals of time), we expect to reach Brownian motion, and as it turns out that is indeed what happens.

Theorem 6.1 (Donsker's Invariance Principle). *Suppose we have a random walk on $\mathbb{R}$ with time interval $1/n$ and step length $\sqrt{1/n}$, with the position at time t given by the sequence $\{X_n(t) : t = \frac{m}{n}, m \in \mathbb{Z}_{\geq 0}\}$. Let $X^{(n)}(t)$ be the curve created by linearly connecting each successive pair of points $(X_n(t), X_n(t+\frac{1}{n}))$. Then the sequence $\{X^{(n)}(t) : n \in \mathbb{Z}_{>0}\}$ converges in distribution*[4] *to a standard Brownian motion.*

Donsker's invariance principle is also sometimes called the functional central limit theorem, as it makes an analogous statement to the central limit theorem in statistics. For suppose we ask what the sequence $\{X^{(n)}(1) : n \in \mathbb{Z}_{>0}\}$ converges to. Using the central limit theorem, we conclude that $\lim_{n\to\infty} X^{(n)}(1)$ is a normally distributed random variable with mean 0 and variance 1, which is precisely $B(1)$ for a standard Brownian motion. Thus, $\{X^{(n)}(1) : n \in \mathbb{Z}_{>0}\}$ converges in distribution to $B(1)$. Similarly, we can show using the central limit theorem that $\{X^{(n)}(s) : n \in \mathbb{Z}_{>0}\}$ converges in distribution to $B(s)$ for any $s \geq 0$. What Donsker's invariance principle does is show that this convergence holds not just pointwise but for the entire function at once.

Thus, Brownian motion is in a sense the functional analogue of the Gaussian distribution, connected to random walks in much the same way that the Gaussian distribution is connected to finite random variables with binomial distributions. This connection is important because Brownian motion is more powerful than random walks, allowing us to use Brownian motion to prove things about random walks.

Another example of the connection between Brownian motion and random walks is the law of the iterated logarithm.

[4] A sequence of random variables $\{X_n\}$ converges in distribution to a random variable X if the cdfs of the X_n's converge to the cdf of X, that is, if the distributions of the X_n's converge to the distribution of X.

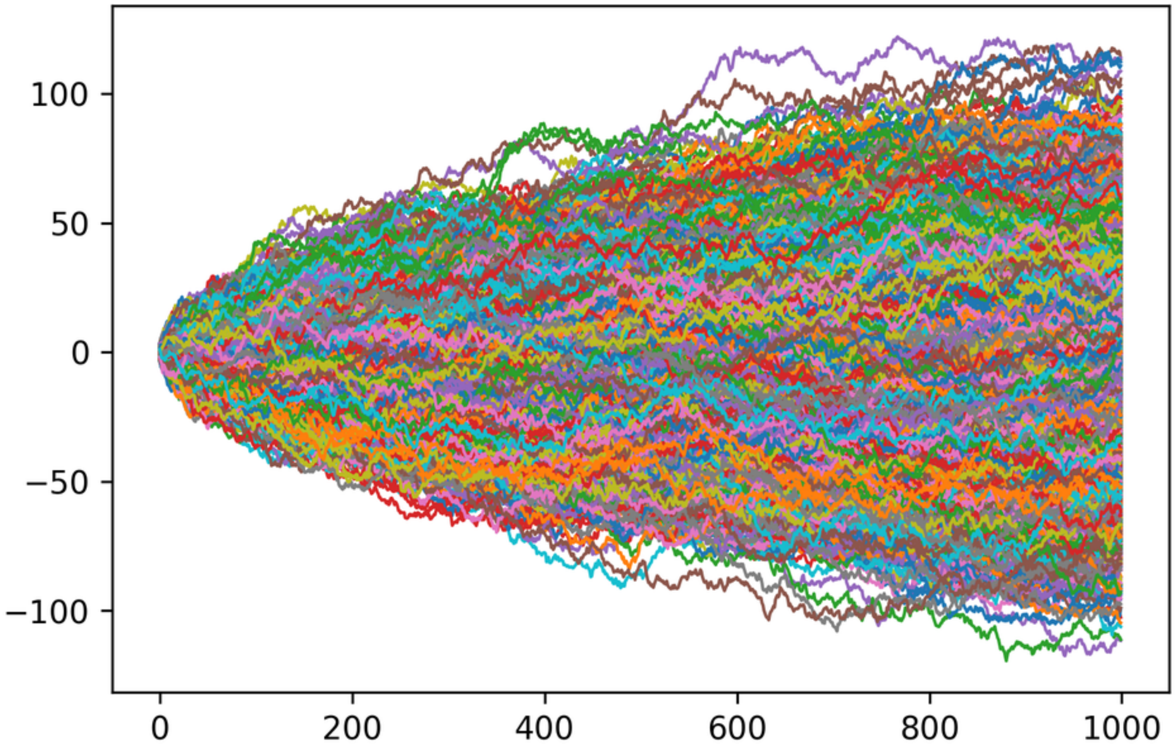

Figure 4. Over many iterations of a Brownian motion, the bounds given by the law of the iterated logarithm become more apparent. Image source: [bml18].

The law of the iterated logarithm is an interesting intermediary between two other bounds on the long-term behavior of random walks and Brownian motion. By the law of large numbers, we have that

$$\lim_{t\to\infty} \frac{B(t)}{t} = \lim_{t\to\infty} \frac{X_n(t)}{t} = 0$$

almost surely. And by the properties of Brownian motion and the central limit theorem, we have

$$\lim_{t\to\infty} \frac{B(t)}{\sqrt{t}} \sim \lim_{t\to\infty} \frac{X_n(t)}{\sqrt{t}} \sim N(0,1)$$

That is, the limits do not converge to any single value, but instead approach the standard normal distribution. While we will not prove this (or either of the preceding statements, for that matter), it is a fact that follows relatively quickly from the above and Kolmogorov's zero-one law that

$$\limsup_{t\to\infty} \frac{X_n(t)}{\sqrt{t}} = \infty, \quad \liminf_{t\to\infty} \frac{X_n(t)}{\sqrt{t}} = -\infty$$

and

$$\limsup_{t\to\infty} \frac{B(t)}{\sqrt{t}} = \infty, \quad \liminf_{t\to\infty} \frac{B(t)}{\sqrt{t}} = -\infty$$

These two limits suggest that somewhere in between $\sqrt{t}$ and t is a value where we can make a more precise statement about the behavior of Brownian motion and random walks. As it turns out, that value is $\sqrt{2t\log\log t}$.

Theorem 6.2 (Law of the Iterated Logarithm for Brownian Motion)**.** *Let $B(t)$ be a standard Brownian motion. Then, almost surely,*

$$\limsup_{t\to\infty} \frac{B(t)}{\sqrt{2t\log\log t}} = 1$$

and

$$\liminf_{t\to\infty} \frac{B(t)}{\sqrt{2t\log\log t}} = -1$$

This implies that for any $\epsilon > 0$, there exists t such that for all $s > t$,

$$|B(s)| \leq (1+\epsilon)\sqrt{2s\log\log s}$$

However, the law of the iterated logarithm also implies that for any $\epsilon > 0$, there exists a time s such that

$$|B(s)| \geq (1-\epsilon)\sqrt{2s\log\log s}$$

What this looks like in practice is that Brownian motion spends most of its time very close to the bound given by $\pm\sqrt{2t\log\log t}$, switching at random between the upper and lower bounds. There is also a (very similar) version of the law of the iterated logarithm for random walks:

Theorem 6.3 (Law of the Iterated Logarithm for Random Walks)**.** *Let $\{X_n(t) : t = \frac{m}{n}, m \in \mathbb{Z}_{\geq 0}\}$ be a random walk with time interval $1/n$ and step length $\sqrt{1/n}$, and probability $\frac{1}{2}$ each of stepping up or down. Then, almost surely,*

$$\limsup_{t\to\infty} \frac{X_n(t)}{\sqrt{2t\log\log t}} = 1$$

and

$$\liminf_{t\to\infty} \frac{X_n(t)}{\sqrt{2t\log\log t}} = -1$$

The law of the iterated logarithm for random walks can be proven from the law of the iterated logarithm for Brownian motion using embeddings. [Str64]

A related and very interesting fact about Brownian motion is the following: the arcsine distribution says that the cumulative distribution function of one-dimensional Brownian Motion is

$$\mathbb{P}\{X \leq x\} = \frac{2}{\pi}\arcsin\sqrt{x}$$

for $x \in (0,1)$. This distribution can be proved with Donksker's Invariance Principle and analysis of the extrema of the function.

7. Itô Calculus

Many stochastic processes are constructed with functions that are not differentiable. In fact, we proved that a Brownian motion is not differentiable anywhere in Proposition 3.9.

Furthermore, ordinary calculus cannot account for the random behavior of stochastic processes. To extend the concepts of calculus to stochastic processes, we need to use stochastic calculus[5]. Since we can't differentiate, we can start by defining integration[6].

Definition 7.1 (Itô Integral)**.** An **Itô integral** of the function $f = f(t)$ with respect to the Brownian motion B on the finite interval $[0, T]$ is a function $W = W(t)$, $t \in [0, T]$ given by:

$$W(t) = \int_0^t f(s)dB(s) = \lim_{N\to\infty} \sum_{k=1}^{N} f(t_{k-1})\left(B(t_k) - B(t_{k-1})\right)$$

where $t_k = \frac{kt}{N}$.

What does this actually mean? $W(t)$ is a random variable that can be approximated in the same way that Riemann sums are used to approximate ordinary integrals. The difference is, when $f(t)$ is a non-random function, we use intervals of random width when evaluating the Itô integral. Each $f(t_{k-1})(B(t_k) - B(t_{k-1})$ is a random variable, so their sum, $W(t)$ is a random variable. When f is a random process (meaning W is the integral of process f with respect to Brownian motion B), then $f(t_{k-1})(B(t_k) - B(t_{k-1}))$ is the product of two random variables. This results in another random variable, so the sum $W(t)$ is also a random variable.

If we were working with Riemann integrals, the fundamental theorem of calculus tells us that the derivative of the $W(t)$ (the integrated function) simply returns the original function. Our intuition tells us that we should be able to take the derivative of $W(t)$, the result of our integration, just by reversing the process. Unfortunately, the process of taking the derivative is a little more complicated in stochastic calculus.

The stochastic calculus analogue of the chain rule for evaluating derivatives is Itô's Lemma:

Lemma 7.2 (Itô's Lemma)**.** *Let $f(B_t)$ be a twice-differentiable function and let $B_t = B(t)$ be a Brownian motion.* ***Itô's Lemma*** *states that*

$$df(B_t) = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt$$

Proof. Since the differentiation $\frac{dB_t}{dt}$ doesn't exist, we cannot use the chain rule. Instead, let's start by considering the Taylor expansion of $f(B_t)$:

$$\Delta f(B_t) = f(B_t + \Delta B_t) - f(B_t) = (\Delta B_t) \cdot f'(B_t) + \frac{(\Delta B_t)^2}{2} f''(B_t) + \frac{(\Delta B_t)^3}{3!} f'''(B_t) + \dots$$

[5]We will only briefly introduce concepts of stochastic calculus that are related to Brownian motion. A more rigorous treatment of stochastic calculus would require more measure theory and more concepts that cannot be adequately summarized so as to fit in a short paper such as this one.

[6]For a more detailed construction of the Itô Integral, see `https://www.stat.cmu.edu/~cshalizi/754/notes/lecture-19.pdf`.

Since B_t is a Brownian motion, Wald's second lemma tells us that $\mathbb{E}[(\Delta B_t)^2] = \Delta t$. Substituting this into the previous equation, we get

$$\begin{aligned}\Delta f(B_t) &= (\Delta B_t) \cdot f'(B_t) + \frac{\Delta t}{2} f''(B_t) + \frac{(\Delta B_t)^3}{3!} f'''(B_t) + \dots \\ &= dB_t f'(B_t) + \frac{dt}{2} f''(B_t) + O(\Delta B_t \Delta t)\end{aligned}$$

We can exclude the terms starting from $\frac{(\Delta B_t)^3}{3!} f'''(B_t)$, as $\frac{\Delta B_t \cdot \Delta t}{3!} f'''(B_t) \leq \Delta B_t \cdot \Delta t < \Delta B_t, \Delta t$.

Restating the previous equation in terms of infinitesimals, we get

$$\Delta f(B_t) = dB_t f'(B_t) + \frac{dt}{2} f''(B_t)$$

Rearranging this gives us Itô's lemma:

$$df(B_t) = f'(B_t) dB_t + \frac{1}{2} f''(B_t) dt$$

■

The stochastic calculus equivalent of differential equations are stochastic differential equations.

Definition 7.3. Given that $X = X(t, \omega)$ is a random variable with initial condition $X(0, \omega) = X_0$ with a probability of one, a **stochastic differential equation (SDE)** is an equation of the form

$$dX(t) = f\left(t, X(t)\right) dt + g\left(t, X(t)\right) dW(t) \tag{7.1}$$

When f, g are real-valued functions on $\mathbb{R}^2$, the SDE above is equivalent to saying that $X(t)$ satisfies the following integral equation

$$X(t) = \int_0^t f\big(s, X(s)\big) ds + \int_0^t g\big(s, X(s)\big) dB(s)$$

8. Geometric Brownian Motion and the Black-Scholes Model

One application of Brownian motion can be found in financial mathematics. The Black-Scholes model is the most commonly used stock price behavior estimation model, and it relies on geometric Brownian motion. Broadly speaking, a geometric Brownian motion is a stochastic process in which the logarithm of the variation in the quantity being modeled can be described with a Brownian motion. The definition of geometric Brownian motion reflects this, as it is based on an SDE.

Definition 8.1 (Geometric Brownian Motion)**.** Let $B(t)$ denote a Brownian motion. A stochastic process $S(t)$ is a **geometric Brownian motion (GBM)** if it satisfies

$$\frac{dS}{S} = \mu dt + \sigma dB$$

In the equation above, μ and σ are constants. μ is known as the percentage drift and σ is known as the percentage volatility.

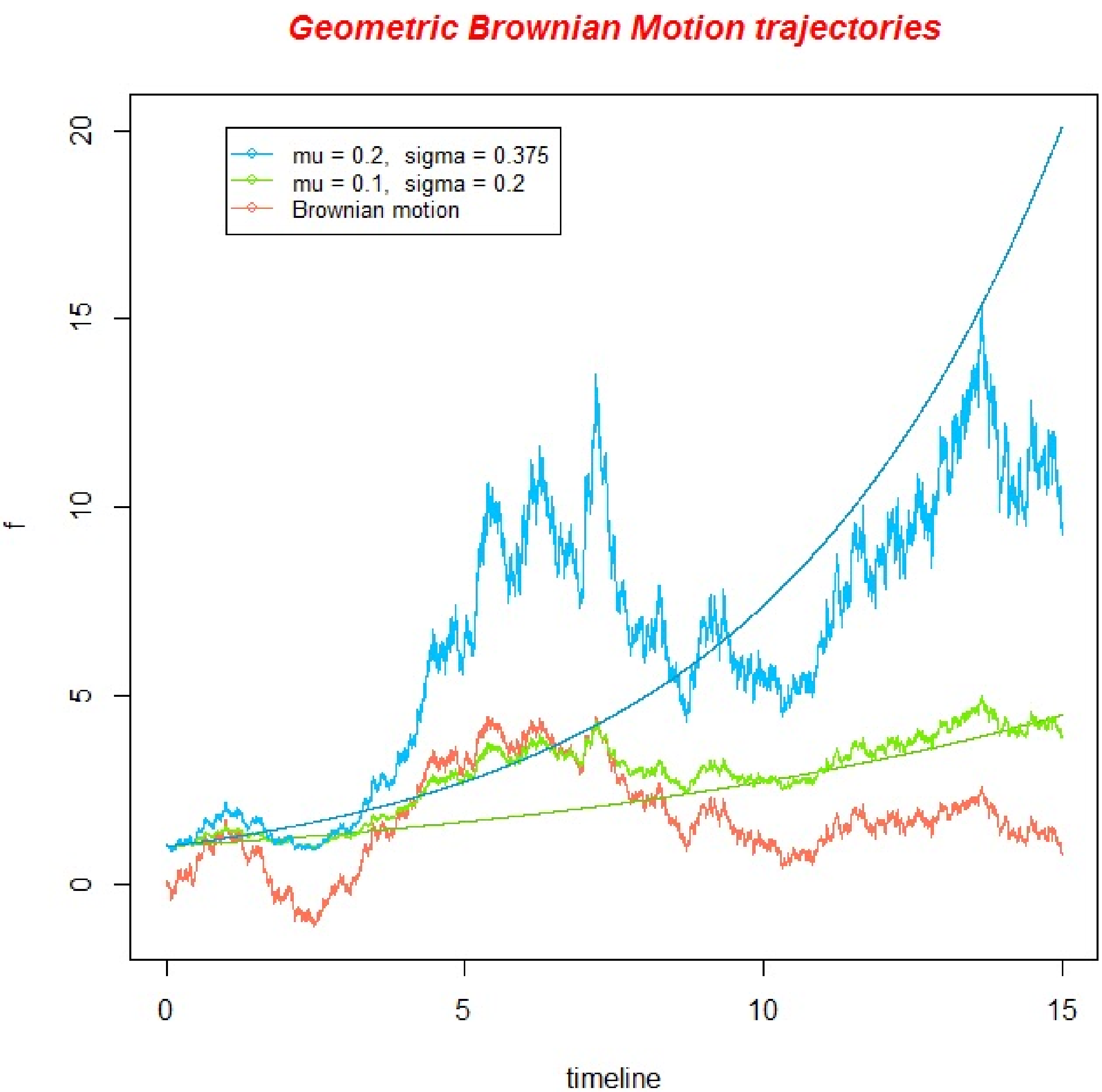

Figure 5. Geometric Brownian motions for different values of μ and σ. Image source: [Wal17].

GBM is excellent for modeling stock prices for several reasons. GBMs are strictly positive, just like real stock prices. Also, the expected return of a GBM process is independent from the current value of the process. This reflects how the expected percentage return[7] required by investors is independent of the stock's price [Hul08].

[7] The expected percentage return (also known as the expected rate of return) is the profit that an investor expects to earn on an investment, represented as a percentage of the initial price of the investment.

8.1. **Black-Scholes Model for Stock Option Pricing.** As this paper is primarily concerned with the mathematical implications of Brownian motion, we shall largely skip over the finance-related background and applications of the Black-Scholes model. However, we will provide a brief explanation of the terms involved in the Black-Scholes equation.

Proposition 8.2. *The Black-Scholes equation states that the option price C for a European-style option*[8] *can be described by*

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0$$

Proof Sketch. Our option price, C, is a function of time, t, and the asset, S (so we have $C = C(S,t)$). We assume that S follows a GBM with drift rate μ and volatility σ. Applying Itô's lemma, we get

$$dC = \frac{\partial C}{\partial t}dt + \frac{\partial C}{\partial S}dS + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}dS^2 \tag{8.1}$$

Since S follows a GBM, we know that

$$\frac{dS}{S} = \mu dt + \sigma dB$$

so we have

$$dS = \mu S dt + \sigma S dB \tag{8.2}$$

Substituting Equation (7.2) into Equation (7.1), we get

$$dC = \frac{\partial C}{\partial t}dt + \frac{\partial C}{\partial S}\left(\mu S dt + \sigma S dB\right) + \frac{1}{2}\frac{\partial^2 C}{\partial S^2}\left(\mu S dt + \sigma S dB\right)^2 \tag{8.3}$$

$$= \left(\frac{\partial C}{\partial t} + \mu S\frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 C}{\partial S^2}\right)dt + \sigma S\frac{\partial C}{\partial S}dB \tag{8.4}$$

Now, consider a delta-hedged portfolio[9]. The value of its holdings is

$$\Pi = -C + \frac{\partial C}{\partial S}S$$

We want to know how much we're making or losing over the time interval $[t, t+\Delta t]$:

$$\Delta\Pi = -\Delta C + \frac{\partial C}{\partial S}\Delta S \tag{8.5}$$

[8] More specifically, a European-style call or put on an underlying stock without dividends. A European-style call option is an agreement that gives the buyer the right to purchase the agreed-upon asset at a specified price during a specified time period, while a European-style put option is an agreement that gives the holder of an asset the right to sell it at a specified price during a specified time interval to the seller of the put.

[9] A delta-hedged portfolio (also known as a delta neutral portfolio) evens out the response to market movements for a certain range so that the net change in the portfolio's value is near zero. For a more detailed explanation of delta hedging, see [Che20].

Taking the discrete version of Equations (7.2) and (7.4), we have

$$\Delta S = \mu S \delta t + \sigma S \Delta B$$
$$\Delta C = \left(\frac{\partial C}{\partial t} + \mu S \frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \right) \Delta t + \sigma S \frac{\partial C}{\partial S} \Delta B$$

Substituting these into Equation (7.5), we get

$$\Delta \Pi = -\left(\left(\frac{\partial C}{\partial t} + \mu S \frac{\partial C}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \right) \Delta t + \sigma S \frac{\partial C}{\partial S} \Delta B \right) + \frac{\partial C}{\partial S} \left(\mu S \delta t + \sigma S \Delta B \right) \tag{8.6}$$

$$= \left(-\frac{\partial C}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \right) \Delta t \tag{8.7}$$

The rate of return on this portfolio must be equivalent to the rate of return on any other riskless investment. This means that if we denote the continuously compounding risk free rate of return as r, we must have

$$r\Pi\Delta t = \Delta\Pi$$

Substituting Equation (7.7) into the equation above, we get

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

This simplifies to the Black-Scholes equation:

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS \frac{\partial C}{\partial S} - rC = 0$$

■

There is an exact formula that gives the solution to the Black-Scholes equation for European options. For an explanation of its derivation and usage, see [Cha].

9. Conclusion

This paper focused on the probabilistic aspects of Brownian Motion and their connections to Markov Chains. There are several more interesting crossovers from other branches of mathematics, like connections to harmonic analysis, which investigates Brownian Motion on a two-dimensional lattice and the Dirichlet problem and applications of the Hausdorff dimension, which is a measure of fractal dimension of sample paths of Brownian motion. This paper doesn't address the most important invariance property: conformal invariance, which allows angle-preserving linear mappings. Brownian motion also has applications in physics through statistical mechanics. Current research using Brownian motion focuses on dynamics and creating better models of moving particles.

References

[bml18] Adapted from post by reddit user omnilogical. a plot of 10,000 random walkers utilizing brownian motion, 2018. [Online; accessed November 20, 2020]. URL: https://www.reddit.com/r/dataisbeautiful/comments/7omtm5/a_plot_of_10000_random_walkers_utilizing_brownian/.

[Cha] Raymond Chan. Chapter 8: Black-scholes equations. URL: https://www.math.cuhk.edu.hk/~rchan/teaching/math4210/chap08.pdf.

[Che20] James Chen. How delta hedging works, Sep 2020. URL: https://www.investopedia.com/terms/d/deltahedging.asp.

[DBS16] Nabyendu Das, Pankaj Bhalla, and Navinder Singh. Memory function approach to correlated electron transport: A comprehensive review. *International Journal of Modern Physics B*, 30, 01 2016. doi:10.1142/S0217979216300152.

[Dun16] Steven R Dunbar. Stochastic processes and advanced mathematical finance, Jul 2016. URL: http://www.math.unl.edu/~sdunbar1/MathematicalFinance/Lessons/BrownianMotion/Definition/definition.pdf.

[Dur19] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[Hul08] John C. Hull. *Chapter 13: The Black-Scholes-Merton Model*, page 277–291. Pearson, 7th edition, 2008.

[McK09] Aaron McKnight. Some basic properties of brownian motion, Aug 2009. URL: http://www.math.uchicago.edu/~may/VIGRE/VIGRE2009/REUPapers/McKnight.pdf.

[MP10] Peter Mörters and Yuval Peres. *Brownian motion*, volume 30. Cambridge University Press, 2010.

[Str64] Volker Strassen. An invariance principle for the law of the iterated logarithm, Sep 1964.

[Wal17] Simulation of a geometric brownian motion in r, Feb 2017. URL: http://delta9hedge.blogspot.com/2013/02/simulation-of-geometric-brownian-motion.html.

[Wik13] Wikipedia, the free encyclopedia. Brownian motion, 2013. [Online; accessed November 20, 2020]. URL: https://en.wikipedia.org/wiki/Markov_property#/media/File:Wiener_process_3d.png.

Euler Circle, Palo Alto, CA 94306