BROWNIAN MOTION

MEHANA ELLIS

Abstract. In this paper, I would like to cover some important topics, theorems, etc. relating to Brownian motion. We will begin by introducing the basic scientific idea behind Brownian motion, then we'll cover the mathematics of it–in particular, Markov processes, Gaussian distribution, the Wiener process, the Lévy characterisation, and martingales. I shall assume basic knowledge of Markov chains and the fundamentals of probability theory.

1. A brief look at the history and science of Brownian motion

In 1827, botanist Robert Brown observed the strange motion of the plant Clarkia pulchella's pollen when placed in water. [Bro28] This fascinating phenomenon, named after Brown, is essentially caused by the random movements of particles in liquid/gas. There are plenty of other examples —coal dust and alcohol, diffusion of calcium through bones, and diffusion of pollutants through the air, to name a few. Around 80 years after Brown first observed this phenomenon, Albert Einstein wrote a paper [1] in which he stated that the particles were being moved by individual water molecules. This was an extremely important paper in science, because Einstein was able to prove the existence of atoms (a debated question originating nearly a century earlier).

Einstein's paper was then used by the physicist Jean Perrin, who proved Dalton's atomic theory[2] using ideas from Einstein's paper and Brownian motion. Brownian motion is occasionally refered to as "pedesis," from Ancient Greek $\pi\eta\delta\eta\sigma\iota\varsigma$, meaning "leaping." This refers to the random movement of the particles. The connection between Brownian motion and random walks should be obvious by now, but we will discuss more about this later in the paper. Let's now turn to the mathematics of Brownian motion.

2. Introducing the mathematics behind Brownian motion

We will begin with a bit of measure-theoretic probability, assuming knowledge of Markov chains and some basic probability. First, let's recall the definition of a σ-algebra, and some related definitions (mostly taken from [Lei]):

Definition 2.1. A σ-algebra on some set S is a subset of the power set $\Sigma \subseteq 2^S$, such that

(1) $S \in \Sigma$;
(2) if $A \in \Sigma$, then $A^c \in \Sigma$;
(3) if $A_1, A_2, \ldots \in \Sigma$, then $\bigcup_{i=1}^{\infty} A_i \in \Sigma$.

Date: November 22, 2020.

[1]Titled "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen." For those interested in the translation, I'd loosely translate this as "On Molecular Kinetic Heat Theory of the movement demanded of particles suspended in static liquids."

[2]A very famous theory in chemistry, essentially stating what we now see as common facts: all matter is composed of atoms, atoms aren't divisible (Greek α (not) + $\tau\epsilon\mu\nu\omega$ (I cut)="atom"), etc.

A *measurable space* consists of a pair (S, Σ), where Σ is a σ-algebra over S. If $\mathcal{C}$ is a collection of subsets of S, then the σ-algebra generated by $\mathcal{C}$, denoted $\sigma(\mathcal{C})$, is the intersection of the σ-algebras on S with $\mathcal{C}$ as a subcollection.

We can also review the definitions of measures/measure spaces and probability measure/space:

Definition 2.2. Let (S, Σ) be a measurable space. A map $\mu : \Sigma \to [0, 1]$ is called a *measure* when $\mu(\{\varnothing\}) = 0$ and is countably additive. So, we have

$$\mu\left(\bigcup_{j=1}^{\infty} F_j\right) = \sum_{j=1}^{\infty} \mu(F_j).$$

Additionally, we call the triple (Ω, Σ, μ) a *measure space*. If X is a function $f : \Omega \to \mathbb{R}$, we say that X is Σ-measurable if $X^{-1}(H) \subseteq \Sigma$ for all $H \in \sigma(\mathbb{R})$. Finally, for a measure space (S, Σ, μ), when $\mu(\Sigma) = 1$, we say this map is a *probability measure* and the associated measure space is called a *probability space*.

Note that going forward, we will be using the more common notation $(\Omega, \mathcal{F}, \mathbb{P})$ for measure spaces, rather than (S, Σ, μ). Another common notation is *i.o.*, which stands for "infinitely often," often when dealing with set-theoretic limits. Let's now recall some standard definitions in measure-theoretic probability.

Definition 2.3. We say that the measure space $(\Omega, \mathcal{F}, \mathbb{P})$ is a *probability triple*, where Ω denotes a *sample space* (and $\omega \in \Omega$ is a *sample point*), $\mathcal{F}$ denotes a σ-algebra called a *family of events*.

Finally, we have stochastic processes:

Definition 2.4. A *stochastic process* is a collection of random variables $\{W_t : t \in \mathcal{T}\}$ on a probability space, where $\mathcal{T}$ is a set of times.

There are some more basic definitions, but we will cover those later. Now we can formally define Brownian motion mathematically:

Definition 2.5. *Brownian motion* started at $x \in \mathbb{R}$ refers to a stochastic process such that the following hold:

(1) $W_0 = x$;
(2) For every $0 \leq s \leq t$, $W_t - W_s$ has normal distribution with mean zero and variance $t - s$, and $|W_t - W_s|$ is independent of $\{W_r : r \leq s\}$;
(3) With probability 1, the function $t \to W_t$ is continuous.

If the Brownian motion begins at 0, we call it *standard Brownian motion*.

There is still more to explore concerning the definition of Brownian motion on the dyadic rationals.

3. Brownian motion on the dyadic rationals

Before we introduce the definitions, let's briefly recall what a dyadic rational is.[3] A dyadic rational is analogous to a 2-adic rational.

In his paper, mathematician Peter Rudzis tells us that "The first ingredient needed for the mathematical description of Brownian motion is Gaussian distribution." [Rud17] This is indeed true; without the definition of Gaussian distribution (which from here on, we shall refer to as *normal distribution*), it is impossible to understand Brownian motion.

Definition 3.1. A random variable X has *normal distribution* with mean μ and variance σ^2 if

$$\mathbb{P}(X > x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_x^\infty e^{\frac{-(u-\mu)^2}{2\sigma^2}} du.$$

We often denote normal distribution as $N(\mu, \sigma^2)$.

Remark 3.2. It is a good idea to understand the definitions thoroughly, but it's also useful to go by the simpler notation (e.g., the notation in the latter definition above) in following definitions. We will mostly use notations without discussing how the actual (often-complicated) definition fits in; nevertheless, we should try to keep the original meanings in mind.

Definition 3.3. We denote the set of non-negative dyadic rationals as $\mathcal{D} = \bigcup_n \mathcal{D}_n$, where $\mathcal{D}_n = \{\frac{k}{2^n} : k = 0, 1, 2 \ldots\}$. A standard, 1-dimensional Brownian motion on the dyadic rationals $\{W_q : q \in \mathcal{D}\}$ is a random process such that for all n, the random variables $W_{k/2^n} - W(k-1)/2^n$, $k \in \mathbb{N}$ are independent and $N(0, \frac{1}{2^n})$.

We will use the following proposition to prove that Brownian motion exists over the dyadic rationals.

Proposition 3.4. *Suppose X and Y are independent normal random variables, each $N(0,1)$. Suppose we have that*

$$Z = \frac{X}{\sqrt{2}} + \frac{Y}{\sqrt{2}}$$

and similarly

$$\hat{Z} = \frac{X}{\sqrt{2}} - \frac{Y}{\sqrt{2}}.$$

Then Z and $\hat{Z}$ are independent $N(0,1)$ variables.

Lemma 3.5. *Standard Brownian motion exists over the dyadic rationals.*

Proof. We will prove this using recursive method. Using Definition 3.3., we can define the following:

$$J(k,n) = 2^{n/2}[W_{k/2^n} - W_{(k-1)/2^n}].$$

Let us assume there exist a countable number of independent normal random variables $\{Z_n : n \in \mathbb{N}\}$. We will use a recursive method to define W_q. Now, for $n = 0$, we have $\{J(k,0) = Z_k : k \in \mathbb{N}\}$. Assume there exists n where $\{J(k,n) : k \in \ltimes\}$ was defined using

[3] I assume the reader's basic knowledge of the p-adic number system. You may want to see https://www.overleaf.com/project/5d7ff7f3b33b1e0001f7bc02 for a quick introduction.

$\{Z_q : q \in \mathbb{D}\}$ and thus are independent $N(0,1)$ variables. This gives us our definition for $J(k, n+1)$:

$$J(2k-1, n+1) = \frac{J(k,n)}{\sqrt{2}} + \frac{Z_{(2k+1)/(2^{n+1}}}{\sqrt{2}}$$
$$J(2k, n+1) = \frac{J(k,n)}{\sqrt{2}} - \frac{Z_{(2k+1)/2^{n+1}}}{\sqrt{2}}.$$

Seeing those $\sqrt{2}$'s in the denominators should remind us of Proposition 3.4; using it repeatedly gives us $\{J(k, n+1) : k \in \mathbb{N}\}$ (independent $N(0,1)$ variables), so we can now define $W_{k/2^n}$ as follows:

$$W_{k/2^n} = 2^{-n/2} \sum_{j=1}^{k} J(j,n).$$

We then have

$$2^{n/2}(W_{k/2^n} - W_{(k-1)/2^n}) = \sum_{j=1}^{k} J(j,n) - \sum_{j=1}^{k-1} J(j,n) = J(k,n),$$

thus proving the existence of standard Brownian motion over the dyadic rationals. [Rud17] □

Here is another lemma, this time about convergence of Brownian motion:

Lemma 3.6. *If W_q, $q \in \mathcal{D}$ is a standard 1-dimensional Brownian motion, then it is almost certain that this function converges uniformly on every closed interval $[a, b]$.*

The proof for this lemma is rather long; it utilizes dyadic rationals, the triangle inequality, the Borel-Cantelli lemma (see Leiner's paper) and involves bounding Brownian motion using integrals. See Leiner's paper for a complete proof. [Lei]

Theorem 3.7. *Standard Brownian motion exists.*

Proof. This is a rather silly statement, for which we can make a fairly straightforward proof. Essentially, this follows from uniform continuity. Let us choose $\frac{\epsilon}{2}$ that yields some δ where $|W_t - W_s| \leq \frac{\epsilon}{2}$ for all $s, t \in \mathcal{D}$. Now pick $n_0 \in \mathbb{N}$ such that $\frac{1}{2^{n_0}} < \delta$. Pick $a \in \mathcal{D}$, and $n, m > n_0$ and $k_{n_0} = 0, 1, \ldots, 2^{n_0}$ such that $0 < a - \frac{k_{n_0}}{2^n} < \frac{1}{2^n}$. We can pick k_n and k_m similarly. We get

$$|k_n - k_{n_0}| < \frac{1}{2^{n_0}}$$
$$|k_m - k_{n_0} < \frac{1}{2^{n_0}}.$$

This gives us $|W_{k_n} - W_{k_m}| < |W_{k_n} - W_{k_{n_0}}| + |W_{k_m} - W_{k_{n_0}} < \epsilon$. We have a Cauchy sequence W_{k_n}, so there is a convergent subsequence with a unique limit. If we define $\{W_t, t \in \mathbb{R}\}$ to be the limit, we get a unique extension of $\{W_q : q \in \mathcal{D}\}$ to $\{W_t : t \in \mathbb{R}\}$ that is continuous. □

One interesting thing about Brownian motion is that, although it is continuous, it is differentiable nowhere. The reader may find Leiner's section on non-differentiability interesting.

4. Brownian motion as a Markov process

Even if one knows only the loosest definition of a Markov chain, one will see that Brownian motion is a perfectly intuitive example of a random walk, which we will recall in the following definition:

Definition 4.1. A *random walk*, loosely defined, is a *stochastic/random* process describing a path consisting of random "steps" on some space.

An intuitive observation about random walks is that they are basically Markov chains, because the $(n+1)^{\text{th}}$ step does not depend on the n^{th} step.

Remark 4.2. Note that many authors may refer to Markov chains as *Markov processes*, but the two are essentially the same. Note that a Markov process might be a more general term (i.e., with a continuous state space and continuous movements), but a Markov chain may assume discreteness in state space or time steps.

Definition 4.3. Let $W_1, \dots, W_d$ be independent Brownian motions started in $x_1, \dots, x_d$, then the random process W_t given by $W_t = (W_1, \dots, W_d)$ is called a *d-dimensional Brownian motion* started in $(x_1, \dots, x_d)$. If W_t starts at the origin, we call this a *standard d-dimensional Brownian motion.*

Definition 4.4. A *filtration* on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a family $\mathcal{F}(t) : t \geq 0)$ of σ-algebras where $\mathcal{F}(S) \subset \sqcup \subset \mathcal{F}$. We call a probability space with a filtration a *filtered probability space*. A random process $\{X_t : t \geq 0\}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is *adapted* if X_t is $\mathcal{F}(t)$ measurable for all $t \geq 0$.

Theorem 4.5 (Simple Markov Property for Brownian motion)**.** *Let $\{W_t : t \geq 0\}$ be a Brownian motion started in $x \in \mathbb{R}^d$. Then the process $\{W_{t+s} - W_s : t, s > 0\}$ is a Brownian motion started at the origin, and it is independent of $\{W_t : 0 \leq t \leq s\}$.*

What this theorem is essentially saying is that we know just as much from the current position as we do from the past positions in Brownian motion. An analogous explanation is that we do not need to know the past positions, we focus entirely on the current one. The latter is the main idea surrounding Markov chains. In addition to the simple Markov property, we also have the strong Markov property. First, let's define the notation $\mathcal{F}^+$:

Definition 4.6. The germ σ-algebra is defined as $\mathcal{F}^+(0)$, where

$$\mathcal{F}^+(t) = \bigcap_{s>t} \mathcal{F}^0(s)$$

and $\{\mathcal{F}^0 : t \geq 0\}$ is the σ-algebra generated by $\{W_t : 0 \leq s \leq t\}$.

Theorem 4.7 (Strong Markov Property for Brownian motion)**.** *For each finite*[4] *stopping time T, we have that the process $\{W_{T+t} - W_T : t \geq 0\}$ is a standard Brownian motion independent of $\mathcal{F}^+ T$.*

Proof. See Leiner's paper for the complete proof. [Lei] □

[4]Almost always, but not absolutely guaranteed to be finite.

5. The Wiener process, another definition of Brownian Motion

Now let's introduce another mathematical model for Brownian motion, called the Wiener process. We will be using the notations a and S, both coming from the following definition:

Definition 5.1. If X is a normal random variable in $\mathbb{R}^n$, we call the notion of one-dimensional variance *covariance*, denoted with the matrix $\text{Cov}(X) = E[XX^T]$. We denote $a = E[X]$.

Definition 5.2. Let $X = \{X_t\}_{t\geq 0}$ is some real valued stochastic process. Then we say X is a *Wiener process* in $\mathbb{R}$ if the following are satisfied:

(1) $X_0 = 0$, or more generally, $X_0 = x$ (meaning that X starts from x).

(2) If $0 \leq t_0 < \cdots < t_m$, then for $1 \leq j \leq m$ the increments $X_{t_j} - X_{t_{j-1}}$ are independent.

(3) If $0 \leq s < t$, then $X_t - X_s \in N_n(a, S)$, where $a = E[X]$ and $S = \text{Cov}(X)$.

(4) The paths of X are almost certainly continuous.

Remark 5.3. Note that this is basically the same as the definition of Brownian motion. The conditions are the same, albeit there being a bit more nuance in the definition of the Wiener process.

6. Introduction to Martingales

Generally speaking, a *martingale* is a sequence of random variables such that, for a given time, the expected value of the next step in the sequence is equal to that of the current step (regardless of the previous steps). This is also a great example of a Markov chain–a sequence in which the n^{th} step is not determined by the $n-k^{\text{th}}$ step. Additionally, we call a process in which we cannot see into the future an *adapted process*. Let's now see the more technical definition:

Definition 6.1. An adapted process $M = \{M_t, t \geq 0\}$ is called a *martingale* with respect to $\mathcal{F}_t$, where $\{\mathcal{F}_t, t \geq 0\}$ denotes a filtration, if the following apply.

(1) For all $t \geq 0$, we have $E(|M_t|) < \infty$.

(2) For each $s \leq t$, we have $E(M_t|\mathcal{F}_s) = M_s$.

A few things to note about about the second condition–first, we can write it as $E(M_t - M_s|\mathcal{F}_s) = 0$. Second, we call M_t a *supermartingale/submartingale* if the second condition is instead $E(M_t|\mathcal{F}_s) \leq M_s$ or $E(M_t|\mathcal{F}_s) \geq M_s$. We have that for any integrable random variable X, $\{E(X|\mathcal{F}_t), t \geq 0\}$ is a martingale. [Nua16]

Martingales give us an interesting theorem concerning stopping times:

Theorem 6.2. *Suppose M_t is a continuous martingale. Let $S \leq T \leq K$ be two bounded stopping times. Then we have*

$$E(M_T|\mathcal{F}_S) = M_S.$$

Furthermore, we simply have that

$$E(M_T) = E(M_S).$$

Proof. Let's show that $E(M_T) = E(M_0)$. Suppose T is some value in the set $0 \leq t_1 \leq \ldots \leq t_n \leq K$. By the martingale property, we have

$$\begin{aligned} E(M_T) &= \sum_{i=1}^{n} E(M_T 1_{\{T=t_i\}}) \\ &= \sum_{i=1}^{n} E(M_{t_i} 1_{\{T=t_i\}}) \\ &= \sum_{i=1}^{n} E(M_{t_n} 1_{\{T=t_i\}}) \\ &= E(M_{t_n}) \\ &= E(M_0). \end{aligned}$$

We can approximate T as follows:

$$\tau_n = \sum_{k=1}^{2^n} \frac{kK}{2^n} 1_{\frac{(k-1)K}{2^n} \leq T < \frac{kK}{2^n}}.$$

We have that $M_{\tau_n} \to M_T$ by continuity. We must now show that M_{τ_N} is integrable:

$$\begin{aligned} E(|M_{\tau_N}| 1_{|M_{\tau_n}| \geq A}) &= \sum_{k=1}^{2^n} E(|M_{\frac{kK}{2^n}}| 1_{\{|M_{\frac{kK}{2^n}}| \geq A, \tau_n = \frac{kK}{2^n}\}}) \\ &\leq \sum_{k=1}^{2^n} E(|M_K| 1_{\{|M_{\frac{kK}{2^n}}| \geq A, \tau_n = \frac{kK}{2^n}\}}) \\ &= E(|M_K| 1_{\{|M_{\tau_n}| \geq A\}}) \\ &\leq E(|M_K| 1_{\{\sup_{0 \leq s \leq K} |M_S| \geq A\}}). \end{aligned}$$

This converges to 0 as $A \to \infty$, uniformly in n. This completes the proof. [Nua16] ☐

Here's another interesting theorem, which we will also give a proof of:

Theorem 6.3. *Let $\{M_t, t \in [0,t]\}$ be a continuous martingale, where $E(|M_T|^p) < \infty$ for some $p \geq 1$. Then, for all $\lambda > 0$, the following holds:*

$$P\left(\sup_{0 \leq t \leq T} |M_t| > \lambda\right) \leq \frac{1}{\lambda^p} E(|M_T|^p).$$

If p is strictly greater than 1, then we have

$$E\left(\sup_{0 \leq t \leq T} |M_t|^p\right) \leq \left(\frac{p}{p-1}\right)^p E(|M_T|^p).$$

Proof. First we look at

$$P\left(\sup_{0 \leq t \leq T} |M_t| > \lambda\right) \leq \frac{1}{\lambda^p} E(|M_T|^p).$$

If we let

$$\tau = \inf\{s \geq 0 : |M_S| \geq \lambda\} \wedge T.$$

Notice that τ is a bounded stopping time, and $|M_t|^p$ is a martingale, so we have

$$E(|M_\tau|^p) \leq E(|M_T|^p.$$

Next, we consider

$$E\left(\sup_{0\le t\le T}|M_t|^p\right)\le\left(\frac{p}{p-1}\right)^p E(|M_T|^p),$$

for $p>1$. We have

$$|M_\tau|^p\ge 1_{\{\sup_{0\le t\le T}|M_t|\ge\lambda\}}\lambda^p+1_{\{\sup_{0\le t\le T}|M_t|<\lambda\}}|M_T|^p$$

by definition of τ. This implies that

$$P\left(\sup_{0\le t\le T}|M_t|>\lambda\right)\le\frac{1}{\lambda^p}E(|M_\tau|^p)\le\frac{1}{\lambda^p}E(|M_T|^p.$$

□

7. Martingales and Brownian motion

Here are some applications of martingales to Brownian motion, in a series of short propositions with quick proofs.

Proposition 7.1. *Let B_t be a Brownian motion. Consider $a\in\mathbb{R}$ and the hitting time*

$$\tau_a=\inf\{t\ge 0:B_t=a\}.$$

If $a<0<b$, then we have

$$P(\tau_a<\tau_b)=\frac{b}{b-a}.$$

Proof. A stopping theorem for martingales states that the expected value of a martingale at a stopping time is equal to its initial expected value. By this theorem, we have

$$E(B_{t\wedge\tau_a})=E(B_0)=0.$$

If we let $t\to\infty$, we have

$$aP(\tau_a<\tau_b)+b(1-P(\tau_a<\tau_b))=0.$$

□

Proposition 7.2. *Let $T=\inf\{t\ge 0:B_t\notin(a,b)\}$, where $a<0<b$. Then we have*

$$E(T)=-ab.$$

Proof. We can use the fact that B_t^2-t is a martingale to see that

$$E(B_{T\wedge t}^2)=E(T\wedge t).$$

Furthermore, we have

$$E(T)=\lim_{t\to\infty}E(B_{T\wedge t}^2)=E(B_T^2)=-ab.$$

□

Here is one more interesting proposition:

Proposition 7.3. *Let $a>0$. Then the hitting time*

$$\tau_a=\inf\{t\ge 0:B_t=a\}$$

satisfies (assuming $\alpha>0$)

$$E[\exp(-\alpha\tau_a)]=e^{-\sqrt{2\alpha}a}.$$

Now, we will finally define the strong Markov property in terms of martingales:

Theorem 7.4. *Let B be a Brownian motion and let T be a finite stopping time where the filtration $\mathcal{F}_t^B$ is generated by B. Then the process*

$$\{B_{T+t} - B_T, t \geq 0\}$$

is a Brownian motion independent of B_T.

Proof. Consider the process $\tilde{B}_t = B_{T+t} - B_T$, and suppose T is bounded. Let $\lambda \in \mathbb{R}$ and $0 \leq s \leq t$. We have

$$E[e^{i\lambda B_{T+t} + \frac{\lambda^2}{2}(T+t)} | \mathcal{F}_{T+s}] = e^{i\lambda B_{T+s} + \frac{\lambda^2}{2}(T+s)}$$

by the optional stopping theorem for the martingale

$$\exp\left(i\lambda\tilde{B}_t + \frac{\lambda^2 t}{2}\right).$$

Therefore, we have

$$E[e^{i\lambda(B_{T+t} - B_{T+s}} | \mathcal{F}_{T+s}] = e^{-\frac{\lambda^2}{2}(t-s)}.$$

[Nua16] □

References

[Bro28] Robert Brown. Xxvii. a brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *The Philosophical Magazine*, 4(21):161–173, 1828.

[Lei] James Leiner. Brownian motion and the strong markov property.

[Nua16] David Nualart. Brownian motion, martingales and markov processes, 2016.

[Rud17] Peter Rudzis. Brownian motion. 2017.