Stochastic Processes and Branching Brownian Motion

by

Abhijeet Anand Mulgund
B.A., Rice University, 2019

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Masters of Science in Mathematics
in the Graduate College of the
University of Illinois at Chicago, 2022

Chicago, Illinois

Defense Committee:
William Perkins, Chair and Advisor
Cheng Ouyang
Marcus Michelen

TODO: A brief dedication to someone you care about. For example, “Dedicated to my cats, Neo and Trinity, who are purrfect in every way.”.

An example of “Dedication” can be found on page 15 of the thesis manual[1].

[1] `http://grad.uic.edu/sites/default/files/pdfs/ThesisManual_rev_06Oct2016.pdf`

ACKNOWLEDGMENT

TODO: A page or two so of shout-outs to people you appreciate. Don't forget your advisor and committee members!

PES

CONTRIBUTIONS OF AUTHORS

This Masters Thesis is a culmination of my studies on various topics of Stochastic Processes, Martingales, and Branching Brownian Motion. It is by no means a comprehensive study, though it hopefully can serve as a resource to others who wish to learn more about these topics.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

CHAPTER **PAGE**

SUMMARY

We will work to first build up the basic theory of Stochastic Processes and Martingales. We will establish conditional expectation, but take most of the basic results and relevant measure theoretic results for granted. From there, we will turn to martingales and start with studying "strategies" on martingales and how they help us prove convergence. From there we will look into decomposition of processes in terms of martingales and a handful of important examples.

With an understanding of martingales, we will then turn to Markov Processes. [[TODO]]

Brownian Motion is both a martingale and markov process, so our prior work will help elucidate some of its properties. Unlike in the previous sections, however, we will not study brownian motion in isolation. Instead, we will examine its deep relationship with a deterministic object, the heat equation.

Finally, we will put together our work from the previous 3 chapters towards a study of Branching Brownian Motion. Branching Brownian motion involves Brownian motions with lifetimes determined by a standard exponential distribution. Upon death, a Brownian motion will split into two Brownian motions, and so on. Like with regular Brownian motion, we will find an interesting connection with diffusion partial differential equations. With a basic understanding of this connection in hand, we will use it to study the distribution of the maximal point of a standard branching Brownian motion.

CHAPTER 1

STOCHASTIC PROCESSES AND MARTINGALES

1.1 Conditional Expectation

Suppose we are given a probability space $(\Omega, \mathcal{B}, \mathbb{P})$ and a random variable $X \in L^1(\mathcal{B})$. We would like to describe the operation of "viewing" this random variable from a sub-sigma algebra $\mathcal{G} \subset \mathcal{B}$. We call this operation *conditional expectation* and define it as follows:

Definition 1.1.1. *The* ***conditional expectation*** *of* X *given* $\mathcal{G}$*, denoted* $\mathbb{E}(X|\mathcal{G})$*, is a random variable with the following properties*

1. $\mathbb{E}(X|\mathcal{G}) \in \mathcal{G}$
2. $\int_A X d\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{G}) d\mathbb{P}$ *for all* $A \in \mathcal{G}$.

The existence and $\mathbb{P}$-a.s. uniqueness of a conditional expectation follows from the following measure-theoretic theorem.

Theorem 1.1.2 (Radon-Nikodym Theorem (1))**.** *Let* $(X, \mathcal{M})$ *be a measure space with sigma-finite nonnegative measure* μ *and sigma-finite signed measure* ν *so that* $\nu << \mu$*. Then there exists* μ*-a.e. unique* $f \in L^1(\mathcal{M})$ *so that for all* $A \in \mathcal{M}$*,*

$$\int_A f d\mu = \nu(A).$$

We denote $\frac{d\nu}{d\mu} := f$.

A thorough proof of the above result can be found in (1). Note that $\nu(A) = \int\limits_A X d\mathbb{P}$ is a signed measure on $(\Omega, \mathcal{B})$ and $\mathbb{P}$ is a finite measure. Restricting ν to $\mathcal{G}$ does not change this fact, and applying 1.1.2 to $\nu\big|_{\mathcal{G}}$ gives us our $\mathbb{P}$-a.s. unique conditional expectation

$$\mathbb{E}(X|\mathcal{G}) = \frac{d\nu\big|_{\mathcal{G}}}{d\mathbb{P}\big|_{\mathcal{G}}}$$

in $L^1(\mathcal{G})$.

Conditional expectation has many useful properties, some of which generalize from expectation:

Proposition 1.1.3. *Let* $X, Y \in L^1(\mathcal{B})$*.* $\mathbb{E}(\cdot|\mathcal{G})$ *has the following properties*

1. *Linearity: if* $c \in \mathbb{R}$*, we have*

$$\mathbb{E}(X + cY|\mathcal{G}) = \mathbb{E}(X|\mathcal{G}) + c\mathbb{E}(Y|\mathcal{G}).$$

2. *Monotonicity: if* $X \leq Y$*, then we have*

$$\mathbb{E}(X|\mathcal{G}) \leq \mathbb{E}(Y|\mathcal{G}).$$

3. *Monotone Convergence: if* $0 \leq X_n \in \mathcal{B}$*,* $X_n \uparrow X$*, we have*

$$\mathbb{E}(X_n|\mathcal{G}) \uparrow \mathbb{E}(X|\mathcal{G}).$$

4. Jensen's inequality: if $\psi : \mathbb{R} \to \mathbb{R}$ *is a convex function, then*

$$\psi(\mathbb{E}(X|\mathcal{G})) \leq \mathbb{E}(\psi(X)|\mathcal{G})$$

5. Smoothing: if $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{B}$*, then*

$$\mathbb{E}(\mathbb{E}(X|\mathcal{F}_2)|\mathcal{F}_1) = \mathbb{E}(X|\mathcal{F}_1) = \mathbb{E}(\mathbb{E}(X|\mathcal{F}_1)|\mathcal{F}_2)$$

6. Known Information: if $X \in \mathcal{G}$*,* $XY \in L^1(\mathcal{B})$*, then*

$$\mathbb{E}(XY|\mathcal{G}) = X\mathbb{E}(Y|\mathcal{G})$$

For any random variable $Z \in \mathcal{B}$, we can define $\mathbb{E}(X|Y) := \mathbb{E}(X|\sigma(Y))$. If we want to condition on the event $[Y = y]$, we can define $\mathbb{E}(X|Y = y) = \mathbb{E}(X|Y)(\omega)$ for any $\omega \in [Y = y]$. For any event $A \in \mathcal{B}$ that is not defined in terms of another random variable, we can define $\mathbb{E}(X|A) = \mathbb{E}(X|1_A = 1)$.

1.2 Regular Conditional Probabilities

TODO: Standard results 1. existence on "nice" spaces 2. relationship to conditional expectation

1.3 Martingales

We start our initial study of martingales by first setting up some underlying concepts. Again, we let $(\Omega, \mathcal{B}, \mathbb{P})$ be our default probability space.

Definition 1.3.1. *Let* T *be a total ordering and let* S, Σ *be a measure space. A* ***stochastic process*** *is a random element* $\mathsf{X} : \Omega, \mathcal{B} \to \mathsf{S}^{\mathsf{T}}, \Sigma^{\mathsf{T}}$*, where* Σ^{T} *denotes the* Σ*-product sigma algebra on* S^{T}*. We call* S *the* ***state space****.*

Normally, however, we think of a stochastic process as some kind random object that progresses through "time". We could alternately have defined it as an indexed collection of random elements $\{\mathsf{X}_{\mathsf{t}} : \Omega, \mathcal{B} \to \mathsf{S}, \Sigma\}_{\mathsf{t}\in\mathsf{T}}$. The definitions are equivalent due to the following proposition

Proposition 1.3.2 ((1))**.** *Let* $(\mathsf{Z}, \mathcal{M})$*,* $(\mathsf{Y}_\alpha, \mathcal{N}_\alpha)$ *be measure spaces for* $\alpha \in \mathsf{A}$*. Let* $\mathsf{Y} = \prod_{\alpha\in\mathsf{A}} \mathsf{Y}_\alpha$ *and let* $\mathcal{N} = \bigotimes_{\alpha\in\mathsf{A}} \mathcal{N}_\alpha$*. Let* $\mathsf{f} : \mathsf{Z} \to \mathsf{Y}$ *and denote* $\mathsf{f}_\alpha := \pi_\alpha \circ \mathsf{f}$*, where* π_α *is the projection map for* $\alpha \in \mathsf{A}$*. Then* f *is* $(\mathcal{M}, \mathcal{N})$*-measureable iff* f_α *is* $(\mathcal{M}, \mathcal{N}_\alpha)$*-measureable for every* $\alpha \in \mathsf{A}$*.*

Therefore, we will use the two definitions interchangeably.

The total ordering in the definition of stochastic process gives us our concept of "time". Usually, $\mathsf{T} = \mathbb{N}$ or $\mathbb{R}$. However, we can imagine more exotic stochastic processes by choosing a total ordering that cannot be order embedded into $\mathbb{R}$. For example, we could consider $\mathbb{R}^2, \preccurlyeq$, endowed with $\preccurlyeq$ is the lexicographic ordering of $\mathbb{R}^2$. Observe that

$$\{\{\mathsf{t}\} \times \mathbb{R} \subset \mathbb{R}^2 | \mathsf{t} \in \mathbb{R}\}$$

is a partition of $\mathbb{R}^2$ into uncountably many nondegenerate $\preccurlyeq$-intervals. Thus, if $\mathbb{R}^2, \preccurlyeq$ could be order embedded into $\mathbb{R}$, then $\mathbb{R}$ would uncountably many nondegenerate intervals, each of which must contain a rational. But then the rationals would be uncountable, giving us a contradiction (2).

A stochastic process defined with this more exotic total ordering would look like a grid of random variables, with each vertical slice representing a stochastic process over $\mathbb{R}$. We could instead choose to think of it as a stochastic process of stochastic processes indexed by $\mathbb{R}$.

A stochastic process on its own does not have much more structure than a random element. We would like to incorporate the idea of measureability with respect to some growing information.

Definition 1.3.3. *A **filtration** is an non-decreasing collection of sub-sigma algebras* $\{\mathcal{F}_\mathsf{t} \subset \mathcal{B}\}_{\mathsf{t}\in\mathsf{T}}$*, indexed by our total ordering* T*.*

Definition 1.3.4. *An **adapted process** with respect to filtration* $\{\mathcal{F}_\mathsf{t}\}_{\mathsf{t}\in\mathsf{T}}$ *is a stochastic process* $\{\mathsf{X}_\mathsf{t}\}_{\mathsf{t}\in\mathsf{T}}$ *so that* $\forall \mathsf{t} \in \mathsf{T}$, $\mathsf{X}_\mathsf{t} \in \mathcal{F}_\mathsf{t}$.

Finally, we are ready to give our definition of a martingale, our concept of a fair game.

Definition 1.3.5. *Let* $\mathsf{X} : \Omega \to \mathbb{R}^\mathsf{T}$ *be an adapted process wrt filtration* $\mathcal{F}_* = (\mathcal{F}_\mathsf{t})_{\mathsf{t}\in\mathsf{T}}$. X *is a **martingale** if it satisfies the following properties:*

1. $\mathsf{X}_\mathsf{t} \in \mathsf{L}^1(\Omega)\ \forall \mathsf{t} \in \mathsf{T}$,
2. $\mathbb{E}(\mathsf{X}_\mathsf{t}|\mathcal{F}_\mathsf{r}) = \mathsf{X}_\mathsf{r}$ *($\mathbb{P}$-a.s.) for any* $\mathsf{r} \leq \mathsf{t}$.

Note that for condition (2) of definition 1.3.5, we need only check that it holds for $\mathsf{r} < \mathsf{t}$, since it already holds by the assumption that X is an adapted process and by the known information property of proposition 1.1.3.

We can also define the related sub(super)-martingale by changing the equality in condition (2) of definition 1.3.5 by a $\geq$ ($\leq$). A submartingale represents a favorable game and a supermartingale represents an unfavorable game. We could have alternatively defined submartingale and supermartingale first, then defined a martingale as an adapted process that is both a submartingale and a supermartingale. We will later prove theorems about submartingales, but they will automatically carry over to martingales. To get the corresponding property for supermartingales, we can just look at the supermartingale's negation to turn it into a submartingale.

When our total ordering $T = \mathbb{N}$, we have the following alternative characterization of a martingale

Proposition 1.3.6. *Let* $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ *be an adapted process wrt filtration* $\mathcal{F}_* = (\mathcal{F}_n)_{n\in\mathbb{N}}$. X *is a martingale if and only if it satisfies the following properties:*

1. $X_n \in L^1(\mathcal{B})\ \forall n \in \mathbb{N}$,
2. $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$ *(*$\mathbb{P}$*-a.s.) for any* $n \in \mathbb{N}$.

Proof. ($\Rightarrow$) Condition (2) of 1.3.6 is just a special case of condition (2) of 1.3.5 when we set $t = n + 1 \geq n = r$.

($\Leftarrow$): This follows by induction. First observe that for $n = 1$, the only $m \in \mathbb{N}$ s.t. $m \leq n$ is $m = 1$. Therefore, $\forall m \leq n$, we have that $\mathbb{E}(X_n|\mathcal{F}_m) = \mathbb{E}(X_1|\mathcal{F}_1) = X_1$. For the inductive step, suppose this claim holds for $n \in \mathbb{N}$. Consider $m \leq n + 1$. If $m = n + 1$, then $\mathbb{E}(X_{n+1}|\mathcal{F}_m) = \mathbb{E}(X_{n+1}|\mathcal{F}_{n+1}) = X_{n+1}$, so we're good. On the other hand, if $m < n + 1$, then we have that

$$\mathbb{E}(X_{n+1}|\mathcal{F}_m) = \mathbb{E}(\mathbb{E}(X_{n+1}|\mathcal{F}_n)|\mathcal{F}_m) = \mathbb{E}(X_n|\mathcal{F}_m) = X_m$$

by the *smoothing* property of proposition 1.1.3 and induction. ∎

We refer to such martingales as **discrete-time** martingales.

TODO: - ~~definition of stochastic process~~ - ~~remark about total ordering~~ - ~~definition of adapted process and filter~~ - ~~definition of martingale (total ordering), super, sub~~ - ~~connection to natural number definition~~

1.4 Decomposition of Discrete-Time martingales

A useful concept for discrete-time martingales is a predictable process

Definition 1.4.1. *Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be an adapted process wrt filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$. X is a* ***predictable process*** *if for all $n \in \mathbb{N}$, $X_{n+1} \in \mathcal{F}_n$*

This concept is capturing the idea that we can "predict" the next value of X using the information we have now. Indeed, if at timestep n we can discern between whether a given event in $\mathcal{F}_n$ occured, then for any $x \in \mathbb{R}$ we can determine whether or not $X = x$ since $[X = x] \in \mathcal{F}_n$.

The following proposition reveals that we can think of any adapted process indexed by $\mathbb{N}$ as a predictable process with some "fair" noise.

Proposition 1.4.2. *Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a discrete-time adapted process wrt filtration $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$. Let $X_0 := 0$ and $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Then there exists a martingale $(M_n)_{n\geq 0}$ and a predictable process $(A_n)_{n\geq 1}$ so that for $n \geq 1$*

$$X_n = M_n + A_n$$

Before we proceed with the proof of proposition 1.4.2, we first establish a useful concept and lemma.

Definition 1.4.3. *Let* $d : \Omega \to \mathbb{R}^{\mathbb{N}}$ *be a discrete-time process adapted to filtration* $(\mathcal{F}_n)_{n\in\mathbb{N}}$*. Then, it is a* ***martingale difference sequence*** *if*

1. $d_j \in L^1(\mathcal{B})\ \forall j \in \mathbb{N}$
2. $\mathbb{E}(d_{j+1}|\mathcal{B}_j) = 0\ \forall j \in \mathbb{N}$

Lemma 1.4.4. $(M_n)_{n\geq 0}$ *is a martingale adapted to filtration* $(\mathcal{F}_n)_{n\geq 0}$ *iff there exists a martingale difference sequence* $d : \Omega \to \mathbb{R}^{\mathbb{N}}$ *adapted to the same filtration so that*

$$M_n := M_0 + \sum_{j=1}^{n} d_j,$$

for all $n \in \mathbb{N}$*.*

Proof. ($\Rightarrow$): Let $d_n := M_n - M_{n-1}$ for $n \geq 1$. Then for $n \geq 1$, the sum $M_0 + \sum_{j=1}^{n} d_j = M_0 + \sum_{j=1}^{n} M_j - M_{j-1}$ is telescoping and is equal to M_n.

Now we need simply show that $(d_n)_{n\geq 1}$ is a martingale difference sequence. Checking condition (1) of 1.4.3, we see for $n \geq 1$ that $\mathbb{E}|d_n| = \mathbb{E}|M_n - M_{n-1}| \leq \mathbb{E}|M_n| + \mathbb{E}|M_{n-1}| < \infty$, so $d_n \in L^1(\mathcal{B})$.

Checking condition (2), we see that

$$\mathbb{E}(d_{n+1}|\mathcal{F}_n) = \mathbb{E}(M_{n+1}|\mathcal{F}_n) - \mathbb{E}(M_n|\mathcal{F}_n)) = M_n - M_n = 0.\checkmark$$

First checking for $n \in \mathbb{N}$ that $M_n \in L^1(\mathcal{B})$, consider

$$\begin{aligned}\mathbb{E}|M_n| &= \mathbb{E}\left|\sum_{j=1}^{n} d_j\right| \\ &\leq \sum_{j=1}^{n} \mathbb{E}|d_j| \\ &< \infty\end{aligned}$$

since $d_j \in L^1(\mathcal{B})\ \forall j \in \mathbb{N}$.

Next we check condition (2) of definition 1.3.5. $\square$

TODO: - predictable process definition - decomp of adapted process into predictable process and martingale - Doob's Decomposition

1.5 Strategies and Discrete-Time martingales

TODO: - Definition - preservation of fairness - relationship to stopping times

1.6 Convergence of Discrete-time Martingales

TODO: - Upcrossing inequality - convergence of discrete-time martingales - pathological examples

1.7 Cotninuous-time martingales, definitions

- indistinguishable - section 1.1 Karatzas shreve

1.8 Continuous-time martingales, stopping times

-section 1.2 Karatzas shreve

1.9 Convergence of Continuous-time Martingales

- section 1.3 karatzas shreve - submartingale inequalities (Karatzas thm 3.8) - convergence of right-continuous martingales

1.10 Polya's Urn

1.11 Borel-Cantelli Lemma

1.12 Differential Equation Method?

TODO: - Azuma's inequality - 2nd case study with graphs

CHAPTER 2

MARKOV PROCESSES

2.1 Basics

CHAPTER 3

BROWNIAN MOTION AND THE HEAT EQUATION

3.1 Basics

- Definition - Existence Remark - Martingale - Markov Process

3.2 Heat Equation

3.3 Brownian Motion as a Solution

CHAPTER 4

BRANCHING BROWNIAN MOTION

4.1 Basics

TODO: - definition

4.2 F-KPP Equation

- F-KPP equation

4.3 McKean Representation

- Mckean Representation

4.4 Kolmogorov's Result

CITED LITERATURE

1. Folland, G. B.: Real analysis: Modern techniques and their applications. "John Wiley & Sons", 2011.

2. (https://math.stackexchange.com/users/39362/haskell curry), H. C.: Is the set of real numbers the largest possible totally ordered set? Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/255152 (version: 2013-02-10).

VITA

NAME	Sheena The Punk Rocker
EDUCATION	B.A., Composition, Rock n' Roll High School, New York City, NYC, 1976
TEACHING	Music lessons (CS342, Summer 1980)
PUBLICATIONS	Author 1, Author 2, and Author 3. "Paper title 1." In Proceedings of the CONFERENCE NAME (CONFERENCE ABBR, YEAR).
	Author 1, Author 2, and Author 3. "Paper title 2." In Proceedings of the CONFERENCE NAME (CONFERENCE ABBR, YEAR).