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. $\int\limits_A X d\mathbb{P} = \int\limits_A \mathbb{E}(X|\mathcal{G}) d\mathbb{P}$ *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* $A \in \mathcal{M}$*,*

$$\int\limits_A f d\mu = \nu(A).$$

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

A thorough proof of the above result can be found in (1). Note that $\nu(A) = \int_A X d\mathbb{P}$ 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, Σ 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 $\{\mathsf{X}_t : \Omega, \mathcal{B} \to \mathsf{S}, \Sigma\}_{t \in \mathsf{T}}$. 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

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

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 disjoint 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* $\{\mathcal{F}_t \subset \mathcal{B}\}_{t \in T}$*, indexed by our total ordering* T*.*

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

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

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

1. $X_t \in L^1(\mathcal{B})\ \forall t \in T$,
2. $\mathbb{E}(X_t | \mathcal{F}_r) = X_r$ *($\mathbb{P}$-a.s.) for any* $r \leq t$.

Note that for condition (2) of definition 1.3.5, we need only check that it holds for $r < t$, since it already holds for $r = t$ 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 to 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.

Note that by applying *smoothing* and the identity $\mathbb{E}(Y) = \mathbb{E}(Y|\{\emptyset, \Omega\})$, we get that for any $r \leq t \in T$,

$$\mathbb{E}(X_t) = \mathbb{E}(X_t|\{\emptyset, \Omega\}) = \mathbb{E}(\mathbb{E}(X_t|\mathcal{F}_r)|\{\emptyset, \Omega\}) = \mathbb{E}(X_r|\{\emptyset, \Omega\}) = \mathbb{E}(X_r)$$

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 $X : \Omega \to S^{\mathbb{N}}$ be an adapted process wrt filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ and state space S, Σ. X is a **predictable process** if for all $n \in \mathbb{N}$, $X_n \in \mathcal{F}_{n-1}$, with $\mathcal{F}_0 = \{\emptyset, \Omega\}$.*

This concept is capturing the idea that we can "predict" the next value of a stochastic process X using the information we have now. Indeed, for a real-valued predictable process, 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 $X_{n+1} = x$ since $[X_{n+1} = x] \in \mathcal{F}_n$. By setting $\mathcal{F}_0 = \{\emptyset, \Omega\}$, we enforce that X_1 is a constant.

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* $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$ *so that* $X_n \in L^1(\mathcal{B})$ *for all* $n \geq 1$*. Then there exists a martingale* $(M_n)_{n\geq 1}$ *and a predictable process* $(A_n)_{n\geq 1}$ *with* $A_1 = 0$ *so that for* $n \geq 1$

$$X_n = M_n + A_n.$$

This decomposition is $\mathbb{P}$*-a.s. unique.*

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. $(M_n)_{n\geq 1}$ *is a martingale adapted to filtration* $\mathcal{F}_* = (\mathcal{F}_n)_{n\geq 1}$ *iff there exists a martingale difference sequence* $d : \Omega \to \mathbb{R}^{\mathbb{N}}$ *adapted to the same filtration and* $c \in \mathbb{R}$ *so that*

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

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

Proof. ($\Rightarrow$): Let $d_n := M_n - M_{n-1}$ for $n \geq 1$, where $M_0 = \mathbb{E}(M_1)$. 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. First note that $(d_n)_{n\geq 1}$ is adapted to $\mathcal{F}_*$ since $M_n - M_{n-1} \in \mathcal{F}_n$ for all $n \geq 1$. 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

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

Both conditions are satisfied so $(d_n)_{n\geq 1}$ is a martingale difference sequence. ✓

($\Leftarrow$): We are given difference sequence $(d_n)_{n\geq 1}$ and constant $c \in \mathbb{R}$. Note that since $(d_n)_{n\geq 1}$ is adapted to $\mathcal{F}_*$, we know $M_n = c + \sum_{j=1}^{n} d_j \in \mathcal{F}_n$ for all $n \geq 0$, so $(M_n)_{n\geq 1}$ is adapted to $\mathcal{F}_*$. Now, checking for $n \in \mathbb{N}$ that $M_n \in L^1(\mathcal{B})$, consider

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

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

Next we check condition (2) of definition 1.3.5. Letting $n \in \mathbb{N}$

$$
\begin{aligned}
\mathbb{E}(M_{n+1}|\mathcal{F}_n) &= c + \sum_{j=1}^{n+1} \mathbb{E}(d_j|\mathcal{F}_n) \\
&= \mathbb{E}(d_{n+1}|\mathcal{F}_n) + c + \sum_{j=1}^{n} d_j \\
&= c + \sum_{j=1}^{n} d_j && ((d_n)_{n\geq 1}\text{is a difference sequence}) \\
&= M_n.
\end{aligned}
$$

Both conditions are satisfied, so $(M_n)_{n\geq 1}$ is a martingale. ✓ □

Note that our proof above also gives a procedure to construct a martingale difference sequence from a martingale. Now we proceed to the proof of proposition 1.4.2

Proof of Proposition 1.4.2. We will deconstruct $(X_n)_{n\geq 1}$ into

$$
\begin{aligned}
M_n &:= X_0 + \sum_{j=1}^{n} X_j - \mathbb{E}(X_j|\mathcal{F}_{j-1}) \\
A_n &:= \sum_{j=1}^{n} \mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1}
\end{aligned}
$$

where we let $\mathcal{F}_0 := \{\emptyset, \Omega\}$ and $X_0 = \mathbb{E}(X_1|\mathcal{F}_0) = \mathbb{E}(X_1)$. Observe that $A_1 = \mathbb{E}(X_1|\mathcal{F}_0) - X_0 = \mathbb{E}(X_1) - \mathbb{E}(X_1) = 0$. With this choice of $(M_n)_{n\geq 1}$ and $(A_n)_{n\geq 1}$ we have that

$$\begin{aligned} M_n + A_n &= X_0 + \sum_{j=1}^{n} X_j - \mathbb{E}(X_j|\mathcal{F}_{j-1}) + \sum_{j=1}^{n} \mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1} \\ &= X_0 + \sum_{j=1}^{n} X_j - \mathbb{E}(X_j|\mathcal{F}_{j-1}) + \mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1} \\ &= X_0 + \sum_{j=1}^{n} X_j - X_{j-1} \\ &= X_n. \end{aligned}$$

Next we show that $(M_n)_{n\geq 1}$ is a martingale and $(A_n)_{n\geq 1}$ is a predictable sequence. To that end, observe that for $n \geq 1$:

1. $X_n - \mathbb{E}(X_n|\mathcal{F}_{n-1}) \in \mathcal{F}_n$ since $(X_n)_{n\geq 1}$ is an adapted process and $\mathcal{F}_{n-1} \subset \mathcal{F}_n$.
2. We are given that $X_n \in L^1(\mathcal{B})$ and $\mathbb{E}(X_n|\mathcal{F}_{n-1}) \in L^1(\mathcal{B})$ by definition, so $X_n - \mathbb{E}(X_n|\mathcal{F}_{n-1}) \in L^1(\mathcal{B})$.
3. $\mathbb{E}(X_n - \mathbb{E}(X_n|\mathcal{F}_{n-1})|\mathcal{F}_{n-1}) = \mathbb{E}(X_n|\mathcal{F}_{n-1}) - \mathbb{E}(X_n|\mathcal{F}_{n-1}) = 0$

so $X_n - \mathbb{E}(X_n|\mathcal{F}_{n-1})$ is a martingale difference sequence. By lemma 1.4.4, noting that $X_0 = \mathbb{E}(X_1)$ is a constant, we have that $M_n = X_0 + \sum_{j=1}^{n} X_j - \mathbb{E}(X_j|\mathcal{F}_{j-1})$ is a martingale.

On the other hand, for all $j \geq 1$, we know that $\mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1} \in \mathcal{F}_j$ since $(X_n)_{n\geq 1}$ is an adapted process. Because $\mathcal{F}_*$ is a filtration, for all $m \geq j$, $\mathcal{F}_j \subset \mathcal{F}_m$. Therefore, for all $n \geq 1$, $A_n = \sum_{j=1}^{n} \mathbb{E}(X_j|\mathcal{F}_{j-1}) - X_{j-1} \in \mathcal{F}_n$ and it is predictable.

Finally we show that this decomposition is (almost surely) unique. Suppose there existed another decomposition with the same properties: $X_n = N_n + B_n$. We will apply induction to show that these decompositions are almost surely the same. For our base case, let $n = 1$. We are given that $A_1 = B_1 = 0$. Therefore, after subtracting the two decompositions, $0 = M_1 + A_1 - N_1 + B_1 = M_1 - N_1$, so $M_1 = N_1$. For our inductive step, assume for some $n \geq 1$ that $M_n = N_n$ and $A_n = B_n$ almost surely. We know that $0 = M_{n+1} + A_{n+1} - N_{n+1} - B_{n+1}$. Taking conditional expectations of both sides conditioned on $\mathcal{F}_n$, we get

$$\begin{aligned}
0 &= \mathbb{E}(M_{n+1} + A_{n+1} - N_{n+1} - B_{n+1}|\mathcal{F}_n) \\
&= M_n - N_n + \mathbb{E}(A_{n+1}|\mathcal{F}_n) - \mathbb{E}(B_{n+1}|\mathcal{F}_n) && ((M_n)_{n\geq 1}, (N_n)_{n\geq 1} \text{ are martingales}) \\
&= M_n - N_n + A_{n+1} - B_{n+1} && ((A_n)_{n\geq 1}, (B_n)_{n\geq 1} \text{ are predictable}) \\
&= A_{n+1} - B_{n+1} && (\text{inductive hypothesis})
\end{aligned}$$

Therefore $A_{n+1} = B_{n+1}$ almost surely. Then $0 = M_{n+1} + A_{n+1} - N_{n+1} - B_{n+1} = M_{n+1} - N_{n+1}$, so $M_{n+1} = N_{n+1}$ almost surely, proving the inductive hypothesis for $n + 1$. Therefore, by induction, we have that $M_{n+1} + A_{n+1}$ is the almost surely unique decomposition of X_n for $n \geq 1$ with the desired properties. $\square$

We can interpret proposition 1.4.2 as saying any adapted process is some kind of "trend" with "fair" additive noise applied on top. This result naturally gives us some well-known decomposition results.

Corollary 1.4.5 (Doob's Submartingale Decomposition)**.** *Let* $(Y_n)_{n\geq 1}$ *be a discrete-time adapted process wrt filtration* $\mathcal{F}_* := (\mathcal{F}_n)_{n\geq 1}$ *so that* $Y_n \in L^1(\mathcal{B})$ *for all* $n \geq 1$*. Then it is a submartingale iff there exists a martingale* $(M_n)_{n\geq 1}$ *and a non-decreasing predictable process* $(A_n)_{n\geq 1}$ *with* $A_1 = 0$ *so that for* $n \geq 1$

$$Y_n = M_n + A_n$$

This decomposition is almost surely unique.

Proof. ($\Rightarrow$): Apply proposition 1.4.2 to get our almost surely unique martingale $(M_n)_{n\geq 1}$ and predictable process $(A_n)_{n\geq 1}$ so that $Y_n = M_n + A_n$ for $n \geq 1$. We must simply show that A_n is non-decreasing. To that end, observe that for $n \geq 2$

$$\begin{aligned}
0 &\leq \mathbb{E}(Y_n|\mathcal{F}_{n-1}) - Y_{n-1} \\
&= \mathbb{E}(M_n + A_n|\mathcal{F}_{n-1}) - A_{n-1} \\
&= \mathbb{E}(M_n|\mathcal{F}_{n-1}) - M_{n-1} + \mathbb{E}(A_n|\mathcal{F}_{n-1}) - A_{n-1} \\
&= M_{n-1} - M_{n-1} + \mathbb{E}(A_n|\mathcal{F}_{n-1}) - A_{n-1} && ((M_n)_{n\geq 1}\text{is a martingale}) \\
&= A_n - A_{n-1}, && ((A_n)_{n\geq 1}\text{is a predictable process})
\end{aligned}$$

giving us that $(A_n)_{n\geq 1}$ is non-decreasing. ✓

($\Leftarrow$): We must show that $Y_n = M_n + A_n$ for $n \geq 1$ is a submartingale. We only need to check condition (2) since we are given $Y_n \in L^1(\mathcal{B})$. Checking, for $n \geq 1$, we see

$$\begin{aligned}
\mathbb{E}(Y_{n+1}|\mathcal{F}_n) &= \mathbb{E}(M_{n+1} + A_{n+1}|\mathcal{F}_n) \\
&= \mathbb{E}(M_{n+1}|\mathcal{F}_n) + A_{n+1} \\
&= M_n + A_{n+1} && ((M_n)_{n\geq 1} \text{ is a martingale}) \\
&\geq M_n + A_n && ((A_n)_{n\geq 1} \text{ is non-decreasing}) \\
&= Y_n,
\end{aligned}$$

so $(Y_n)_{n\geq 1}$ is indeed a submartingale. ✓ □

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

1.5 Strategies and Discrete-Time martingales

To further analyze martingales, we introduce the concept of a "strategy" for playing the "fair game", where we our strategy determines our move at the end of each turn.

Definition 1.5.1. *Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a discrete-time adapted process wrt filtration $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$ and let $H : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a predictable process on the same filtration. Then we call H a **strategy** for X with **winnings***

$$(H \cdot X)_n := \sum_{m=1}^{n} H_m(X_m - X_{m-1})$$

for $n \geq 1$ *(where* $X_0 = 0$*).*

The following proposition will establish that we cannot change the fairness of a game by choosing the "best" (or "worst") strategy.

Proposition 1.5.2. *Let* $(Y_n)_{n\geq 1}$ *be a discrete-time martingale wrt filtration* $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$ *and let* $(H_n)_{n\geq 1}$ *be bounded strategy for* $(Y_n)_{n\geq 1}$*. Then* $((H \cdot Y)_n)_{n\geq 1}$ *is a martingale.*

Proof. First note that since $Y_0, Y_1, \ldots, Y_n, H_1, \ldots, H_n \in \mathcal{F}_n$, we have that $(H \cdot Y)_n \in \mathcal{F}_n$, so $(H \cdot Y)_n$ is adapted to $\mathcal{F}_*$.

Next note that for $n \geq 1$ because H_n is bounded, say by $M \in \mathbb{R}_{\geq 0}$

$$
\begin{aligned}
\mathbb{E}|(H \cdot Y)_n| &= \mathbb{E}\left|\sum_{m=2}^{n} H_m(Y_m - Y_{m-1})\right| \\
&\leq M \sum_{m=1}^{n} \mathbb{E}|Y_m - Y_{m-1}| && \text{(Triangle Inequality)} \\
&\leq M \sum_{m=1}^{n} \mathbb{E}|Y_m| + \mathbb{E}|Y_{m-1}| && \text{(Triangle Inequality)} \\
&< \infty
\end{aligned}
$$

since $Y_m \in L^1(\mathcal{B})$ $\forall m \geq 0$. So $(H \cdot Y)_n \in L^1(\mathcal{B})$ as well. ✓

Observe that for $n \geq 1$

$$
\begin{aligned}
\mathbb{E}((H \cdot Y)_{n+1} | \mathcal{F}_n) &= \mathbb{E}(\sum_{m=1}^{n+1} H_m(Y_m - Y_{m-1}) | \mathcal{F}_n) \\
&= \sum_{m=1}^{n+1} \mathbb{E}(H_m(Y_m - Y_{m-1}) | \mathcal{F}_n) \\
&= \sum_{m=1}^{n} H_m(Y_m - Y_{m-1}) + \mathbb{E}(H_{n+1}(Y_{n+1} - Y_n) | \mathcal{F}_n) \\
&= (H \cdot Y)_n
\end{aligned}
$$

since $(Y_n)_{n \geq 1}$ is a martingale, so

$$
\begin{aligned}
\mathbb{E}(H_{n+1}(Y_{n+1} - Y_n) | \mathcal{F}_n) &= H_{n+1}(\mathbb{E}(Y_{n+1} | \mathcal{F}_n) - Y_n) \\
&= 0. \checkmark
\end{aligned}
$$

Therefore, we have that $((H \cdot Y)_n)_{n \geq 1}$ is a martingale. $\square$

Note that the boundedness of our strategy was only used to ensure the winnings process is integrable. We could replace the boundedness condition with alternate conditions. For example, if we had $H_n, Y_n \in L^2(\mathcal{B})$ $\forall n \in \mathbb{N}$, then we could apply the cauchy-schwarz inequality to get that $(H \cdot Y)_n \in L^1(\mathcal{B})$.

Using the Doob decomposition, we can easily generalize this result to submartingales (and therefore supermartingales by negation):

Corollary 1.5.3. *Let* $(Y_n)_{n\geq 1}$ *be a discrete-time submartingale wrt filtration* $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$ *and let* $(H_n)_{n\geq 1}$ *be non-negative bounded strategy for* $(Y_n)_{n\geq 1}$*. Then* $((H\cdot Y)_n)_{n\geq 1}$ *is a submartingale.*

Proof. Using corollary 1.4.5, decompose $(Y_n)_{n\geq 1}$ into $(M_n + A_n)_{n\geq 1}$, where $(M_n)_{n\geq 1}$ is a martingale and $(A_n)_{n\geq 1}$ is a predictable process with $A_1 = 0$. Denote $M_0 = A_0 = 0$ to make $Y_0 = M_0 + A_0 = 0$. Then,

$$\begin{aligned}
(H \cdot Y)_n &= \sum_{m=1}^{n} H_m(Y_m - Y_{m-1}) \\
&= \sum_{m=1}^{n} H_m(M_m + A_m - M_{m-1} - A_{m-1}) \\
&= \sum_{m=1}^{n} H_m(M_m - M_{m-1}) + \sum_{m=1}^{n} H_m(A_m - A_{m-1}) \\
&= (H \cdot M)_n + (H \cdot A)_n
\end{aligned}$$

We know by proposition 1.5.2 that $((H \cdot M)_n)_{n\geq 1}$ is a martingale. Furthermore, we have that

1. $((H \cdot A)_n)_{n\geq 1}$ is a predictable process because $(H_n)_{n\geq 1}$ and $(A_n)_{n\geq 1}$ are predictable processes.

2. $(H \cdot A)_1 = H_1 A_1 = 0$.

3. For all $n \geq 1$, $(H\cdot A)_{n+1} = (H\cdot A)_n + H_{n+1}(A_{n+1} - A_n) \geq (H\cdot A)_n$ because $H_{n+1} \geq 0$ and $A_{n+1} \geq A_n$ (it is non-decreasing). Thus $((H \cdot A)_n)_{n\geq 1}$ is a non-decreasing sequence.

Therefore, by the converse case of corollary 1.4.5, we have that $((H\cdot Y)_n)_{n\geq 1} = ((H\cdot M)_n + (H \cdot A)_n)_{n\geq 1}$ is a submartingale. $\square$

Corresponding results hold for submartingales and martingales. If $(Y_n)_{n\geq 1}$ is a martingale, then bounded $(H_n)_{n\geq 1}$ will suffice, it does not need to be non-negative.

We can look at playing the game up to a stopping point as a special kind of strategy. The classical definition of a "stopping time" is as follows

Definition 1.5.4. *Let $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration. A stopping time on $\mathcal{F}_*$ is a random variable $\nu : \Omega \to \bar{\mathbb{N}}$ so that for each $n \in \mathbb{N}$, $[\nu = n] \in \mathcal{F}_n$. Here, $\bar{\mathbb{N}}$ is the extended natural numbers.*

There are a couple equivalent conditions for being a stopping time.

Proposition 1.5.5. *Suppose $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$ is a filtration on Ω. Let $\nu : \Omega \to \bar{\mathbb{N}}$. The following conditions are equivalent:*

1. *ν is a stopping time.*
2. *$[\nu \leq n] \in \mathcal{F}_n, \forall n \in \mathbb{N}$*
3. *$[\nu > n] \in \mathcal{F}_n, \forall n \in \mathbb{N}$*

Proof. $(1) \Rightarrow (2)$: We have for $n \in \mathbb{N}$

$$[\nu \leq n] = \bigcup_{m\leq n} [\nu = m] \in \mathcal{F}_n$$

since $[\nu = m] \in \mathcal{F}_m \subset \mathcal{F}_n$ $\forall m \leq n$. ✓

$(2) \Rightarrow (3)$: Let $n \in \mathbb{N}$. Since $[\nu \leq n] \in \mathcal{F}_n$, we know by definition of a sigma-algebra that $[\nu > n] = [\nu \leq n]^C \in \mathcal{F}_n$. ✓

(3) $\Rightarrow$ (1): Let $n \in \mathbb{N}$. Since $[\nu > n] \in \mathcal{F}_n$ and because $\mathcal{F}_n$ forms a filtration, we have that

$$[\nu = n] = [\nu > n-1] \cap ([\nu > n])^C \in \mathcal{F}_n,$$

where $[\nu > 0] = \Omega$. ✓ □

We can alternatively view a stopping time as a special kind of strategy:

Proposition 1.5.6. *Suppose $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$ is a filtration on Ω. Let $\nu : \Omega \to \bar{\mathbb{N}}$. If ν is a stopping time, then $(1_{\nu\geq n})_{n\geq 1}$ is a predictable process.*

Proof. We need only check that $1_{\nu\geq n} \in \mathcal{F}_{n-1}$ for $n \geq 1$ (with $\mathcal{F}_0 = \{\emptyset, \Omega\}$). To that end we have $1_{\nu\geq 0} = 1$ since $\nu(\omega) \in \bar{\mathbb{N}}$ for all $\omega \in \Omega$, so $1_{\nu\geq 0} \in \mathcal{F}_0$. For $n \geq 1$, we have that $1_{\nu\geq n} = 1 - 1_{\nu\leq n-1} \in \mathcal{F}_{n-1}$ since $[\nu \leq n-1] \in \mathcal{F}_{n-1}$ by proposition 1.5.5. □

For a adapted process $X : \Omega \to \mathbb{R}^{\mathbb{N}}$, we can denote the process resulting from stopping the game after ν turns as:

$$X_{n\wedge\nu} := (1_{\nu\geq\cdot} \cdot X)_n = \sum_{m=1}^{n} 1_{\nu\geq n}(X_m - X_{m-1}).$$

Combining with our result about strategies, we get the following result:

Corollary 1.5.7. *Let $(X_n)_{n\geq 1}$ be a discrete-time (sub/super)martingale wrt filtration $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$ and let ν be stopping time on $\mathcal{F}_*$. Then $(X_{n\wedge\nu})_{n\geq 1}$ is a (sub/super)martingale.*

Proof. This follows from the definition of $X_{n\wedge\nu}$ and application of proposition 1.5.2 or corollary 1.5.3. □

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

1.6 Convergence of Discrete-time Martingales

So far we have concerned ourselves with properties of martingales without looking at any limiting behavior. Under the right conditions, we will find that discrete-time martingales do converge to a random variable. To establish this, we'll take approach of analyzing upcrossings.

Proposition 1.6.1. *Let $(\mathcal{F}_n)_{n=1}^{\infty}$ be a filtration and let $(X_n)_{n=1}^{\infty}$ be adapted to $(\mathcal{F}_n)_{n=1}^{\infty}$. Let $a < b \in \mathbb{R}$. Then the following random variables are stopping times for $k \geq 1$:*

$$N_{2k-1} := \inf\{m > N_{2k-2} : X_m \leq a\}$$

$$N_{2k} := \inf\{m > N_{2k-1} : X_m \geq b\}$$

with $N_0 = 0$. We define the $\mathcal{F}_n$-random variable

$$U_n[a, b] = \sup\{k \geq 0 : N_{2k} \leq n\}$$

to be the number of ***complete upcrossings*** *by time $n \in \mathbb{N}$.*

Proof. We wish to check that N_{2k-1}, N_{2k} are stopping times for $k \geq 1$ and that $U_n[a,b] \in \mathcal{F}_n$. We establish the first claim by induction. First observe that, letting $n \in \mathbb{N}$,

$$[N_1 = n] = [X_1 > a, \ldots, X_{n-1} > a] \cap [X_n \leq a] \in \mathcal{F}_n$$

by virtue of $(\mathcal{F}_n)_{n=1}^{\infty}$ being a filtration, so N_1 is a stopping time. Now suppose $N_1, \ldots, N_{2k-1}$ are stopping times for $k \geq 1$. Then, letting $n \in \mathbb{N}$,

$$\begin{aligned}[N_{2k} = n] &= [X_{N_{2k-1}+1} < b, \ldots, X_{n-1} < b] \cap [X_n \geq b] \in \mathcal{F}_n \\ &\bigcup_{m=1}^{n-1} [X_{m+1} < b, \ldots, X_{n-1} < b] \cap [X_n \geq b] \cap [N_{2k-1} = m] \in \mathcal{F}_n\end{aligned}$$

by virtue of N_{2k-1} being a stopping time and $(\mathcal{F}_n)_{n=1}^{\infty}$ being a filtration. Thus, N_{2k} is a stopping time. A similar approach will show that N_{2k+1} must also be a stopping time.

Now we check that $U_n[a,b] \in \mathcal{F}_n$. Let $k \in \mathbb{N}$, then

$$[U_n[a,b] = k] = [N_{2k} \leq n] \cap [N_{2k+2} > n] \in \mathcal{F}_n,$$

which holds because N_{2k} and N_{2k+2} are both stopping times. $\square$

Proposition 1.6.2. *Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be discrete-time process adapted to $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$. Let*

$$H_n := \begin{cases} 1 & \textit{if } \exists k \geq 1 \textit{ s.t. } N_{2k-1} < n \leq N_{2k} \\ \\ 0 & \textit{otherwise} \end{cases}$$

*Then $(H_n)_{n\geq 1}$ is a predictable process, and thus a strategy for $(X_n)_{n\geq 1}$. We call it the **upcrossing strategy**.*

Proof.

$$H_n = \sum_{k=1}^{\infty} 1_{N_{2k-1}\leq n-1} 1_{n-1<N_{2k}}$$

and N_j are all stopping times for $j \geq 1$. Therefore $H_n \in \mathcal{F}_{n-1}$ and it is predictable. □

We now prove some sufficient criteria for a submartingale (and thus a martingale) to converge almost surely to a random variable in $L^1(\mathcal{B})$. First a useful lemma about submartingales

Lemma 1.6.3. *Let $(X_n)_{n\geq 1}$ be a discrete-time submartingale wrt filtration $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$. Let $\psi : \mathbb{R} \to \mathbb{R}$ be a non-decreasing convex function so that $\psi(X_n) \in L^1(\mathcal{B})$ for all $n \geq 1$. Then $(\psi(X_n))_{n\geq 1}$ is a submartingale.*

Proof. Since all real convex functions are continuous, we have that $\psi(X_n) \in \mathcal{F}_n$ since $X_n \in \mathcal{F}_n$ and ψ is borel measureable for all $n \geq 1$. We have already assumed that $\psi(X_n) \in L^1(\mathcal{B})$ for all $n \geq 1$. Finally, letting $n \geq 1$, observe that

$$\mathbb{E}(\psi(X_{n+1})|\mathcal{F}_n) \geq \psi(\mathbb{E}(X_{n+1}|\mathcal{F}_n)) \geq \psi(X_n).$$

$\square$

Proposition 1.6.4 (Upcrossing Inequality). *Let* $(X_n)_{n\geq 1}$ *be a discrete-time submartingale wrt filtration* $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$. *Then* $\forall a < b \in \mathbb{R}$,

$$(b-a)\mathbb{E}(U_n[a,b]) \leq \mathbb{E}(X_n - a)^+ - \mathbb{E}(X_1 - a)^+$$

Proof. Let $(H_n)_{n\geq 1}$ be the upcrossing strategy for $(X_n)_{n\geq 1}$. We let $Y_n = a + (X_n - a)^+$ for $n \geq 1$. Because of the convexity of $(\cdot)^+$, we have that $(Y_n)_{n\geq 1}$ is a submartingale. The only difference between X_n and Y_n is that $Y_n = a$ when $X_n \leq a$ for all $n \geq 1$. Therefore $(Y_n)_{n\geq 1}$ has a the same number of complete upcrossings as $(X_n)_{n\in\mathbb{N}}$. Since $(Y_n)_{n\geq 1}$ cannot dip below a, we have that

$$(b-a)U_n[a,b] \leq (H\cdot Y)_n = \sum_{m=1}^{n} H_m(Y_m - Y_{m-1}) \tag{1.1}$$

with $Y_0 = 0$ as in the definition 1.5.1, since $H_n = 1$ if and only if Y_n is rising up to b after a fall from above b down to a for $n \geq 2$. That is, we will include a $(b-a)$ for each upcrossing in the sum along with any extra that comes from breaking above b and any remaining incomplete upcrossing. Furthermore, for $n \geq 1$

$$Y_n = \sum_{m=1}^{n} Y_m - Y_{m-1} = \sum_{m=1}^{n} (H_m - (1-H_m))(Y_m - Y_{m-1}) = (H\cdot Y)_n + ((1-H)\cdot Y)_n.$$

Denote $K := 1 - H$. Thus, $\mathbb{E}(Y_n) = \mathbb{E}((H \cdot Y)_n) + \mathbb{E}((K \cdot Y)_n)$. Since K is just another strategy, corollary 1.5.3 tells us that $((K \cdot Y)_n)_{n \geq 1}$ is another submartingale. Then, we have that $\mathbb{E}((K \cdot Y)_n) \geq \mathbb{E}((K \cdot Y)_1)$ by smoothing. Note that $H_1 = 0$ since all $N_{2k-1} \geq 1$ for $k \geq 1$. Thus

$$(K \cdot Y)_1 = Y_1 - Y_0 = Y_1.$$

Therefore, $\mathbb{E}(Y_n) - \mathbb{E}(Y_1) \geq \mathbb{E}(H \cdot Y)_n$. Putting it together with Equation 1.1, we get

$$(b-a)\mathbb{E}(U_n[a,b]) \leq \mathbb{E}((H \cdot Y)_n) \leq \mathbb{E}(Y_n) - \mathbb{E}(Y_1)$$

Substituting our original expression for Y_n,

$$(b-a)\mathbb{E}(U_n[a,b]) \leq a + \mathbb{E}(X_n - a)^+ - a - \mathbb{E}(X_1 - a)^+ = \mathbb{E}(X_n - a)^+ - \mathbb{E}(X_1 - a)^+$$

□

Using proposition 1.6.4, we can prove convergence submartingales with uniformly bounded first moment:

Theorem 1.6.5 (Martingale Convergence Theorem). *Let* $(X_n)_{n \geq 1}$ *be a discrete-time submartingale wrt filtration* $\mathcal{F}_* := (\mathcal{F}_n)_{n \in \mathbb{N}}$ *so that* $\sup_{n \geq 1} \mathbb{E}(X_n)^+ < \infty$. *Then there exists* $X \in L^1(\mathcal{B})$ *so that* $X_n \to X$ *almost surely.*

Before preceding with the proof of theorem 1.6.5, we first establish a helpful lemma.

Lemma 1.6.6. *Let* $(X_n)_{n\geq 1}$ *be a discrete-time submartingale wrt filtration* $\mathcal{F}_* := (\mathcal{F}_n)_{n\in\mathbb{N}}$. *Then* $\sup_{n\geq 1}\mathbb{E}(X_n)^+ < \infty$ *iff* $\sup_{n\geq 1}\mathbb{E}|X_n| < \infty$.

Proof. ($\Rightarrow$): Because $(X_n)_{n\geq 1}$ is a submartingale, $\forall n \geq 1$,

$$\begin{aligned}
&\mathbb{E}(X_n) \geq \mathbb{E}(X_1) \\
\Rightarrow &\mathbb{E}(X_n)^+ - \mathbb{E}(X_n)^- \geq \mathbb{E}(X_1) \\
\Rightarrow &\mathbb{E}(X_n)^+ - \mathbb{E}(X_1) \geq \mathbb{E}(X_n)^- \\
\Rightarrow &\infty > \sup_{n\geq 1}\mathbb{E}(X_n)^+ - \mathbb{E}(X_1) \geq \sup_{n\geq 1}\mathbb{E}(X_n)^-
\end{aligned}$$

Thus,

$$\begin{aligned}
\sup_{n\geq 1}\mathbb{E}|X_n| &= \sup_{n\geq 1}(\mathbb{E}(X_n)^+ + \mathbb{E}(X_n)^-) \\
&\leq \sup_{n\geq 1}(\mathbb{E}(X_n)^+) + \sup_{n\geq 1}(\mathbb{E}(X_n)^-) \\
&< \infty
\end{aligned}$$

✓

($\Leftarrow$): This follows because

$$\infty > \sup_{n\geq 1}\mathbb{E}|X_n| \geq \sup_{n\geq 1}(\mathbb{E}(X_n)^+ + \mathbb{E}(X_n)^-) \geq \sup_{n\geq 1}\mathbb{E}(X_n)^+$$

✓ □

Proof of theorem 1.6.5. Letting $a < b \in \mathbb{R}$, we uniformly bound the number of complete up-crossings with proposition 1.6.4. First note that if $a \geq 0$

$$(X_n - a)^+ = 1_{X_n \geq a} X_n - 1_{X_n \geq a} a \leq 1_{X_n \geq a} X_n \leq (X_n)^+$$

and if $a < 0$

$$(X_n - a)^+ = 1_{X_n \geq a} X_n - 1_{X_n \geq a} a = (X_n)^+ + 1_{0 > X_n \geq a} X_n - 1_{X_n \geq a} a \leq (X_n)^+ + |a|.$$

So $\forall n \geq 1$

$$\begin{aligned}(b-a)\mathbb{E}(U_n[a,b]) &\leq \mathbb{E}(X_n - a)^+ - \mathbb{E}(X_1 - a)^+ \\ &\leq \mathbb{E}(X_n)^+ + |a| \\ &\leq M + |a|\end{aligned}$$

for $\infty > M \geq \sup\limits_{n \geq 1} \mathbb{E}(X_n)^+$. Therefore, we have that $\mathbb{E}(U_n[a,b]) \leq \frac{M+|a|}{b-a}$. Since $U_n[a,b]$ is non-decreasing as $n \to \infty$, it converges (possibly to ∞), say to $U[a,b]$. By monotone convergence theorem, we have that

$$\frac{M+|a|}{b-a} \geq \lim_{n \to \infty} \mathbb{E}(U_n[a,b]) = \mathbb{E}(U[a,b]).$$

Therefore, $U[a,b]$ must be finite almost surely (otherewise $\mathbb{E}(U[a,b]) = \infty$). This means that

$$\mathbb{P}[\liminf_{n\to\infty} X_n \leq a < b \leq \limsup_{n\to\infty} X_n] = 0,$$

since that event occurs iff $U[a,b] = \infty$.

Therefore

$$\bigcup_{a,b\in\mathbb{Q}} [\liminf_{n\to\infty} X_n \leq a < b \leq \limsup_{n\to\infty} X_n]$$

is a null set. But then we must have that $\liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n =: X$ almost surely.

Finally, we show that $X \in L^1(\mathcal{B})$. By lemma 1.6.6, $\infty > \sup_{n\geq 1} \mathbb{E}|X_n|$. By Fatou's Lemma, we have that

$$\mathbb{E}|X| \leq \liminf_{n\to\infty} \mathbb{E}|X_n| \leq \sup_{n\geq 1} \mathbb{E}|X_n| < \infty.$$

so $X \in L^1(\mathcal{B})$. $\square$

We end this section by looking at a couple examples to help us understand what theorem 1.6.5 tells us (and what it doesn't).

Example 1: Consider the simple random walk $(X_n)_{n\geq 1}$ starting at 0 with increments of ± 1. Let $\nu := \inf\{n \geq 1 : X_n \geq 2\}$. Since $(X_n)_{n\geq 1}$ is a martingale and ν is a stopping time, so $X_{n\wedge\nu}$ is a martingale as well. Since $(X_n)_{n\geq 1}$ starts at 0 and moves by at most 1, we must have that $\nu \geq 3$. So $0 = \mathbb{E}(X_1) = \mathbb{E}(X_{1\wedge\nu}) = \mathbb{E}(X_{n\wedge\nu})$ $\forall n \geq 1$. Now certainly $\sup_{n\geq 1} \mathbb{E}(X_{n\wedge\nu})^+ \leq 2$, so $X_{n\wedge\nu} \to X$ almost surely by theorem 1.6.5. Since $X_{n\wedge\nu}$ only takes integral values, X must (almost surely) as well. But since $X_{n\wedge\nu}$ changes by ± 1 if it is not already at 2, X must be 2. Thus, $\mathbb{E}(X) = 2 \neq 0 = \mathbb{E}(X_{n\wedge\nu})$ for $n \geq 1$. Therefore $(X_{n\wedge\nu})_{n\geq 1}$ cannot converge in $L^1(\mathcal{B})$.

This example seems to fly in our face of what it means to be a "fair game". How can $(X_{n\wedge\nu})_{n\geq 1}$ be a martingale, but converge to something with larger expectation almost surely? If we were to play finitely many games, then the probability mass assigned to positive results 1 and 2 will be matched equally by the probability mass asssigned to negative results. That is, although it is almost sure that at some n we will hit $X_n = 2$, we may suffer very large losses before that happens.

Example 2:

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?

In this section we examine an example studied in (3) about counting the number of connected components of size k in an Erdos-Renyi graph. First we construct our probability space.

We let $\mathbb{G}_n$ be the set of graphs on $[n]$ for some $n \geq 0$. Denote $\binom{[n]}{2}$ the set of edges of the complete graph on $[n]$. Each graph in $G \in \mathbb{G}_n$ has the same vertex set, so they can be identified with their edge sets $E(G) \subset \binom{[n]}{2}$. Thus, we simply refer to a graph in $\mathbb{G}_n$ by its set of edges.

Let $n \in \mathbb{N}$, let $\Omega := \binom{[n]}{2}$, and let $M = \binom{n}{2}!$. Let $\mathbb{P} : \mathcal{P}(\Omega^M) \to [0, 1]$ be defined for $\omega \in \Omega^M$ so that

$$\mathbb{P}(\{\omega\}) = \begin{cases} \frac{1}{M!} & \text{if } \omega \text{ consists of distinct edges} \\ \\ 0 & \text{otherwise} \end{cases}$$

This probability measure simply assigns uniform probability to all permutations of edges. Let $\mathcal{F}_0 = \{\emptyset, \Omega^M\}$ and $\mathcal{F}_i = \{S \times \Omega^{M-i} : S \in \mathcal{P}(\Omega^i)\}$ make up the filtration $\mathcal{F}_* := (\mathcal{F}_i)_{i=0}^M$.

Definition 1.12.1. *An **Erdos-Renyi random graph** on $m \in \{0, \ldots, M\}$ edges, $G_{n,m} : \Omega \to \mathbb{G}_n$, is defined on $\omega \in \Omega$ as*

$$G_{n,m}(\omega) = \{\omega_1, \ldots, \omega_m\}.$$

That is, it is a random graph sampled uniformly from all graphs on n vertices with m edges.

We can think of sampling a particular Erdos-Renyi graph as the result of picking an edge m times without replacement. This is captured by the process defined below:

Definition 1.12.2. *The **Erdos-Renyi graph process** is the stochastic process* $(G_{n,i} : \Omega \to \mathbb{G}_n)_{i=0}^{M}$ *adapted to* $\mathcal{F}_*$.

We would like to characterize properties about this process as $n \to \infty$. To do so, we use the following concept:

Definition 1.12.3. *Let* $(X_n : \Omega \to \mathbb{R})_{n\geq 1}$ *be a stochastic process and let* $(x_n)_{n\geq 1} \subset \mathbb{R}$. *Then* $X_n = x_n$ ***asymptotically almost surely*** *if there exists* $(\epsilon_n)_{n\geq 1} \subset \mathbb{R}$ *so that* $\epsilon_n = o(1)$ *and*

$$\mathbb{P}((1-\epsilon_n)x_n \leq X_n \leq (1+\epsilon_n)x_n) \to 1$$

as $n \to \infty$. *It may be abbreviated a.a.s..*

We would like to prove the following a.a.s. characterization of the number of connected components of a fixed size of an Erdos-Renyi random graph.

Theorem 1.12.4. *Let* $\kappa \geq 1$ *and* $c \in \mathbb{R}^{\geq 0}$. *Then a.a.s. the number of connected components of order* k, *for* $1 \leq k \leq \kappa$, *in* $G(n, \lfloor cn \rfloor)$ *is*

$$\frac{k^{k-2}}{k!}(2c)^{k-1}e^{-2kc}n$$

To prove this theorem, we will apply the *differential equation method.* This method works by approximating a random discrete process (in our case the number of connected components of size k in our random graph as we sample edges)with a continuous, deterministic, differential equation. We construct the differential equation using the expected change in our tracked random variable. We'll use some process-related heuristic to find a solution to the differential equation. We then show that our error incurred from our continuous approximation is small by bounding it from above and below by a supermartingale and submartingale respectively, then applying the following:

Theorem 1.12.5 (Azuma-Hoeffding's Inequality (3))**.** *Let* $(Y_n)_{n\geq 0}$ *be a supermartingale and suppose there exists real* $C \geq 0$ $|Y_{n+1} - Y_n| \leq C$ *a.s. for* $n \geq 0$*. Then, for all* $\lambda \in \mathbb{R}^{\geq 0}$ *and* $n \geq 1$,

$$\mathbb{P}(Y_n - Y_0 \geq \lambda) \leq \exp\left(-\frac{\lambda^2}{2C^2 n}\right)$$

TODO: Proof?

The following will also be useful:

Theorem 1.12.6 (Taylor's Theorem (3))**.** *Let* $f \in C^2([a, b])$ *for* $a < b \in \mathbb{R}$*. Then there exists* $\tau \in (a, b)$ *so that*

$$f(b) = f(a) + f'(a)(b - a) + \frac{f''(\tau)}{2}(b - a)^2.$$

Now we come back to our task of studying the number of fixed size connected components in an Erdos-Renyi graph. Recall we are given $c \in \mathbb{R}^{\geq 0}$ and suppose $n \in \mathbb{N}$. Let $i \leq \lfloor cn \rfloor$ and define $Y_k(i)$ to be the number of components with exactly k vertices in $G_{n,i}$.

We would like to show that $Y_k(i) = ny_k(t_i)$ a.a.s. for $t_i := \frac{i}{n}$ and $y_k(t) = \frac{k^{k-2}}{k!}(2t)^{k-1}e^{-2kt}$. In our Erdos-Renyi process, we'll let i progress to $\lfloor cn \rfloor$ (i.e. let $t \to c$) to get the desired result. To prove $Y_k(i) = ny_k(t_i)$ a.a.s., we'll make a (good) guess that $\epsilon_n(t) = n^{-1/3}e^{4\kappa^3 t}$ and show

$$n(y_k(t_i) - \epsilon_n(t_i)) \leq Y_k(i) \leq n(y_k(t_i) + \epsilon_n(t_i))$$

tends to probability 1. This reasoning behind this choice of ϵ_n will be clear as we proceed with the computations.

For now, we'll assume such a y_k exists and is in $C^2([0, c])$, but take it to be unknown. Let $\nu_n := \inf\{i \geq 0 : Y_k(i) \notin [n(y_k(t_i) - \epsilon_n(t_i)), n(y_k(t_i) + \epsilon_n(t_i))]\}$. Define

$$Y_k^{\pm}(i) = Y_k(i \wedge \nu_n) - n(y_k(t_{i\wedge\nu_n}) \pm \epsilon_n(t_{i\wedge\nu_n}))$$

Effectively, we're freezing our error once $Y_k(i)$ leaves our error bound. We will show that $Y_k^+(i)$ is a supermartingale. Using that, we will use Azuma-Hoeffding (theorem 1.12.5) to show that $\mathbb{P}(\nu_n = O(n)) = o(1)$. The corresponding computations for Y_k^- are symmetric.

Let's take a look at the one-step change of Y_k from i to $i+1$, which we'll call $\Delta Y_k(i)$. Given $\mathcal{F}_i$ (that is, the first i choices of edges), we have the possible outcomes for new edge ω_{i+1}:

1. ω_{i+1} connects a component of size k to one of size that's not k. There are $kY_k(i)(n - kY_k(i))$ such edges. In this case $\Delta Y_k(i) = -1$.

2. ω_{i+1} connects a component of size k to another component of size k. There are $\frac{1}{2}k^2(Y_k(i) - 1)Y_k(i)$ such edges. In this case $\Delta Y_k(i) = -2$.

3. ω_{i+1} connects a component of size j and size $k-j$ (for some $j<k$). There are $jY_j(i)(k-j)Y_{k-j}(i)$ such edges. In this case $\Delta Y_k(i)=1$.

4. Otherwise, $\Delta Y_k(i)=0$.

There are $\binom{n}{2}-i$ possible choices for edges. This means the probability of picking a particular edge that hasn't been chosen yet is

$$\frac{1}{\binom{n}{2}-i}=\frac{2}{n^2+n-2i}\leq\frac{2}{n^2}\left(\frac{n^2}{n^2-n}\right)=\frac{2}{n^2}\left(1+O(\frac{1}{n})\right)$$

First looking at the contribution of cases 1 and 2 to $\mathbb{E}(\Delta Y_k(i)|\mathcal{F}_i)$

$$\begin{aligned}
&-\frac{2}{n^2}\left(1+O(\frac{1}{n})\right)\left[nkY_k(i)-k^2Y_k(i)^2+k^2Y_k(i)^2-k^2Y_k(i)\right]\\
=&-\frac{2}{n^2}\left(1+O(\frac{1}{n})\right)\left[nkY_k(i)-k^2Y_k(i)\right]\\
=&\left(1+O(\frac{1}{n})\right)\left[-2k\frac{Y_k(i)}{n}+k^2\frac{2Y_k(i)}{n^2}\right]\\
\leq&\left(1+O(\frac{1}{n})\right)\left[-2k\frac{Y_k(i)}{n}+\kappa^2\frac{2}{n}\right]\\
\leq&-2k\frac{Y_k(i)}{n}+O(\frac{1}{n})+O(\frac{1}{n^2})\\
=&-2k\frac{Y_k(i)}{n}+O(\frac{1}{n})
\end{aligned}$$

where we make use of the fact that $Y_k(i)\leq n$ and $k\leq\kappa$.

Next, we look at the contribution of case 3:

$$\begin{aligned}
&\frac{2}{n^2}\left(1+O(\frac{1}{n})\right)\frac{1}{2}\sum_{j=1}^{k-1} jY_j(i)(k-j)Y_{k-j}(i)\\
&=\left(1+O(\frac{1}{n})\right)\sum_{j=1}^{k-1} j(k-j)\frac{Y_j(i)}{n}\frac{Y_{k-j}(i)}{n}\\
&\leq \sum_{j=1}^{k-1}\left[j(k-j)\frac{Y_j(i)}{n}\frac{Y_{k-j}(i)}{n}\right]+\kappa^3 O(\frac{1}{n})\\
&=\sum_{j=1}^{k-1}\left[j(k-j)\frac{Y_j(i)}{n}\frac{Y_{k-j}(i)}{n}\right]+O(\frac{1}{n})
\end{aligned}$$

where we again use the fact that $Y_j(i), Y_{k-j}(i) \leq n$ and $j, k-j \leq \kappa$. Thus, we have that

$$\mathbb{E}(\Delta Y_k(i)|\mathcal{F}_i) = -2k\frac{Y_k(i)}{n} + \sum_{j=1}^{k-1}\left[2j(k-j)\frac{Y_j(i)}{n}\frac{Y_{k-j}(i)}{n}\right]+O(\frac{1}{n})$$

Assuming $i < \nu_n$, we have that

$$\begin{aligned}
\mathbb{E}(\Delta Y_k(i)|\mathcal{F}_i, [i<\nu_n]) &= -2k(y_k(t_i)+\epsilon_n(t_i)) + \sum_{j=1}^{k-1}\Big[j(k-j)(y_j(t_i)+\epsilon_n(t_i))(y_{k-j}(t_i)+\epsilon_n(t_i))\Big]+O(\frac{1}{n})\\
&\leq -2ky_k(t_i) + \sum_{j=1}^{k-1}\Big[j(k-j)y_j(t_i)y_{k-j}(t_i)\Big] - 2k\epsilon_n(t_i) + \epsilon_n(t_i)3k^3 + O(\frac{1}{n})\\
&\leq -2ky_k(t_i) + \sum_{j=1}^{k-1}\Big[j(k-j)y_j(t_i)y_{k-j}(t_i)\Big] + \epsilon_n(t_i)3k^3 + O(\frac{1}{n})
\end{aligned}$$

using the fact that $y_l(t) \leq 1$ for all $1 \leq l \leq \kappa, 0 \leq t \leq c$ and $0 \leq \epsilon_n(t) \leq 1$ for sufficiently large n for all $0 \leq t \leq c$.

Now note that given $[i < \nu_n]$, $\Delta Y_k^+(i) = \Delta Y_k(i) - (y_k(t_i + \frac{1}{n}) - y_k(t_i) + \epsilon_n(t_i + \frac{1}{n}) - \epsilon_n(t_i))$. We can simplify that last term using Taylor's Theorem (theorem 1.12.6) to get

$$\begin{aligned} n(y_k(t_i + \frac{1}{n}) - y_k(t_i)) &= n(\frac{y_k'(t_i)}{n} + \frac{y_k''(\tau)}{n^2} \\ &= y_k'(t_i) + O(\frac{1}{n}) \end{aligned}$$

for some $\tau \in (t_i, t_{i+1})$. Note that $y_k \in C^2([0, c])$, so y_k'' is continuous and it is bounded on $[0, c]$ (by the extreme value theorem). This justifies replacing $n\frac{y_k''(\tau)}{n^2}$ with $O(\frac{1}{n^2})$. Applying Taylor's Theorem to ϵ_n as well:

$$n(\epsilon_n(t_i + \frac{1}{n}) - \epsilon_n(t_i)) = \epsilon_n'(t_i) + O(\frac{1}{n})$$

Thus, we get

$$\begin{aligned} &\mathbb{E}(\Delta Y_k^+(i) | \mathcal{F}_i, [i < \nu_n]) \\ &\quad = -2ky_k(t_i) + \sum_{j=1}^{k-1} \Big[j(k-j) y_j(t_i) y_{k-j}(t_i) \Big] + \epsilon_n(t_i) 3k^3 - y_k'(t_i) - \epsilon_n'(t_i) + O(\frac{1}{n}) \end{aligned}$$

Note that if we choose our $(y_k)_{k=1}^{\kappa}$ so that it solves the ODE system

$$y_k' = -2ky_k + \sum_{j=1}^{k-1} \Big[j(k-j) y_j y_{k-j} \Big] \; \forall 1 \leq k \leq \kappa$$

then we will get

$$\begin{aligned}
&\mathbb{E}(\Delta Y_k^+(i)|\mathcal{F}_i, [i < \nu_n]) \\
&\quad = \epsilon_n(t_i)3k^3 - \epsilon_n'(t_i) + O(\frac{1}{n}) \\
&\quad = 3k^3 n^{-1/3} e^{4\kappa^3 t_i} - 4\kappa^3 n^{-1/3} e^{4\kappa^3 t_i} + O(\frac{1}{n}) \\
&\quad \le -\kappa^3 n^{-1/3} e^{4\kappa^3 t_i} + O(\frac{1}{n}) \\
&\quad \le 0
\end{aligned}$$

for sufficiently large n. Note that we applied $k \le \kappa$ above. From this computation the particular choice for the exponential part of our error term should now make sense. For the event $[i \ge \nu_n]$, Y_k^+ is frozen and we have that $\mathbb{E}(\Delta Y_k^+(i)|\mathcal{F}_i, [i \ge \nu_n]) = 0$. This means $\mathbb{E}(\Delta Y_k^+(i)|\mathcal{F}_i) \le 0$ and is thus a supermartingale.

Now to solve our ODE system

$$y_k' = -2ky_k + \sum_{j=1}^{k-1} \left[2j(k-j)y_j y_{k-j}\right] \; \forall 1 \le k \le \kappa.$$

We claim (unsurprisingly) that $y_k(t) = \frac{k^{k-2}}{k!}(2t)^{k-1}e^{-2kt}$ solves the ODE system.

Claim 1.12.7. $y_k(t) = \frac{k^{k-2}}{k!}(2t)^{k-1}e^{-2kt}$ *is the unique solution to the ODE system*

$$y_k' = -2ky_k + \sum_{j=1}^{k-1} \left[j(k-j)y_j y_{k-j}\right] \; \forall 1 \le k \le \kappa.$$

Proof of claim 1.12.7. For $k = 1$, the relevant ODE has form

$$y_1' = -2y_1$$

which is solved by $y_1 = e^{-2t}$. This agrees with our claimed solution after substituting $k = 1$. For $k \geq 2$, we get

$$\begin{aligned}
-2ky_k + &\sum_{j=1}^{k-1}\left[j(k-j)y_j y_{k-j}\right] \\
&= -2k\frac{k^{k-2}}{k!}(2t)^{k-1}e^{-2kt} + \sum_{j=1}^{k-1}\left[j(k-j)\frac{j^{j-2}}{j!}(2t)^{j-1}e^{-2jt}\frac{(k-j)^{(k-j)-2}}{(k-j)!}(2t)^{(k-j)-1}e^{-2(k-j)t}\right] \\
&= -2k\frac{k^{k-2}}{k!}(2t)^{k-1}e^{-2kt} + \sum_{j=1}^{k-1}\left[\frac{j^{j-1}(k-j)^{(k-j)-1}}{j!(k-j)!}(2t)^{k-2}e^{-2kt}\right] \\
&= -2k\frac{k^{k-2}}{k!}(2t)^{k-1}e^{-2kt} + (2t)^{k-2}\frac{e^{-2kt}}{k!}\sum_{j=1}^{k-1}\left[\binom{k}{j}j^{j-1}(k-j)^{(k-j)-1}\right]
\end{aligned}$$

We would like to show

$$\sum_{j=1}^{k-1}\left[\binom{k}{j}j^{j-1}(k-j)^{(k-j)-1}\right] = 2(k-1)k^{k-2}.$$

Then,

$$\begin{aligned}
-2ky_k + \sum_{j=1}^{k-1}\left[j(k-j)y_j y_{k-j}\right] \\
&= -2k\frac{k^{k-2}}{k!}(2t)^{k-1}e^{-2kt} + (2t)^{k-2}\frac{e^{-2kt}}{k!}2(k-1)k^{k-2} \\
&= y_k'
\end{aligned}$$

by product rule.

We will prove the closed form of our sum by examining labelled trees on [k]. Observe that for $1 \leq j \leq k-1$

1. we have $\binom{k}{j}$ choices subsets $J \subset [k]$ of size $|J| = j$
2. we have j^{j-2} choices for labelled trees on J (Cayley's Formula)
3. we have $(k-j)^{k-j-2}$ choices for labelled trees on $[k] \setminus J$
4. we have $j(k-j)$ choices for edges between J and $[k] \setminus J$.

Making a choice at each step will join two labelled trees to make a labelled tree on [k]. Each labelled tree on [k] has $k-1$ edges, so there are $2(k-1)$ ways to create a given labelled tree on [k] using the above procedure. The choices that produce a given labelled tree T on [k] can be reverse engineered by

1. choosing one of the $k-1$ edges of T, call it e
2. removing edge e from T and choosing one of the 2 remaining components as J.

Since there are k^{k-2} labelled trees on [k], we have that

$$2(k-1)k^{k-2} = \sum_{j=1}^{k-1} \left[\binom{k}{j} j^{j-2}(k-j)^{(k-j)-2} j(k-j) \right]$$
$$= \sum_{j=1}^{k-1} \left[\binom{k}{j} j^{j-1}(k-j)^{(k-j)-1} \right].$$

Substituting back into our system and applying product rule gets us the desired result. $\square$

So we have that for $y_k(t) = \frac{k^{k-2}}{k!}(2t)^{k-1}e^{-2kt}$ and $\epsilon_n(t) = n^{-1/3}e^{4\kappa^3 t}$, $(Y_k^+(i))_{i=0}^{\lfloor cn \rfloor}$ is a supermartingale.

We would now like to apply Azuma-Hoeffding (Theorem 1.12.5). First we need to show $|\Delta Y_k^+(i)| = O(1)$ a.s. $\forall 0 \leq i \leq \lfloor cn \rfloor$. This is straightforward since $\Delta Y_k^+(i) = 0$ if $i \geq \nu_n$, otherwise

$$\begin{aligned}
|\Delta Y_k^+(i)| &= |\Delta Y_k(i) - (y_k(t_i + \frac{1}{n}) - y_k(t_i) + \epsilon_n(t_i + \frac{1}{n}) - \epsilon_n(t_i))| \\
&= |\Delta Y_k(i) - y_k'(t_i) - \epsilon_n'(t_i) + O(\frac{1}{n})| \\
&\leq |\Delta Y_k(i)| + |y_k'(t_i)| + |\epsilon_n'(t_i)| + O(\frac{1}{n}) \\
&\leq 2 + O(1) + O(1) + O(\frac{1}{n}) \\
&= O(1)
\end{aligned}$$

because y_k and ϵ_n are in $C^2([0, c])$ and Y_k changes by at most -2.

TODO: Complete application of Azuma to wrap it up.

TODO: - Erdos-renyi graph - Thm 1.2 (and defn of a.a.s.) about number of connected components - Erdos-renyi process and equivalence at step to G(n, i) - What is differential equation Method and informally how we will use item - Useful theorems - Taylor's Theorem (no proof) - Azuma's inequality (with proof) - Chernoff-Cramer bound - Hoeffding's Lemma - Azuma-Hoeffding inequality - Cayley's formula (no proof) - stopping time for bad event - proof it is supermartingale - proof guessed solution solves differential equation - application of azuma-hoeffding

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).

3. Bennett, P. and Dudek, A.: A gentle introduction to the differential equation method and dynamic concentration, 2020.

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).