

III Advanced Probability

Ishan Nath, Michaelmas 2024

Based on Lectures by Prof. Jason Miller

November 8, 2024

Contents

0	Introduction	2
1	Conditional Expectation	3
1.1	Discrete Case	4
1.2	Existence and Uniqueness	5
1.3	Properties of Expectation	7
1.4	Examples of Conditional Expectation	11
2	Discrete Time Martingales	14
2.1	Stopping Times	15
2.2	Martingale Convergence Theorem	19
2.3	\mathcal{L}^p Convergence of Martingales	23
2.4	Uniformly Integrable Martingales	25
2.5	Backwards Martingales	28
2.6	Applications of Martingale Theory	29
3	Continuous Time Processes	34
3.1	Martingale Regularization Theorem	38
3.2	Convergence and Doob's Inequalities	43
3.3	Kolmogorov Continuity Criterion	45
4	Weak Convergence	47
	Index	49

0 Introduction

We will be following the lecture notes by Perla Sousi.

This course is divided into seven main topics:

- Chapter 1. Conditional expectation. We have seen how to define $\mathbb{E}[Y|X = x]$ for X a discrete-valued random variable, or one with continuous density wrt. the Lebesgue measure. We will define more generally $\mathbb{E}[Y|\mathcal{G}]$, for \mathcal{G} a σ -algebra.
- Chapter 2. Discrete-time martingales. A martingale is a sequence (X_n) of integrable random variables such that $\mathbb{E}[X_{n+1} | X_1, \dots, X_n] = X_n$, for all n . A basic example is a simple symmetric random walk. We will be looking at properties, bounds and computations.
- Chapter 3. Continuous time processes. Examples are
- Martingales in continuous time.
 - Some continuous continuous-time processes.
- Chapter 4. Weak convergence. This is a notion of convergence which extends convergence in distribution, and is related to other notions of convergence.
- Chapter 5. Large deviations. Here we estimate how unlikely rare events are. An example if the sample mean for (ξ_j) iid, with mean μ , which converges to μ by SLLN. We can ask what the probability is that the sample mean deviates from μ by ε , which can be answered by Cramer's theorem.
- Chapter 6. Brownian motion. This is a fundamental object, which is a continuous time stochastic process. We will look at the definition, Markov processes, and its relationship with PDEs.
- Chapter 7. Poisson random measures. This is the generalization of the standard Poisson process on \mathbb{R}_+ .

1 Conditional Expectation

Recall that a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ consists of a set Ω , a σ -algebra on Ω , meaning:

- $\Omega \in \mathcal{F}$,
- $A \in \mathcal{F} \implies A^c \in \mathcal{F}$,
- if (A_n) is in \mathcal{F} , then $\bigcup A_n \in \mathcal{F}$,

and \mathbb{P} is a probability measure on (Ω, \mathcal{F}) , meaning:

- $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ such that
- $\mathbb{P}(\Omega) = 1, \mathbb{P}(\emptyset) = 0$,
- if (A_n) are pairwise disjoint in \mathcal{F} , then $\mathbb{P}(\bigcup A_n) = \sum \mathbb{P}(A_n)$.

An important σ -algebra is the Borel σ -algebra \mathcal{B} , which for \mathbb{R} is the intersection of all σ -algebra on \mathbb{R} which contain the open sets in \mathbb{R} .

Recall that $X : \Omega \rightarrow \mathbb{R}$ is a random variable if $X^{-1}(U) \in \mathcal{F}$ for all $U \subseteq \mathbb{R}^n$ open. This is equivalent to $X^{-1}(B) \in \mathcal{F}$ for all $B \in \mathcal{B}$.

Definition 1.1. Suppose that \mathcal{A} is a collection of sets in Ω . Then

$$\sigma(\mathcal{A}) = \bigcap \{ \mathcal{E} \mid \mathcal{E} \text{ is a } \sigma\text{-algebra containing } \mathcal{A} \}.$$

If (X_i) is a collection of random variables, then

$$\sigma(X_i \mid i \in I) = \sigma(\{ \omega \in \Omega \mid X_i(\omega) \in B \}, i \in I, B \in \mathcal{B}).$$

This is the smallest σ -algebra which makes X_i measurable for all $i \in I$.

Let $A \in \mathcal{F}$. Set

$$\mathbb{1}_A(x) = \mathbb{1}(x \in A) = \begin{cases} 1 & x \in A, \\ 0 & \text{else.} \end{cases}$$

We can define expectation as follows:

- For $C_i \geq 0, A_i \in \mathcal{F}$, set

$$\mathbb{E} \left[\sum_{i=1}^n C_i \mathbb{1}_{A_i} \right] = \sum_{i=1}^n C_i \mathbb{P}(A_i).$$

- For $X \geq 0$, set $X_n = (2^{-n} \lfloor 2^n X \rfloor) \wedge n$, so that $X_n \uparrow X$ as $n \rightarrow \infty$. Set

$$\mathbb{E}[X] = \lim_{n \rightarrow \infty} \mathbb{E}[X_n].$$

- If X is a general random variable, write $X = X^+ - X^-$, where $X^+ = \max(X, 0)$, $X^- = \max(-X, 0)$. If $\mathbb{E}[X^+]$ or $\mathbb{E}[X^-]$ is finite, set

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-].$$

Call X integrable if $\mathbb{E}[|X|] < \infty$.

Recall that if $A, B \in \mathcal{F}$ with $\mathbb{P}(B) > 0$, then

$$\mathbb{P}(A | B) = \mathbb{P}(A \cap B) / \mathbb{P}(B).$$

If X is integrable and $A \in \mathcal{F}$ with $\mathbb{P}(A) > 0$, then

$$\mathbb{E}[X|A] = \mathbb{E}[X\mathbb{1}_A] / \mathbb{P}(A).$$

Our goal is to extend this definition to

$$\mathbb{E}[X|\mathcal{G}],$$

where \mathcal{G} is a σ -algebra in \mathcal{F} . The main idea is that $\mathbb{E}[X|\mathcal{G}]$ is the best prediction of X given \mathcal{G} .

Hence it should be a \mathcal{G} -measurable random variable Y , which minimizes

$$\mathbb{E}[(X - Y)^2].$$

1.1 Discrete Case

As a warm-up, we consider the discrete case.

Suppose that X is integrable, (B_i) is countable and disjoint with $\Omega = \bigcup B_i$, and $\mathcal{G} = (B_i, i \in I)$. Set $X' = \mathbb{E}[X|\mathcal{G}]$ by

$$X' = \sum_{i \in I} \mathbb{E}[X|B_i] \mathbb{1}_{B_i},$$

with the convention that $\mathbb{E}[X|B_i] = 0$ if $\mathbb{P}(B_i) = 0$. If $\omega \in \Omega$, then

$$X'(\omega) = \sum_i \mathbb{E}[X|B_i] \mathbb{1}(\omega \in B_i).$$

We look at some important properties of X' .

1. X' is \mathcal{G} -measurable, as it is a linear combination of the \mathcal{G} -measurable random variables $\mathbb{1}_{B_i}$.
2. X' is integrable, as

$$\mathbb{E}[|X'|] \leq \sum_{i \in I} \mathbb{E}[|X| \mathbb{1}_{B_i}] = \mathbb{E}[|X|] < \infty.$$

3. If $G \in \mathcal{G}$, then

$$\mathbb{E}[X \mathbb{1}_G] = \mathbb{E}[X' \mathbb{1}_G].$$

1.2 Existence and Uniqueness

We need a couple of important facts from measure theory:

Theorem 1.1 (Monotone Convergence Theorem). *If (X_n) is a sequence of random variables with $X_n \geq 0$ for all n and $X_n \uparrow X$ almost-surely, then $\mathbb{E}[X_n] \uparrow \mathbb{E}[X]$.*

Theorem 1.2 (Dominated Convergence Theorem). *Suppose that (X_n) is a sequence of random variables with $X_n \rightarrow X$ almost-surely and $|X_n| \leq Y$ for all n where Y is an integrable random variable, then $\mathbb{E}[X_n] \rightarrow \mathbb{E}[X]$ as $n \rightarrow \infty$.*

The construction of $\mathbb{E}[X|\mathcal{G}]$ requires a few steps; in the case that $X \in L^2$, then this can be defined by orthogonal projection.

We need to recall a few things about L^p spaces. Suppose that $p \in [1, \infty)$, and X is a random variables in (Ω, \mathcal{F}) . Then,

$$\|X\|_p = \mathbb{E}[|X|^p]^{1/p}.$$

Now

$$L^p = L^p(\Omega, \mathcal{F}, \mathbb{P}) = \{X \text{ with } \|X\|_p < \infty\}.$$

For $p = \infty$,

$$\|X\|_\infty = \inf\{\lambda \mid |X| \leq \lambda \text{ almost surely}\},$$

and

$$L^\infty = L^\infty(\Omega, \mathcal{F}, \mathbb{P}) = \{X \text{ with } \|X\|_\infty < \infty\}.$$

Two random variables in L^p are equivalent if they agree almost-surely. Then \mathcal{L}^p is the set of equivalence classes under this equivalence relation.

Theorem 1.3. *The space $(\mathcal{L}^2, \|\cdot\|_2)$ is a Hilbert space with inner product*

$$\langle X, Y \rangle = \mathbb{E}[XY].$$

If \mathcal{H} is a closed subspace of \mathcal{L}^2 , then for all $X \in \mathcal{L}^2$, there exists $Y \in \mathcal{H}$ such that

$$\|X - Y\|_2 = \inf_{Z \in \mathcal{H}} \|X - Z\|_2,$$

and $\langle X - Y, Z \rangle = 0$ for all $Z \in \mathcal{H}$. Y is the orthogonal projection of X onto \mathcal{H} .

Theorem 1.4. *Let X be an integrable random variable, and $\mathcal{G} \subseteq \mathcal{F}$ a σ -algebra. Then there exists a random variable Y such that:*

- (i) Y is \mathcal{G} -measurable.
- (ii) Y is integrable, with

$$\mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[Y \mathbb{1}_A]$$

for all $A \in \mathcal{G}$.

If Y' satisfies these two properties, then $Y = Y'$ almost surely. Y is called (a version of) the conditional expectation of X given \mathcal{G} , and we write

$$Y = \mathbb{E}[X|\mathcal{G}].$$

If $\mathcal{G} = \sigma(G)$ for a random variable G , we write $Y = \mathbb{E}[X|G]$.

Proof: We begin by proving uniqueness. Suppose Y, Y' are two conditional expectation. Then

$$A = \{Y > Y'\} \in \mathcal{G},$$

so by (ii),

$$\mathbb{E}[(Y - Y')\mathbb{1}_A] = \mathbb{E}[X\mathbb{1}_A] - \mathbb{E}[X\mathbb{1}_A] = 0,$$

but the LHS is non-negative, so the only way this can hold is if $Y \leq Y'$ almost surely. Similarly, $Y \geq Y'$ almost surely, so $Y = Y'$ almost surely.

Now we are ready to tackle existence. The first step is by restricting to \mathcal{L}^2 function.

Take $X \in \mathcal{L}^2$. Note that $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ is a closed subspace of $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$.

Set $Y = \mathbb{E}[X|\mathcal{G}]$ to be the orthogonal projections of X onto $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$.

It is automatically true that (i) holds. Suppose $A \in \mathcal{G}$. Then

$$\mathbb{E}[(X - Y)\mathbb{1}_A] = 0,$$

by orthogonality, since $\mathbb{1}_A$ is \mathcal{G} -measurable. So

$$\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A],$$

so (ii) holds. Finally Y is integrable as it is specified to be in \mathcal{L}^2 .

As an aside, notice if $X \geq 0$, then $Y = \mathbb{E}[X|\mathcal{G}] \geq 0$ almost-surely. Since $\{Y < 0\} \in \mathcal{G}$ and

$$\mathbb{E}[Y\mathbb{1}_{\{Y < 0\}}] = \mathbb{E}[X\mathbb{1}_{\{Y < 0\}}] \geq 0,$$

which can only happen if $\mathbb{P}(Y < 0) = 0$.

Our second step is to prove this for $X \geq 0$. In this case, let $X_n = X \wedge n$. Then $X_n \in \mathcal{L}^2$ for all n , so there exists \mathcal{G} -measurable random variables Y_n so that

$$\mathbb{E}[Y_n\mathbb{1}_A] = \mathbb{E}[(X \wedge n)\mathbb{1}_A],$$

for all $A \in \mathcal{G}$. Now (X_n) is increasing in n , so (Y_n) are also increasing by the above aside. Set

$$Y = \limsup_n Y_n.$$

We now have to check the definitions. Note Y is \mathcal{G} -measurable as it is a limsup of \mathcal{G} -measurable functions. So we check the second definition.

For $A \in \mathcal{G}$,

$$\mathbb{E}[Y \mathbb{1}_A] \stackrel{MCT}{=} \lim_n \mathbb{E}[Y_n \mathbb{1}_A] = \lim_n \mathbb{E}[(X \wedge n) \mathbb{1}_A] \stackrel{MCT}{=} \mathbb{E}[X \mathbb{1}_A].$$

We did not check integrability, but notice that setting $A = G$, $\mathbb{E}[Y] = \mathbb{E}[X]$, and X is non-negative, hence so is Y . Thus Y is integrable.

Finally we prove the result for $X \in \mathcal{L}^1$. We can apply step 2 to X^+ and X^- , and so can set

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X^+|\mathcal{G}] - \mathbb{E}[X^-|\mathcal{G}].$$

This is \mathcal{G} -measurable and integrable as it is the difference of two \mathcal{G} -measurable and integrable random variables, and satisfies (ii) since both of the random variables satisfy (ii).

Remark.

- The proof also works for $X \geq 0$, and not necessarily integrable. Then $\mathbb{E}[X|\mathcal{G}]$ is not necessarily integrable.
- The property

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}] \mathbb{1}_A] = \mathbb{E}[X \mathbb{1}_A]$$

for all $A \in \mathcal{G}$, is equivalent to

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]Y] = \mathbb{E}[XY],$$

for all Y bounded and \mathcal{G} -measurable.

1.3 Properties of Expectation

Definition 1.2. A collection of σ -algebras (\mathcal{G}_i) in \mathcal{G} is *independent* if whenever $G_i \in \mathcal{G}_i$ and i_1, \dots, i_n distinct, then

$$\mathbb{P}(G_{i_1} \cap \dots \cap G_{i_n}) = \prod_{k=1}^n \mathbb{P}(G_{i_k}).$$

We say that a random variable X is *independent* of a σ -algebra \mathcal{G} if $\sigma(X)$ is independent of \mathcal{G} .

Proposition 1.1. Let $X, Y \in \mathcal{L}^2$, and $\mathcal{G} \subseteq \mathcal{F}$ a σ -algebra. Then:

- (i) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$.
- (ii) If X is \mathcal{G} -measurable, then $\mathbb{E}[X|\mathcal{G}] = X$ almost surely.
- (iii) If X is independent of \mathcal{G} , then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.
- (iv) If $X \geq 0$ almost surely, then $\mathbb{E}[X|\mathcal{G}] \geq 0$ almost surely.
- (v) For any $\alpha, \beta \in \mathbb{R}$, $\mathbb{E}[\alpha X + \beta Y | \mathcal{G}] = \alpha \mathbb{E}[X|\mathcal{G}] + \beta \mathbb{E}[Y|\mathcal{G}]$.
- (vi) $|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X||\mathcal{G}]$ almost-surely.

Theorem 1.5 (Fatou's Lemma). *If (X_n) is a sequence of random variables with $X_n \geq 0$, then*

$$\mathbb{E}[\liminf_n X_n] \leq \liminf_n \mathbb{E}[X_n].$$

Theorem 1.6 (Jensen's Inequality). *Let $X \in \mathcal{L}^2$ and let $\phi : \mathbb{R} \rightarrow (-\infty, \infty]$ be a convex function. Then,*

$$\mathbb{E}[\phi(X)] \geq \phi(\mathbb{E}[X]).$$

Proposition 1.2. *Let $\mathcal{G} \subseteq \mathcal{F}$ be a σ -algebra.*

1. *If (X_n) is an increasing sequence of random variables with $X_n \geq 0$ for all n and $X_n \uparrow X$, then*

$$\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}],$$

almost-surely (conditional MCT).

2. *If $X_n \geq 0$, then*

$$\mathbb{E}[\liminf_n X_n|\mathcal{G}] \leq \liminf_n \mathbb{E}[X_n|\mathcal{G}]$$

(conditional Fatou's lemma).

3. *If $X_n \rightarrow X$ and $|X_n| \leq Y$ almost-surely for all n and $Y \in \mathcal{L}^1$, then*

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}],$$

almost-surely (conditional DCT).

4. *If $X \in \mathcal{L}^\infty$ and $\phi : \mathbb{R} \rightarrow (-\infty, \infty]$ is convex such that either $\phi(X) \in \mathcal{L}^1$ or $\phi(X) \geq 0$, then:*

$$\mathbb{E}[\phi(X)|\mathcal{G}] \geq \phi(\mathbb{E}[X|\mathcal{G}])$$

almost surely (convex Jensen's). In particular, for all $1 \leq p < \infty$,

$$\|\mathbb{E}[X|\mathcal{G}]\|_p \leq \|X\|_p.$$

Proof:

1. Let Y_n be a version of $\mathbb{E}[X_n|\mathcal{G}]$. Since $0 \leq X_n \uparrow X$ as $n \rightarrow \infty$, we have that $Y_n \geq 0$ and are increasing.

Define $Y = \limsup_{n \rightarrow \infty} Y_n$. We will show that $Y = \mathbb{E}[X|\mathcal{G}]$.

- Y is \mathcal{G} -measurable as it is a lim sup of \mathcal{G} -measurable random variables.
- For $A \in \mathcal{G}$,

$$\mathbb{E}[X \mathbb{1}_A] \stackrel{MCT}{=} \lim_{n \rightarrow \infty} \mathbb{E}[X_n \mathbb{1}_A] = \lim_{n \rightarrow \infty} \mathbb{E}[Y_n \mathbb{1}_A] \stackrel{MCT}{=} \mathbb{E}[Y \mathbb{1}_A].$$

2. The sequence $\inf_{k \geq n} X_k$ is increasing in n . Moreover,

$$\lim_{n \rightarrow \infty} \inf_{k \geq n} X_k = \liminf_{n \rightarrow \infty} X_n.$$

By 1,

$$\lim_{n \rightarrow \infty} \mathbb{E}[\inf_{k \geq n} X_k | \mathcal{G}] = \mathbb{E}[\liminf_{n \rightarrow \infty} X_n | \mathcal{G}].$$

But also,

$$\mathbb{E}[\inf_{k \geq n} X_k | \mathcal{G}] \leq \inf_{k \geq n} \mathbb{E}[X_k | \mathcal{G}]$$

almost-surely, by monotonicity. Hence taking limits, we get Fatou's lemma.

3. Since $X_n + Y$, $Y - X_n$ give a sequence of random variables which are non-negative,

$$\mathbb{E}[X + Y | \mathcal{G}] = \mathbb{E}[\liminf_n (X_n + Y) | \mathcal{G}] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[X_n + Y | \mathcal{G}]$$

almost-surely, and similarly

$$\mathbb{E}[Y - X | \mathcal{G}] = \mathbb{E}[\liminf_{n \rightarrow \infty} (Y - X_n) | \mathcal{G}] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[Y - X_n | \mathcal{G}]$$

almost-surely. Combining these inequalities,

$$\limsup_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}] \leq \mathbb{E}[X | \mathcal{G}] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}].$$

This can only hold if $\lim_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}] = \mathbb{E}[X | \mathcal{G}]$, almost-surely.

4. We use the fact that a convex function can be written as a supremum of countably many affine functions:

$$\phi(x) = \sup_i (a_i x + b_i).$$

Hence we get

$$\mathbb{E}[\phi(X)|\mathcal{G}] \geq a_i \mathbb{E}[X|\mathcal{G}] + b_i$$

for all i almost-surely, hence

$$\mathbb{E}[\phi(X)|\mathcal{G}] \geq \sup_i (a_i \mathbb{E}[X|\mathcal{G}] + b_i) = \phi(\mathbb{E}[X|\mathcal{G}]),$$

almost-surely.

In particular, for $1 \leq p < \infty$,

$$\begin{aligned} \|\mathbb{E}[X|\mathcal{G}]\|_p^p &= \mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|^p] \leq \mathbb{E}[\mathbb{E}[|X|^p|\mathcal{G}]] \\ &= \mathbb{E}[|X|^p] = \|X\|_p^p. \end{aligned}$$

Proposition 1.3 (Tower Property). *Let $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ be σ -algebras, and $X \in \mathcal{L}^1$. Then,*

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$$

almost-surely.

Proof: Note that $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}]$ is \mathcal{H} -measurable. For $A \in \mathcal{H}$, we have

$$\begin{aligned} \mathbb{E}[\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] \mathbb{1}_A] &= \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \mathbb{1}_A] \\ &= \mathbb{E}[X \mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}] \mathbb{1}_A], \end{aligned}$$

since $A \in \mathcal{G}$.

Proposition 1.4 (Taking out what is known). *Let $X \in \mathcal{L}^1$, $\mathcal{G} \subseteq \mathcal{F}$ a σ -algebra. If Y is bounded and \mathcal{G} -measurable, then*

$$\mathbb{E}[XY|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]Y \quad \text{a.s.}$$

Proof: Y is \mathcal{G} -measurable, so $\mathbb{E}[X|\mathcal{G}]Y$ is \mathcal{G} -measurable.

For $A \in \mathcal{G}$,

$$\mathbb{E}[(\mathbb{E}[X|\mathcal{G}]Y) \mathbb{1}_A] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}](Y \mathbb{1}_A)] = \mathbb{E}[XY \mathbb{1}_A].$$

Definition 1.3. Let \mathcal{A} be a collection of subsets of Ω . Then \mathcal{A} is a π -system if for all $A, B \in \mathcal{A}$, $A \cap B \in \mathcal{A}$, and $\emptyset \in \mathcal{A}$.

Theorem 1.7. Let μ_1, μ_2 be measures on (E, \mathcal{E}) . Suppose that \mathcal{A} is a π -system which generates \mathcal{E} , and

$$\mu_1(A) = \mu_2(A)$$

for all $A \in \mathcal{A}$, and $\mu_1(E) = \mu_2(E)$ is finite.

Then $\mu_1 = \mu_2$.

Proposition 1.5. Let $X \in \mathcal{L}^1$, and $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ be σ -algebras. If $\sigma(X, \mathcal{G})$ are independent of \mathcal{H} , then

$$\mathbb{E}[X|\sigma(\mathcal{G}, \mathcal{H})] = \mathbb{E}[X|\mathcal{G}] \quad \text{a.s.}$$

Proof: Without loss of generality, $X \geq 0$, since the general case comes from decomposing $X = X^+ - X^-$. Let $A \in \mathcal{G}$, $B \in \mathcal{H}$. Then,

$$\begin{aligned} \mathbb{E}[\mathbb{E}[X|\sigma(\mathcal{G}, \mathcal{H})]\mathbb{1}_{A \cap B}] &= \mathbb{E}[X\mathbb{1}_{A \cap B}] = \mathbb{E}[X\mathbb{1}_A]\mathbb{P}(B) \text{ (independence)} \\ &= \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_A]\mathbb{P}(B) = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_{A \cap B}]. \end{aligned}$$

Define two measures

$$\begin{aligned} \mu_1(F) &= \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}_F], \\ \mu_2(F) &= \mathbb{E}[\mathbb{E}[X|\sigma(\mathcal{G}, \mathcal{H})]\mathbb{1}_F]. \end{aligned}$$

These are measures on \mathcal{F} which agree on the π -system $\{A \cap B \mid A \in \mathcal{G}, B \in \mathcal{H}\}$, generating $\sigma(\mathcal{G}, \mathcal{H})$. Moreover since $X \in \mathcal{L}^1$,

$$\mu_1(F) = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X] = \mu_2(F) < \infty.$$

We can now use uniqueness of measures to see these two measures agree on $\sigma(\mathcal{G}, \mathcal{H})$, which can only occur if

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X|\sigma(\mathcal{G}, \mathcal{H})] \quad \text{a.s.}$$

1.4 Examples of Conditional Expectation

Example 1.1. (Gaussians)

Let (X, Y) be a Gaussian random vector in \mathbb{R}^2 . Our goal is to compute

$$X' = \mathbb{E}[X|\mathcal{G}],$$

where $\mathcal{G} = Y$. Since X' is a \mathcal{G} -measurable function, there exists a Borel

measurable function f so that $X' = f(Y)$. We want to find f .

We guess that $X' = aY + b$, for $a, b \in \mathbb{R}$. Then,

$$a\mathbb{E}[Y] + b = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X],$$

and moreover

$$\begin{aligned}\mathbb{E}[(X - X')Y] &= 0 \\ \implies \text{Cov}(X - X', Y) &= 0 \\ \implies \text{Cov}(X, Y) &= \text{Cov}(X', Y) = a \text{Var}(Y).\end{aligned}$$

We can now take a to satisfy this equation, so $\text{Cov}(X - X', Y) = 0$.

But since $(X - X', Y)$ is a Gaussian random variable with covariance 0, $X - X'$ and Y are independent.

Suppose that Z is a $\sigma(Y)$ -measurable random variable. Then Z is independent of $X - X'$, so

$$\mathbb{E}[(X - X')Z] = \mathbb{E}[X - X']\mathbb{E}[Z] = 0.$$

This shows the projection property. Hence

$$\mathbb{E}[X|\mathcal{G}] = aY + b,$$

where a, b are determined as before.

Example 1.2. (Conditional Density Functions)

Suppose that X, Y are random variables with a joint density function $f_{X,Y}$ on \mathbb{R}^2 . Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be Borel measurable so that $h(X)$ is integrable.

Our goal is to compute

$$\mathbb{E}[h(X)|Y] = \mathbb{E}[h(X)|\sigma(Y)].$$

The density for Y is given by

$$f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y) \, dx.$$

Let g be a bounded, measurable function. Then,

$$\begin{aligned}\mathbb{E}[h(X)g(Y)] &= \int h(x)g(y)f_{X,Y}(x,y) \, dx \, dy \\ &= \int \left(\int h(x) \frac{f_{X,Y}(x,y)}{f_Y(y)} \, dx \right) g(y)f_Y(y) \, dy,\end{aligned}$$

where if $f_Y(y) = 0$, the inner integral is 0. We set

$$\phi(y) = \begin{cases} \int h(x)f_{X,Y}(x,y)/f_Y(y) \, dx & \text{if } f_Y(y) > 0, \\ 0 & \text{else.} \end{cases}$$

Then,

$$\mathbb{E}[h(X)|Y] = \phi(Y) \quad \text{a.s.}$$

since $\phi(Y)$ is $\sigma(Y)$ -measurable and satisfies the property defining the conditional expectation.

We interpret this computation as giving that

$$\mathbb{E}[h(X)|Y] = \int_{\mathbb{R}} h(x)\nu(Y, dx),$$

where

$$\begin{aligned}\nu(y, dx) &= \frac{f_{X,Y}(x,y)}{f_Y(y)} \mathbb{1}_{f_Y(y) > 0} \, dx \\ &= f_{X|Y}(x|y) \, dx.\end{aligned}$$

$\nu(y, dx)$ gives the *conditional distribution* of X given $Y = y$, and $f_{X|Y}(x|y)$ is the *conditional density function* of X given $Y = y$.

In this case, the conditional expectation corresponds to an actual expectation. This corresponds to a regular conditional probability distribution.

Also note $f_{X|Y}(x|y)$ is only defined up to a set of measure 0.

2 Discrete Time Martingales

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and (E, \mathcal{E}) a measurable space.

A sequence of random variables $X = (X_n)$ with values in E is a (discrete time) *stochastic process*. A *filtration* (\mathcal{F}_n) is a sequence of σ -algebras with $\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \mathcal{F}$.

The *natural filtration* (\mathcal{F}_n^X) associated with X is

$$\mathcal{F}_n^X = \sigma(X_1, \dots, X_n).$$

We say that X is *adapted* to (\mathcal{F}_n) if X_n is \mathcal{F}_n -measurable for all n . X is always adapted to its natural filtration. Say that X is *integrable* if $X_n \in \mathcal{L}^1$ for all n .

Definition 2.1. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$ be a filtered probability space and let X be a stochastic process which is integrable, adapted and real valued.

- X is a *martingale* (MG) if

$$\mathbb{E}[X_n | \mathcal{F}_m] = X_m \quad \text{a.s.}$$

whenever $n \geq m$.

- X is a *supermartingale* (sup MG) if

$$\mathbb{E}[X_n | \mathcal{F}_m] \leq X_m \quad \text{a.s.}$$

whenever $n \geq m$.

- X is a *submartingale* (sub MG) if

$$\mathbb{E}[X_n | \mathcal{F}_m] \geq X_m \quad \text{a.s.}$$

whenever $n \geq m$.

If X is a MG/sup MG/sub MG, then it is also a MG/sup MG/sub MG with respect to its natural filtration.

Example 2.1.

1. Say

$$X_n = \sum_{i=1}^n \xi_i,$$

where (ξ_i) are IID, integrable and $\mathbb{E}[\xi_i] = 0$.

2. We could also have

$$X_n = \prod_{i=1}^n \xi_i,$$

where (ξ_i) are IID, integrable and $\mathbb{E}[\xi_i] = 1$.

3. Choosing

$$X_n = \mathbb{E}[Z | \mathcal{F}_n],$$

where $Z \in \mathcal{L}^1$ and (\mathcal{F}_n) is a filtration, this gives a martingale.

Martingales are very useful for:

- computations (optimal stopping theorem),
- bounds (Doob's inequalities),
- proving theorems (martingale convergence theorem).

2.1 Stopping Times

Definition 2.2. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$ be a filtered probability space. A *stopping time* is a random variable $T : \Omega \rightarrow \mathbb{Z}_+ \cup \{\infty\}$ with

$$\{T \leq n\} \in \mathcal{F}_n,$$

for all n .

This is equivalent to $\{T = n\} \in \mathcal{F}_n$, for discrete time.

Example 2.2.

1. Constant (deterministic) times.
2. First hitting times: if (X_n) is an adapted stochastic process with values in \mathbb{R} , and $A \in \mathcal{B}(\mathbb{R})$, then

$$T_A = \inf\{n \geq 0 \mid X_n \in A\}.$$

3. Last exit times are not always stopping times.

Proposition 2.1. Suppose that $S, T, (T_n)$ are stopping times on $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$. Then:

$$\begin{array}{lll} S \wedge T, & S \vee T & \inf T_n, \\ \sup T_n, & \liminf T_n, & \limsup T_n, \end{array}$$

are stopping times

Definition 2.3. Let T be a stopping time on $(\Omega, \mathcal{F}, (\mathcal{F}_n), \mathbb{P})$. Then

$$\mathcal{F}_T = \{A \in \mathcal{F} \mid A \cap \{T \leq n\} \in \mathcal{F}_n \text{ for all } n\}$$

is called the *stopped σ -algebra*.

For $T = n$ deterministic, $\mathcal{F}_T = \mathcal{F}_n$. For X a stochastic process, $X_T = X_{T(\omega)}(\omega)$, whenever $T(\omega) < \infty$.

The *stopped process* X^T is defined by

$$X_n^T = X_{n \wedge T}.$$

Proposition 2.2. Let S, T be stopping times and X an adapted process. Then,

- (i) If $S \leq T$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$.
- (ii) $X_T \mathbb{1}_{\{T < \infty\}}$ is \mathcal{F}_T -measurable.
- (iii) X^T is adapted.
- (iv) If X is integrable, so is X^T .

Proof:

(i) This is immediate from definition.

(ii) Take $A \in \mathcal{E}$, then

$$\{X_T \mathbb{1}_{\{T < \infty\}} \in A\} \cap \{T \leq n\} = \bigcup_{k=0}^n (\{X_k \in A\} \cap \{T = k\}) \in \mathcal{F}_n.$$

Since X is adapted, $\{T = k\} \in \mathcal{F}_k$.

(iii) For all n , $X_{T \wedge n}$ is $\mathcal{F}_{T \wedge n}$ -measurable by (ii), so they are \mathcal{F}_n -measurable by (i).

(iv) Note

$$\begin{aligned} \mathbb{E}[|X_{T \wedge n}|] &= \mathbb{E}\left[\sum_{k=0}^{n-1} |X_k| \mathbb{1}(T = k)\right] + \mathbb{E}\left[\sum_{k=n}^{\infty} |X_n| \mathbb{1}(T = k)\right] \\ &\leq \sum_{k=0}^n \mathbb{E}[|X_k|] < \infty. \end{aligned}$$

Theorem 2.1 (Optional Stopping Theorem). *Let (X_n) be a MG. Then:*

(i) *If T is a stopping time, then X^T is a MG, hence*

$$\mathbb{E}[X_{T \wedge n}] = \mathbb{E}[X_0]$$

for all n .

(ii) *If $S \leq T$ are bounded stopping times, then*

$$\mathbb{E}[X_T | \mathcal{F}_S] = X_S \quad \text{a.s.}$$

(iii) *If $S \leq T$ are bounded stopping times, then*

$$\mathbb{E}[X_T] = \mathbb{E}[X_S].$$

(iv) *If there exists an integrable random variable Y such that $|X_n| \leq Y$ for all n , then for all almost-surely finite stopping times T ,*

$$\mathbb{E}[X_T] = \mathbb{E}[X_0].$$

(v) *If X has bounded increments, so $|X_{n+1} - X_n| \leq M$, and T is a stopping time with finite expectation, then*

$$\mathbb{E}[X_T] = \mathbb{E}[X_0].$$

Proof:

(i) By the tower property, we only need to show that

$$\mathbb{E}[X_{T \wedge n} | \mathcal{F}_{n-1}] = X_{T \wedge (n-1)} \quad \text{a.s.}$$

We have

$$\begin{aligned} \mathbb{E}[X_{T \wedge n} | \mathcal{F}_{n-1}] &= \mathbb{E} \left[\sum_{k=0}^{n-1} X_k \mathbb{1}(T = k) \mid \mathcal{F}_{n-1} \right] + \mathbb{E}[X_n \mathbb{1}(T > n-1) | \mathcal{F}_{n-1}] \\ &= X_T \mathbb{1}_{(T < n-1)} + X_{n-1} \mathbb{1}_{(T > n-1)} = X_{T \wedge (n-1)}. \end{aligned}$$

Since $(T > n-1) \in \mathcal{F}_{n-1}$, $\mathbb{E}[X_n | \mathcal{F}_{n-1}] = X_{n-1}$.

(ii) Suppose that $T \leq n$, and $S \leq T$. Then,

$$\begin{aligned} X_T &= (X_T - X_{T-1}) + \cdots + (X_{S+1} - X_S) + X_S \\ &= X_S + \sum_{k=0}^n (X_{k+1} - X_k) \mathbb{1}_{S \leq k < T}. \end{aligned}$$

Let $A \in \mathcal{F}_S$. Then,

$$\begin{aligned} \mathbb{E}[X_T \mathbb{1}_A] &= \mathbb{E}[X_S \mathbb{1}_A] + \mathbb{E} \left[\sum_{k=0}^n (X_{k+1} - X_k) \mathbb{1}_A \mathbb{1}(S \leq k < T) \right] \\ &= \mathbb{E}[X_S \mathbb{1}_A], \end{aligned}$$

since $\{S \leq k < T\} \cap A \in \mathcal{F}_k$ for all k , and X is a MG.

(iii) This follows from taking expectations in (ii).

(iv) and (v) are on the example sheet.

Remark.

1. The theorem holds with supermartingales or submartingales in place of a martingale, with the corresponding inequalities.
2. The assumptions in the theorem are important. Let (ξ_k) be iid, with $\mathbb{P}(\xi_1 = 1) = \mathbb{P}(\xi_1 = -1) = \frac{1}{2}$, and

$$X_n = \sum_{j=1}^n \xi_j.$$

Let $T = \inf\{n \geq 0 \mid X_n = 1\}$. Then T is a stopping time with $\mathbb{P}(T < \infty) = 1$. From the optional stopping theorem,

$$\mathbb{E}[X_{T \wedge n}] = \mathbb{E}[X_0] = 0$$

for all n , but

$$\mathbb{E}[X_T] = 1 \neq 0.$$

We cannot apply (v) above as $\mathbb{E}[T] = \infty$.

Proposition 2.3. *Suppose that X is a non-negative supermartingale. Then for any almost-surely finite stopping time T , we have that*

$$\mathbb{E}[X_T] \leq \mathbb{E}[X_0].$$

Proof: From the optional stopping theorem, $\mathbb{E}[X_{T \wedge n}] \leq \mathbb{E}[X_0]$ for all n .
Apply Fatou's lemma, since $X \geq 0$.

Example 2.3. (Gambler's Ruin)

Let (ξ_k) be iid with $\mathbb{P}(\xi_1 = 1) = \mathbb{P}(\xi_1 = -1) = \frac{1}{2}$, and

$$X_n = \sum_{i=1}^n \xi_i.$$

Let $T_c = \inf\{n \geq 0 \mid X_n = c\}$, and $T = T_{-a} \wedge T_b$ for $a, b > 0$.

X clearly has bounded increments. We now claim that $\mathbb{E}[T] < \infty$. This is as T is at most the first time that $a + b$ consecutive $\xi_i = 1$.

Hence the probability of $a + b$ consecutive $+1$'s is $2^{-(a+b)}$, hence

$$T \leq (a + b) \cdot \text{Geo}(2^{-(a+b)}).$$

Hence,

$$\mathbb{E}[T] \leq (a + b)2^{a+b} < \infty.$$

Hence from the optional stopping theorem,

$$\mathbb{E}[X_T] = \mathbb{E}[X_0] = 0.$$

Also,

$$\mathbb{E}[X_T] = (-a)\mathbb{P}(T_{-a} < T_b) + b\mathbb{P}(T_b < T_{-a}) = 0.$$

We also know that $\mathbb{P}(T_{-a} < T_b) + \mathbb{P}(T_b < T_{-a}) = 1$. Solving this system of equations,

$$\mathbb{P}(T_{-a} < T_b) = \frac{b}{a + b}.$$

We can also compute $\mathbb{E}[T]$ using the fact that $X_n^2 - n$ is also a martingale.

2.2 Martingale Convergence Theorem

The goal of this section is to prove the following:

Theorem 2.2 (Martingale Convergence Theorem). *Let $X = (X_n)$ be a super-*

martingale with

$$\sup_n \mathbb{E}[|X_n|] < \infty.$$

Then $X_n \rightarrow X_\infty$ almost surely as $n \rightarrow \infty$, where $X_\infty \in \mathcal{L}^1(\mathcal{F}_\infty)$, where

$$\mathcal{F}_\infty = \sigma(\mathcal{F}_n \mid n \geq 0).$$

Corollary 2.1. *Let X be a non-negative supermartingale. Then X converges almost surely to a finite limit.*

Let's show this corollary first.

Proof: Since $X_n \geq 0$,

$$\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \leq \mathbb{E}[X_0] < \infty.$$

So X is \mathcal{L}^1 bounded, and we can apply martingale convergence theorem.

Quick detour for some real analysis. Suppose that $x = (x_n)$ is a sequence in \mathbb{R} . Fix $a < b$, set $T_0(x) = 0$, and

$$\begin{aligned} S_{k+1}(x) &= \inf\{n \geq T_k(x) \mid x_n \leq a\}, \\ T_{k+1}(x) &= \inf\{n \geq S_{k+1}(x) \mid x_n \geq b\}, \\ N_n([a, b], x) &= \sup\{k \geq 0 \mid T_k(x) \leq n\}, \end{aligned}$$

the number of up-crossings of $[a, b]$ by x before time n . Then,

$$N_n([a, b], x) \uparrow N([a, b], x) = \sup\{k \geq 0 \mid T_k < \infty\},$$

which is the number of up-crossings of $[a, b]$.

Lemma 2.1. *A sequence in $x = (x_n)$ in \mathbb{R} converges in $\mathbb{R} = \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ if and only if $N([a, b], x) < \infty$ for all $a < b$, $a, b \in \mathbb{Q}$.*

Proof: We do the easy direction. Suppose that x converges. Then if there exists $a < b$ with $N([a, b], x) = \infty$, then

$$\liminf_n x_n \leq a, \quad \limsup_n x_n \geq b,$$

a contradiction to convergence.

Conversely, if x does not converge, then there exists $a < b$ with $\liminf x_n <$

$a < b < \limsup x_n$, and $a, b \in \mathbb{Q}$. Then

$$N([a, b], x) = \infty.$$

One key inequality for the proof is the following:

Theorem 2.3 (Doob's Upcrossing Inequality). *Let X be a supermartingale, and let $a < b$. Then for all n ,*

$$(b - a)\mathbb{E}[N_n([a, b], X)] \leq \mathbb{E}[(X_n - a)^-].$$

Proof: Let $N = N_n([a, b], X)$. Then,

$$X_{T_k} - X_{S_k} \geq b - a.$$

By definition of N ,

$$\sum_{k=1}^n (X_{T_k \wedge n} - X_{S_k \wedge n}) = \sum_{k=1}^N (X_{T_k} - X_{S_k}) + (X_n - X_{S_{N+1}})\mathbb{1}_{(S_{N+1} \leq n)}. \quad (**)$$

Now $(T_k), (S_k)$ are stopping times, and $S_k \wedge n \leq T_k \wedge n$ are bounded stopping times, so from OST,

$$\mathbb{E}[X_{S_k \wedge n}] \geq \mathbb{E}[X_{T_k \wedge n}]$$

for all k . Taking the expectation of $(**)$,

$$\begin{aligned} 0 &\geq \mathbb{E} \left[\sum_{k=1}^n (X_{T_k \wedge n} - X_{S_k \wedge n}) \right] \\ &\geq (b - a)\mathbb{E}[N] - \mathbb{E}[(X_n - a)^-], \end{aligned}$$

since from the first equality,

$$(X_n - X_{S_{N+1}})\mathbb{1}_{S_{N+1} \leq n} \geq -(X_n - a)^-.$$

Now we prove martingale convergence theorem.

Proof: Let $a < b$, $a, b \in \mathbb{Q}$. Doob's upcrossing inequality means

$$\mathbb{E}[N_n([a, b], X)] \leq \frac{1}{b - a} \mathbb{E}[(X_n - a)^-].$$

From monotone convergence theorem,

$$\mathbb{E}[N([a, b], X)] \leq \frac{1}{b-a} \sup_n \mathbb{E}[|X_n| + a] < \infty.$$

In particular, $N([a, b], X) < \infty$ almost surely, for all $a < b$, $a, b \in \mathbb{Q}$. Hence

$$\Omega_0 = \bigcap_{\substack{a < b \\ a, b \in \mathbb{Q}}} \{N([a, b], X) < \infty\} \implies \mathbb{P}(\Omega_0) = 1,$$

since it is an intersection of countably many almost-sure events. Set

$$X_\infty = \begin{cases} \lim_n X_n & \text{on } \Omega_0, \\ 0 & \text{on } \Omega \setminus \Omega_0. \end{cases}$$

Then X_∞ is \mathcal{F}_∞ -measurable, and

$$\begin{aligned} \mathbb{E}[|X_\infty|] &= \mathbb{E}[\liminf_n |X_n|] \\ &\leq \liminf_n \mathbb{E}[|X_n|] < \infty, \end{aligned}$$

by Fatou's, so $X_\infty \in \mathcal{L}^1$.

From conditional Jensen's on $|x|$, we for a martingale X ,

$$\mathbb{E}[|X_n| | \mathcal{F}_m] \geq |X_m| \quad \text{a.s.}$$

So if X is a martingale, $|X|$ is a non-negative submartingale.

Theorem 2.4 (Doob's Maximal Inequality). *Let X be a non-negative submartingale, and $X_n^* = \sup_{0 \leq k \leq n} X_k$. Then for all $\lambda \geq 0$,*

$$\lambda \mathbb{P}[X_n^* \geq \lambda] \leq \mathbb{E}[X_n \mathbb{1}(X_n^* \geq \lambda)] \leq \mathbb{E}[X_n].$$

This can be thought of as a better version of Markov's inequality, that applies to submartingales.

Proof: Let $T = \inf\{k \geq 0 \mid X_k \geq \lambda\}$. Then $T \wedge n$ is a bounded stopping time. By optional stopping theorem,

$$\begin{aligned} \mathbb{E}[X_n] &\geq \mathbb{E}[X_{T \wedge n}] = \mathbb{E}[X_T \mathbb{1}_{T \leq n}] + \mathbb{E}[X_n \mathbb{1}_{T > n}] \\ &\geq \lambda \mathbb{P}(T \leq n) + \mathbb{E}[X_n \mathbb{1}_{T > n}]. \end{aligned}$$

Note that $\{T \leq n\} = \{X_n^* \geq \lambda\}$, so rearranging,

$$\lambda \mathbb{P}(X_n^* \geq \lambda) \leq \mathbb{E}[X_n \mathbb{1}_{T \leq n}] \leq \mathbb{E}[X_n].$$

Theorem 2.5 (Doob's \mathcal{L}^p Inequality). *Let X be a non-negative submartingale. For $p > 1$ and with $X_n^* = \sup_{0 \leq k \leq n} X_k$, we have that*

$$\|X_n^*\|_p \leq \frac{p}{p-1} \|X_n\|_p.$$

Proof: Fix $k < \infty$. We have that

$$\begin{aligned} \mathbb{E}[(X_n^* \wedge k)^p] &= \mathbb{E} \left[\int_0^k p x^{p-1} \mathbb{1}(X_n^* \geq x) dx \right] \\ &= \int_0^k p x^{p-1} \mathbb{P}(X_n^* \geq x) dx && \text{(Fubini's)} \\ &\leq \int_0^k p x^{p-2} \mathbb{E}[X_n \mathbb{1}(X_n^* > x)] dx \\ &= \frac{p}{p-1} \mathbb{E}[X_n (X_n^* \wedge k)^{p-1}] && \text{(Fubini's)} \\ &\leq \frac{p}{p-1} \|X_n\|_p \|X_n^* \wedge k\|_p^{p-1} && \text{(Hölder's).} \end{aligned}$$

This gives us

$$\|X_n^* \wedge k\|_p \leq \frac{p}{p-1} \|X_n\|_p.$$

Applying MCT, we can send $k \rightarrow \infty$, and get

$$\|X_n^*\|_p \leq \frac{p}{p-1} \|X_n\|_p.$$

2.3 \mathcal{L}^p Convergence of Martingales

Theorem 2.6. *Let X be a martingale, and $p > 1$. The following are equivalent:*

(i) X is \mathcal{L}^p bounded:

$$\sup_{n \geq 0} \|X_n\|_p < \infty.$$

(ii) X converges almost-surely and in \mathcal{L}^p to X_∞ .

(iii) There exists $Z \in \mathcal{L}^p$ such that $X_n = \mathbb{E}[Z | \mathcal{F}_n]$ almost-surely, for all n .

Proof:

(i) \implies (ii): Suppose that X is \mathcal{L}^p bounded. Then from Jensen, X is also \mathcal{L}^1 . So martingale convergence theorem gives $X_n \rightarrow X_\infty$ almost-surely. Then,

$$\mathbb{E}[|X_\infty|^p] = \mathbb{E}[\liminf_n |X_n|^p] \leq \liminf_n \mathbb{E}[|X_n|^p] \leq \sup_n \|X_n\|_p^p < \infty.$$

Doob's \mathcal{L}^p inequality gives

$$\|X_n^*\|_p \leq \frac{p}{p-1} \|X_n\|_p,$$

where $X_n^* = \sup_{0 \leq k \leq n} |X_k|$. From monotone convergence theorem,

$$\|X_\infty^*\|_p \leq \frac{p}{p-1} \sup_n \|X_n\|_p < \infty.$$

Thus, $|X_n - X_\infty| \leq 2X_\infty^* \in \mathcal{L}^p$. So from dominated convergence theorem, $X_n \rightarrow X_\infty$ in \mathcal{L}^p .

(ii) \implies (iii): Let $Z = X_\infty$, where $Z \in \mathcal{L}^p$. We will show that $X_n = \mathbb{E}[Z|\mathcal{F}_n]$. Now, for any $m \geq n$,

$$\begin{aligned} \|X_n - \mathbb{E}[X_\infty|\mathcal{F}_n]\|_p &= \|\mathbb{E}[X_m - X_\infty|\mathcal{F}_n]\|_p \\ &\leq \|X_m - X_\infty\|_p \rightarrow 0, \end{aligned}$$

hence $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$ almost-surely.

(iii) \implies (i): Let $X_n = \mathbb{E}[Z|\mathcal{F}_n]$, then

$$\|X_n\|_p \leq \|Z\|_p < \infty,$$

so these are \mathcal{L}^p -bounded.

Definition 2.4. A martingale of the form $\mathbb{E}[Z|\mathcal{F}_n]$ for $Z \in \mathcal{L}^p$ is *closed* in \mathcal{L}^p .

Corollary 2.2. If $Z \in \mathcal{L}^p$, $X_n = \mathbb{E}[Z|\mathcal{F}_n]$, then if we let $\mathcal{F}_\infty = \sigma(\mathcal{F}_n, n \geq 0)$, we have that

$$X_n \rightarrow X_\infty = \mathbb{E}[Z|\mathcal{F}_\infty] \quad \text{a.s.}$$

as $n \rightarrow \infty$, and also in \mathcal{L}^p .

Proof: From the theorem, $X_n \rightarrow X_\infty$ almost-surely, and in \mathcal{L}^p . What we want to show is that $X_\infty = \mathbb{E}[Z|\mathcal{F}_\infty]$.

Clearly X_∞ is \mathcal{F}_∞ -measurable. Fix $A \in \bigcup \mathcal{F}_n$, then $A \in \mathcal{F}_N$ for some N . For $n \geq N$,

$$\mathbb{E}[Z\mathbb{1}_A] = \mathbb{E}[X_n\mathbb{1}_A] \rightarrow \mathbb{E}[X_\infty\mathbb{1}_A],$$

hence

$$\mathbb{E}[X_\infty\mathbb{1}_A] = \mathbb{E}[\mathbb{E}[Z|\mathcal{F}_\infty]\mathbb{1}_A].$$

Note $\bigcup \mathcal{F}_n$ is a π -system which generates \mathcal{F}_∞ , so this holds for all A . Hence this gives $\mathbb{E}[Z|\mathcal{F}_\infty] = X_\infty$.

2.4 Uniformly Integrable Martingales

Definition 2.5. A collection of random variables $(X_i)_{i \in I}$ is called *uniformly integrable* (UI) if

$$\sup_{i \in I} \mathbb{E}[|X_i|\mathbb{1}(|X_i| > \alpha)] \rightarrow 0$$

as $\alpha \rightarrow \infty$. Alternatively, (X_i) is UI if (X_i) is \mathcal{L}^1 -bounded, and for all $\varepsilon > 0$, there exists $\delta > 0$ such that if $\mathbb{P}(A) < \delta$, then

$$\sup_{i \in I} \mathbb{E}[|X_i|\mathbb{1}_A] < \varepsilon.$$

If (X_i) is \mathcal{L}^p bounded for some $p > 1$, then the family is UI. If (X_i) is UI, then the family is \mathcal{L}^1 -bounded. However the converse is not true.

Theorem 2.7. Let $X \in \mathcal{L}^1$. Then

$$\{\mathbb{E}[X|\mathcal{G}] \mid \mathcal{G} \subseteq \mathcal{F} \text{ a } \sigma\text{-algebra}\}$$

is UI.

Proof: Since $X \in \mathcal{L}^1$, for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\mathbb{P}(A) < \delta \implies \mathbb{E}[|X|\mathbb{1}_A] < \varepsilon.$$

Choose $\lambda > 0$ such that $\mathbb{E}[|X|] \leq \lambda \cdot \delta$. Suppose that $\mathcal{G} \subseteq \mathcal{F}$, and let $Y = \mathbb{E}[X|\mathcal{G}]$. We know that $\mathbb{E}[|Y|] < \mathbb{E}[|X|] < \infty$.

Applying Markov's inequality,

$$\mathbb{P}(|Y| \geq \lambda) \leq \frac{\mathbb{E}[|Y|]}{\lambda} \leq \delta,$$

so

$$\mathbb{E}[|Y|\mathbb{1}(|Y| > \lambda)] \leq \mathbb{E}[|X|\mathbb{1}(|Y| \geq \lambda)] \leq \varepsilon,$$

by conditional Markov's.

Lemma 2.2. *Let (X_n) , X be \mathcal{L}^1 with $X_n \rightarrow X$ almost surely, then*

$$X_n \rightarrow X \text{ in } \mathcal{L}^\infty \iff (X_n) \text{ uniformly integrable.}$$

Definition 2.6. A martingale $X = (X_n)$ is called a *uniformly integrable martingale* if and only if X is a martingale, and (X_n) is a uniformly integrable family.

Theorem 2.8. *Let X be a martingale. The following are equivalent:*

- (i) X is a UI martingale.
- (ii) X_n converges almost surely and in \mathcal{L}^1 to X_∞ .
- (iii) There exists $Z \in \mathcal{L}^1$ such that $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ almost surely, for all n

Proof:

(i) \implies (ii): Since X is UI, it is \mathcal{L}^1 bounded, so $X_n \rightarrow X_\infty$ almost-surely, from martingale convergence theorem. Hence also $X_n \rightarrow X_\infty$ in \mathcal{L}^1 , as X is uniformly integrable.

(ii) \implies (iii): Set $Z = X_\infty$, so $Z \in \mathcal{L}^1$. We want to show that

$$X_n = \mathbb{E}[Z|\mathcal{F}_n] \quad \text{a.s.}$$

From the martingale property,

$$\begin{aligned} \|X_n - \mathbb{E}[X_\infty|\mathcal{F}_n]\|_1 &= \|\mathbb{E}[X_n - X_\infty|\mathcal{F}_n]\| \leq \|X_n - X_\infty\| \\ &\rightarrow 0, \end{aligned}$$

so $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$.

(iii) \implies (i): From the tower property, $\mathbb{E}[Z|\mathcal{F}_n]$ is a martingale, which is uniformly integrable from the previous theorem.

Remark.

1. As before, X is a UI martingale $\implies \mathbb{E}[Z|\mathcal{F}_\infty] = X_\infty$.
2. If X is a UI sub MG or sup MG, then $X_n \rightarrow X_\infty$ almost surely, and in \mathcal{L}^1 , where $\mathbb{E}[X_\infty|\mathcal{F}_n] \leq X_n$ for a submartingale.

3. Let (X_i) be iid with $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 2) = \frac{1}{2}$. Then

$$Y_n = \prod_{j=1}^n X_j$$

is a martingale which is \mathcal{L}^1 bounded and converges to 0 almost-surely. But $\mathbb{E}[Y_n] = 1$ for all n , so $Y_n \not\rightarrow Y_\infty$ in \mathcal{L}^1 .

For X a UI martingale, and T a stopping time, then we can consider

$$X_T = \sum_{n=0}^{\infty} X_n \mathbb{1}(T = n) + X_\infty \mathbb{1}(T = \infty).$$

Theorem 2.9 (Optional Stopping Theorem for UI Martingales). *Let X be a UI martingale, and S, T stopping times with $S \leq T$. Then,*

$$\mathbb{E}[X_T | \mathcal{F}_S] = X_S \quad \text{a.s.}$$

Proof: We claim that

$$\mathbb{E}[X_\infty | \mathcal{F}_T] = X_T \quad \text{a.s.}$$

for all stopping times T . First we show that $X_T \in \mathcal{L}^1$. Note that

$$|X_n| \leq \mathbb{E}[|X_\infty| | \mathcal{F}_n],$$

so

$$\begin{aligned} \mathbb{E}[|X_T|] &= \sum_{n=0}^{\infty} \mathbb{E}[|X_n| \mathbb{1}(T = n)] + \mathbb{E}[|X_\infty| \mathbb{1}(T = \infty)] \\ &\leq \sum_{n=0}^{\infty} \mathbb{E}[|X_\infty| \mathbb{1}(T = n)] = \mathbb{E}[|X_\infty|]. \end{aligned}$$

Let $B \in \mathcal{F}_T$. Then

$$\begin{aligned} \mathbb{E}[X_T \mathbb{1}_B] &= \sum_n \mathbb{E}[X_n \mathbb{1}(T = n) \mathbb{1}_B] \\ &= \sum_n \mathbb{E}[X_\infty \mathbb{1}(T = n) \mathbb{1}_B] \\ &= \mathbb{E}[X_\infty \mathbb{1}_B]. \end{aligned}$$

Since X_T is \mathcal{F}_T -measurable, this implies $\mathbb{E}[X_\infty|\mathcal{F}_T] = X_T$ almost-surely.

Suppose that $S \leq T$ are stopping times. Then from the tower property,

$$\begin{aligned}\mathbb{E}[X_T|\mathcal{F}_S] &= \mathbb{E}[\mathbb{E}[X_\infty|\mathcal{F}_T]|\mathcal{F}_S] \\ &= \mathbb{E}[X_\infty|\mathcal{F}_S] = X_S \quad \text{a.s.}\end{aligned}$$

2.5 Backwards Martingales

Let

$$\cdots \subseteq \mathcal{G}_{-2} \subseteq \mathcal{G}_{-1} \subseteq \mathcal{G}_0$$

be a sequence of σ -algebras indexed by \mathbb{Z}_- . Say that $X = (X_n)_{n \leq 0}$ is a *backwards martingale* if it is adapted, $X_n \in \mathcal{L}^1$, and

$$\mathbb{E}[X_{n+1}|\mathcal{G}_n] = X_n \quad \text{a.s.}$$

From the tower property, we have

$$\mathbb{E}[X_0|\mathcal{G}_n] = X_n \quad \text{a.s.}$$

for all $n \leq 0$. Since $X_0 \in \mathcal{L}^\infty$, X is UI.

Theorem 2.10. *Let X be a backwards martingale, $X_0 \in \mathcal{L}^p$, for $p \geq 1$. Then X_n converges almost-surely and in \mathcal{L}^p as $n \rightarrow -\infty$ to*

$$X_{-\infty} = \mathbb{E}[X_0|\mathcal{G}_{-\infty}],$$

where

$$\mathcal{G}_{-\infty} = \bigcap_{n \leq 0} \mathcal{G}_n.$$

Proof: We have multiple steps.

First we use Doob's upcrossing inequality for $a < b$. Let $N_{-n}([a, b], X)$ be the number of upcrossings of X across $[a, b]$ in the time interval from $-n$ to 0, and let

$$\mathcal{F}_k = \mathcal{G}_{-n+k},$$

for $0 \leq k \leq n$. Then:

- (\mathcal{F}_k) is increasing.
- (X_{-n+k}) is a martingale.

So from Doob,

$$(b - a)\mathbb{E}[N_{-n}([a, b], X)] \leq \mathbb{E}[(X_0 - a)^-].$$

As $n \rightarrow \infty$, $N_{-n}([a, b], X)$ goes to the number of upcrossings of X across $[a, b]$. So by monotone convergence theorem,

$$\mathbb{E}[N_{-\infty}([a, b], X)] < \infty.$$

Hence $X_m \rightarrow X_{-\infty}$ as $m \rightarrow -\infty$ almost-surely.

Next we need to argue \mathcal{L}^p convergence. If $X_0 \in \mathcal{L}^p$, then $X_n \in \mathcal{L}^p$ for all $n \leq 0$, by conditional Jensen's.

Hence by Fatou, $X_{-\infty} \in \mathcal{L}^p$. So to show \mathcal{L}^p convergence, we need

$$|X_n - X_{-\infty}|^p = |\mathbb{E}[X_0 - X_{-\infty} | \mathcal{G}_n]|^p \leq \mathbb{E}[|X_0 - X_{-\infty}|^p | \mathcal{F}_n].$$

However the item on the right is uniformly integrable, as it is a random variable conditioned on different σ -algebras. Hence $|X_n - X_{-\infty}|^p$ is a UI family that converges to 0 a.s., hence converges to 0 in \mathcal{L}^1 . So $X_n \rightarrow X_{-\infty}$ in \mathcal{L}^p .

The final thing to show is that

$$X_{-\infty} = \mathbb{E}[X_0 | \mathcal{G}_{-\infty}] \quad \text{a.s.}$$

We need to show that for all $A \in \mathcal{G}_{-\infty}$, that $\mathbb{E}[X_0 \mathbb{1}_A] = \mathbb{E}[X_{-\infty} \mathbb{1}_A]$.

As $A \in \mathcal{G}_{-n}$ for all n , by properties of martingales

$$\mathbb{E}[X_0 \mathbb{1}_A] = \mathbb{E}[X_n \mathbb{1}_A].$$

As $n \rightarrow -\infty$, from \mathcal{L}^1 convergence we get

$$\mathbb{E}[X_0 \mathbb{1}_A] = \mathbb{E}[X_{-\infty} \mathbb{1}_A].$$

2.6 Applications of Martingale Theory

Here are some nice applications.

Theorem 2.11 (Kolmogorov 0-1 Law). *Let (X_i) be IID random variables. Let*

$\mathcal{F}_n = \sigma(X_k \mid k \geq n)$, and

$$\mathcal{F}_\infty = \bigcap_{n \geq 0} \mathcal{F}_n.$$

Then \mathcal{F}_∞ is trivial, i.e. for all $A \in \mathcal{F}_\infty$, $\mathbb{P}(A) \in \{0, 1\}$.

Proof: Let $\mathcal{G}_n = \sigma(X_k \mid k \leq n)$, and $A \in \mathcal{F}_\infty$.

Since \mathcal{G}_n is independent of \mathcal{F}_{n+1} , we have

$$\mathbb{E}[\mathbb{1}_A | \mathcal{G}_n] = \mathbb{P}(A) \quad \text{a.s.}$$

From martingale convergence theorem,

$$\mathbb{E}[\mathbb{1}_A | \mathcal{G}_n] \rightarrow \mathbb{E}[\mathbb{1}_A | \mathcal{G}_\infty] \quad \text{a.s.}$$

where $G_\infty = \sigma(G_n \mid n \geq 0)$. So

$$\mathbb{E}[\mathbb{1}_A | \mathcal{G}_\infty] = \mathbb{1}_A = \mathbb{P}(A) \quad \text{a.s.}$$

since $F_\infty \subseteq G_\infty$. This gives $\mathbb{P}(A) \in \{0, 1\}$.

Theorem 2.12 (Strong Law of Large Numbers). *Let (X_i) be IID in \mathcal{L}^1 , and let $\mu = \mathbb{E}[X_1]$. Let $S_n = X_1 + X_2 + \dots + X_n$*

Then $S_n/n \rightarrow \mu$ almost-surely, and in \mathcal{L}^1 .

Proof: Let

$$\mathcal{G}_n = \sigma(S_n, S_{n+1}, S_{n+2}, \dots) = \sigma(S_n, X_{n+1}, X_{n+2}, \dots).$$

We claim that

$$(M_n)_{n \leq -1} = \frac{S_{-n}}{-n}$$

is a $(\mathcal{F}_n) = (\mathcal{G}_{-n})$ backwards martingale.

Indeed, fix $m \leq -1$, then

$$\begin{aligned} \mathbb{E}[M_{m+1} | \mathcal{F}_m] &= \mathbb{E}\left[\frac{S_{-m-1}}{-m-1} | \mathcal{G}_{-m}\right] = \mathbb{E}\left[\frac{S_{n-1}}{n-1} | \mathcal{G}_n\right] \\ &= \mathbb{E}\left[\frac{S_n - X_n}{n-1} | \mathcal{G}_n\right] = \frac{S_n}{n-1} - \mathbb{E}\left[\frac{X_n}{n-1} | \mathcal{G}_n\right]. \end{aligned}$$

Since X_n is independent of (X_{n+1}, \dots) , we have

$$\mathbb{E} \left[\frac{X_n}{n-1} | \mathcal{G}_n \right] = \frac{1}{n-1} \mathbb{E}[X_n | S_n].$$

By symmetry, $\mathbb{E}[X_k | S_n] = \mathbb{E}[X_1 | S_n]$ for all $1 \leq k \leq n$. Also, $\mathbb{E}[S_n | S_n] = S_n$, so since

$$\mathbb{E}[X_1 | S_n] + \mathbb{E}[X_2 | S_n] + \dots + \mathbb{E}[X_n | S_n] = S_n \implies \mathbb{E}[X_n | S_n] = S_n/n.$$

Putting this into the formula,

$$\mathbb{E} \left[\frac{S_{n-1}}{n-1} | \mathcal{G}_n \right] = \frac{S_n}{n-1} - \frac{S_n}{n(n-1)} = \frac{S_n}{n},$$

hence (S_n/n) is a backwards martingale.

By the backward martingale convergence theorem, $S_n/n \rightarrow Y$ almost-surely and in \mathcal{L}^1 . Since

$$Y = \lim_{n \rightarrow \infty} \frac{S_n}{n} = \lim_{n \rightarrow \infty} \frac{X_{k+1} + \dots + X_{k+n}}{n}$$

almost-surely, we have that Y is measurable with respect to $\mathcal{T}_k = \sigma(X_{k+1}, X_{k+2}, \dots)$, hence it is measurable with respect to

$$\bigcap_k \mathcal{T}_k.$$

From the Kolmogorov 0-1 law, this means that $\mathbb{P}(Y = c) = 1$. Then

$$c = \mathbb{E}[Y] = \mathbb{E}[\lim_{n \rightarrow \infty} S_n/n] = \lim_n \mathbb{E}[S_n/n] = \mu,$$

by \mathcal{L}^1 convergence.

Theorem 2.13 (Radon-Nikodym Theorem). *Let \mathbb{P}, \mathbb{Q} be probability measure on (Ω, \mathcal{F}) . Assume that \mathcal{F} is countably generated, i.e.*

$$\mathcal{F} = \sigma(F_n),$$

where (F_n) is countable. Then the follow are equivalent:

- (i) $\mathbb{P}(A) = 0 \implies \mathbb{Q}(A) = 0$, for all $A \in \mathcal{F}$, i.e. $\mathbb{Q} \ll \mathbb{P}$.

(ii) For all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $A \in \mathcal{F}$,

$$\mathbb{P}(A) < \delta \implies \mathbb{Q}(A) < \varepsilon.$$

(iii) There exists $X \geq 0$ so that

$$\mathbb{Q}(A) = \mathbb{E}[X \mathbb{1}_A]$$

for all $A \in \mathcal{F}$, where the expectation is taking with respect to \mathbb{P} .

The random variable X is called the *Radon-Nikodym derivative*, and is usually written $X = d\mathbb{Q}/d\mathbb{P}$.

This can be extended to finite or σ -finite measures, and the assumption that \mathcal{F} is countably generated can be removed.

Proof:

(i) \implies (ii): Suppose that (ii) does not hold, then there exists $\varepsilon > 0$ such that for all n , there is A_n with $\mathbb{P}(A_n) < 1/n^2$ and $\mathbb{Q}(A_n) \geq \varepsilon$.

By Borel-Cantelli, $\mathbb{P}(A_n \text{ i.o.}) = 0$, so $\mathbb{Q}(A_n \text{ i.o.}) = 0$. But,

$$\mathbb{Q}(A_n \text{ i.o.}) = \mathbb{Q}\left(\bigcap_n \bigcup_{k \geq n} A_k\right) = \lim_n \mathbb{Q}\left(\bigcup_{k \geq n} A_k\right) \geq \varepsilon,$$

which is a contradiction.

(ii) \implies (iii): Let $\mathcal{F}_n = \sigma(F_k \mid k \leq n)$. If we write

$$\mathcal{A}_n = \{H_1 \cap \cdots \cap H_N \mid H_i \in \{F_i, F_i^c\}\},$$

then $\mathcal{F}_n = \sigma(\mathcal{A}_n)$. Note that sets in \mathcal{A}_n are disjoint. Let

$$X_n(\omega) = \sum_{A \in \mathcal{A}_n} \frac{\mathbb{Q}(A)}{\mathbb{P}(A)} \mathbb{1}_A(\omega).$$

Since the sets in \mathcal{A}_n are disjoint, we have that

$$\mathbb{Q}(A) = \mathbb{E}[X_n \mathbb{1}_A],$$

for all $A \in \mathcal{F}_n$. Write $X_n = d\mathbb{Q}/d\mathbb{P}$ on \mathcal{F}_n . Note that $X_n \geq 0$.

Moreover, (X_n) is a martingale. Indeed, if $A \in \mathcal{F}_n$, then

$$\mathbb{E}[X_{n+1} \mathbb{1}_A] = \mathbb{Q}(A) = \mathbb{E}[X_n \mathbb{1}_A],$$

hence $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$. Moreover (X_n) is \mathcal{L}^1 bounded as $\mathbb{E}[X_n] = \mathbb{Q}(\Omega) = 1$.

So by martingale convergence theorem, $X_n \rightarrow X_\infty$ almost-surely as $n \rightarrow \infty$. We are also going to show that (X_n) is UI.

Fix $\varepsilon > 0$. Then there is $\delta > 0$ so that condition (ii) holds. Let $\lambda = 1/\varepsilon$. Then,

$$\mathbb{P}(X_n \geq \lambda) \leq \frac{\mathbb{E}[X_n]}{\lambda} = \frac{1}{\lambda} = \delta.$$

So from (ii),

$$\mathbb{E}[X_n \mathbb{1}(X_n > \lambda)] = \mathbb{Q}(X_n > \lambda) \leq \varepsilon.$$

Hence (X_n) is UI, and $X_n \rightarrow X_\infty$ in \mathcal{L}^1 , with $\mathbb{E}[X_\infty] = 1$.

Now we show that X_∞ is the required derivative. Indeed, for all $A \in \mathcal{F}_n$, $\mathbb{E}[X_n \mathbb{1}_A] = \mathbb{E}[X_\infty \mathbb{1}_A]$. Now set

$$\tilde{\mathbb{Q}}(A) = \mathbb{E}[X_\infty \mathbb{1}_A],$$

then $\mathbb{Q} = \tilde{\mathbb{Q}}$ on $\bigcup \mathcal{F}_n$, which is a generating π -system. Hence $\mathbb{Q} = \tilde{\mathbb{Q}}$.

(iii) \implies (i): Obvious.

3 Continuous Time Processes

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We want to define stochastic process indexed by $t \geq 0$, rather than $n \in \mathbb{N}$.

A *filtration* $(\mathcal{F}_t)_{t \geq 0}$ is an increasing sequence of σ -algebras, so if $s \leq t$, then $\mathcal{F}_s \subseteq \mathcal{F}_t$.

A collection of random variables (X_t) is called a *stochastic process*. We say that (X_t) is *adapted* if X_t is \mathcal{F}_t -measurable for all $t \geq 0$.

A stopping time T is a random variable with values in $[0, \infty]$ with

$$\{T \leq t\} \in \mathcal{F}_t.$$

Why is continuous time so different? There are many different measurability issues that arise.

In particular, in discrete time we can equip \mathbb{N} with the σ -algebra $\mathcal{P}(\mathbb{N})$, and $(\omega, n) \mapsto X_n(\omega)$ is measurable with respect to $\mathcal{F} \otimes \mathcal{P}(\mathbb{N})$.

For continuous time, then for fixed $t \geq 0$ the map $\omega \mapsto X_t(\omega)$ is a random variable, but it need not be that $(\omega, t) \mapsto X_t(\omega)$ is measurable with respect to $\mathcal{F} \otimes \mathcal{B}(\mathbb{R})$. An example is

$$X_t = \mathbb{1}_A(t),$$

where $A \subseteq \mathbb{R}$ is not Borel.

For $A \subseteq \mathbb{R}$, this means that

$$T_A = \inf\{t \geq 0 \mid X_t \in A\}$$

need not be a stopping time. This is as

$$\{T_A \leq t\} = \bigcup_{0 \leq s \leq t} \{T_A = s\} \notin \mathcal{F}_t$$

in general, since we have an uncountable union.

How do we fix these problems? We can:

- Require that $t \mapsto X_t(\omega)$ is continuous as this implies that $(\omega, t) \mapsto X_t(\omega)$ is measurable.
- Require that $t \mapsto X_t(\omega)$ is a *cadlag* (right continuous, and left hand limits exist). Also written as RCLL. So,

$$X_t = \lim_{s \downarrow t} X_s, \quad \lim_{s \uparrow t} X_s \text{ exists.}$$

In both cases, the process is determined by its values at a countable set of time, i.e. \mathbb{Q}_+ .

To see that $t \mapsto X_t(\omega)$ being continuous implies $(\omega, t) \mapsto X_t(\omega)$ is measurable, we first restrict to $t \in [0, 1]$. Now, note that

$$X_t(\omega) = \lim_{n \rightarrow \infty} \sum_{k=0}^{2^n-1} \mathbb{1}_{(t \in (k2^{-n}, (k+1)2^{-n}])} X_{k2^{-n}}(\omega)$$

is $\mathcal{F} \otimes \mathcal{B}([0, 1])$ -measurable, as it is a pointwise limit of measurable function.

We will define a few classes of random variables. Let

$$C(\mathbb{R}_+, E) = \{\text{continuous } x : \mathbb{R}_+ \rightarrow E\}$$

with the smallest σ -algebra which makes $\Pi_t : X \rightarrow X_t$ measurable for all t , and let

$$D(\mathbb{R}_+, E) = \{\text{cadlag functions } x : \mathbb{R}_+ \rightarrow E\}$$

with the same σ -algebra.

For a stopping time its σ -algebra is still

$$\mathcal{F}_T = \{A \in \mathcal{F} \mid A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t\}.$$

For X cadlag, set

$$X_T(\omega) = X_{T(\omega)}(\omega)$$

whenever $T(\omega) < \infty$. Set $X_t^T = X_{t \wedge T}$.

Proposition 3.1. *Let S, T be stopping times, and X a cadlag, adapted process. Then,*

- (i) $S \wedge T$ is a stopping time.
- (ii) $S \leq T$ implies $\mathcal{F}_S \subseteq \mathcal{F}_T$.
- (iii) $X_T \mathbb{1}_{(T < \infty)}$ is \mathcal{F}_T -measurable.
- (iv) X^T is adapted.

Proof:

(i) and (ii) follow from the definition, and (iv) follows from (iii): Note that $X_{T \wedge t}$ is $\mathcal{F}_{T \wedge t}$ -measurable from (iii), and $\mathcal{F}_{T \wedge t} \subseteq \mathcal{F}_t$.

So we are only required to show (iii). To do this, we claim the following: Z

is \mathcal{F}_T -measurable if and only if $Z\mathbb{1}(T \leq t)$ is \mathcal{F}_t -measurable for all t .

In the forwards direction, if Z is \mathcal{F}_T -measurable, then $Z\mathbb{1}(T \leq t)$ is \mathcal{F}_t -measurable for all t by definition.

For the other way, if $Z = c\mathbb{1}_A$, then $Z\mathbb{1}(T \leq c)$ is \mathcal{F}_t -measurable for all t , then Z is \mathcal{F}_T -measurable.

For a general random variable Z , we can take finite linear combinations to prove this for simple functions, and then take limits.

To prove that $X_T\mathbb{1}(T < \infty)$ is \mathcal{F}_T -measurable, we will show that $X_T\mathbb{1}(T \leq t)$ is \mathcal{F}_t -measurable for all t . We can write

$$X_T\mathbb{1}(T \leq t) = X_T\mathbb{1}(T < t) + X_t\mathbb{1}(T = t).$$

Clearly $X_t\mathbb{1}(T = t)$ is \mathcal{F}_t -measurable, so it suffices to show $X_T\mathbb{1}(T < t)$ is \mathcal{F}_t -measurable.

Let $T_n = 2^{-n}\lceil 2^n T \rceil$. Then T_n is a stopping time with values in $\mathcal{D}_n = \{k2^{-n} \mid k \in \mathbb{N}\}$. Indeed,

$$\{T_n \leq t\} = \{\lceil 2^n T \rceil \leq 2^n t\} = \{T \leq 2^{-n}\lceil 2^n t \rceil\} \in \mathcal{F}_{2^{-n}\lceil 2^n t \rceil} \subseteq \mathcal{F}_t.$$

Now $T_n \downarrow T$ as $n \rightarrow \infty$. Since X is cadlag,

$$X_T = \mathbb{1}(T < t) = \lim_{n \rightarrow \infty} X_{T_n \wedge t} \mathbb{1}(T < t).$$

Since T_n takes on countably many values,

$$X_{T_n \wedge t} \mathbb{1}(T < t) = \sum_{\substack{d \in \mathcal{D}_n \\ d < t}} X_d \mathbb{1}(T_n = d) + X_t \mathbb{1}(T_n > t, T < t).$$

Since T_n is a stopping time, $X_{T_n \wedge t} \mathbb{1}(T < t)$ is \mathcal{F}_t -measurable for all n .

Example 3.1. (Hitting times are not always stopping times)

Let J be a random variable with $\mathbb{P}(J = 1) = \mathbb{P}(J = -1) = 1/2$. Let

$$X_t = \begin{cases} t & 0 \leq t \leq 1, \\ 1 + J(t - 1) & t > 1. \end{cases}$$

Let $\mathcal{F}_t = \sigma(X_s \mid s \leq t)$, the natural filtration, and let $A = (1, 2)$, with

$$T_A = \inf\{t \geq 0 \mid X_t \in A\}.$$

Then $\{T_A \leq 1\} \notin \mathcal{F}_1$.

The fix is to impose regularity conditions on the process, or the filtration.

Proposition 3.2. *Let A be closed and let X be a continuous, adapted process. Then*

$$T_A = \inf\{t \geq 0 \mid X_t \in A\}$$

is a stopping time.

Proof: It suffices to show that

$$\{T_A \leq t\} = \left\{ \inf_{\substack{s \in \mathbb{Q} \\ s \leq t}} d(X_s, A) = 0 \right\},$$

where $d(x, A)$ is the distance from x to A . Then the latter element is \mathcal{F}_t -measurable.

Suppose that $T_A = s \leq t$. Then there exists (s_n) such that $X_{s_n} \in A$ for all n , and $s_n \downarrow s$ as $n \rightarrow \infty$. Since X is continuous, $X_{s_n} \rightarrow X_s$, and since A is closed, $X_s = X_{T_A} \in A$.

Then we can find $q_n \in \mathbb{Q}$ such that $q_n \uparrow T_A$. Since $d(X_{T_A}, A) = 0$, $d(X_{q_n}, A) \rightarrow 0$ as $n \rightarrow \infty$. This shows one inclusion.

Now suppose that $\inf d(X_s, A) = 0$. Then there exists $s_n \in \mathbb{Q}$, with $s_n \leq t$ such that $d(X_{s_n}, A) \rightarrow 0$. By compactness, we can restrict to $s_n \rightarrow s$, and by continuity, $X_{s_n} \rightarrow X_s$, where $d(X_s, A) = 0$.

Since A is closed, this means $X_s \in A$, so $T_A \leq t$.

Definition 3.1. Let (\mathcal{F}_t) be a filtration. Set

$$\mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s.$$

If $\mathcal{F}_{t+} = \mathcal{F}_t$ for all t , then (\mathcal{F}_t) is *right-continuous*,

Proposition 3.3. *Let A be an open set and X a continuous process. Then*

$$T_A = \inf\{t \geq 0 \mid X_t \in A\}$$

is a stopping time with respect to (\mathcal{F}_{t+}) .

Proof: Note that

$$\{T_a < t\} = \bigcup_{\substack{q \in \mathbb{Q} \\ q < t}} \{X_q \in A\} \in \mathcal{F}_t,$$

by the continuity of X and the fact A is open, and this being a countable union. Hence

$$\{T_a \leq t\} = \bigcap_n \{T_a < t + 1/n\} \in \mathcal{F}_{t+}.$$

3.1 Martingale Regularization Theorem

A stochastic process is a random variable with values $\{f : \mathbb{R}_+ \rightarrow E\}$ with the σ -algebra which makes the evaluation maps $f \mapsto f(t)$ measurable for all $t \geq 0$.

The *law* of a stochastic process X is the measure $\mu(A) = \mathbb{P}(X \in A)$, where A is an event in this σ -algebra. In practice, this is not very easy to work with.

Suppose that μ is a probability measure on $D(\mathbb{R}_+, E)$. For each $J \subseteq \mathbb{R}_+$, let μ_J be the law $(X_t \mid t \in J)$. Then the family of probability measures (μ_J) are the finite dimensional distributions of μ . Note that

$$\left\{ \bigcap_{s \in J} \{X_s \in \mathcal{A}_s\} \mid J \text{ finite, } \mathcal{A}_s \in \mathcal{B}(\mathbb{R}) \right\}$$

is a π -system, which generates the product σ -algebra on $D(\mathbb{R}_+, E)$. Hence Cad lag processes are define by their finite dimensional distributions.

Definition 3.2. Let X, X' be processes on the same probability space. Then X' is a *version* of X if $\mathbb{P}(X_t = X'_t) = 1$ for all $t \geq 0$.

Example 3.2.

Let $X = (X_t)_{t \in [0,1]}$ be the process $X_t \equiv 0$. Let $U \sim U[0, 1]$, and $X'_t = \mathbb{1}(U = t)$.

Then the finite dimensional distributions of X, X' are the same and X' is a version of X . But X' is not continuous, and

$$\mathbb{P}(X'_t = 0 = X_t \text{ for all } t \in [0, 1]) = 0.$$

In practice, it is useful to have a cadlag version of a process.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a filtered probability space. Let $\mathcal{N} \subseteq \mathcal{F}$ be the collection of events with probability 0, and define

$$\tilde{\mathcal{F}}_t = \sigma(\mathcal{F}_{t+}, \mathcal{N}).$$

Definition 3.3. If $\mathcal{F}_t = \tilde{\mathcal{F}}_t$ for all $t \geq 0$, we say (\mathcal{F}_t) satisfies the *usual conditions*.

A continuous times process X is a *martingale* if

$$\mathbb{E}[X_t | \mathcal{F}_s] = X_s \quad \text{a.s.}$$

for all $0 \leq s \leq t$, and a submartingale or a supermartingale if the usual definitions hold.

Theorem 3.1 (Martingale Regularization). *Let $X = (X_t)_{t \geq 0}$ be a martingale with respect to (\mathcal{F}_t) .*

Then there exists a cadlag process \tilde{X} which is a martingale with respect to $(\tilde{\mathcal{F}}_t)$, and satisfies

$$X_t = \mathbb{E}[\tilde{X}_t | \mathcal{F}_t] \quad \text{a.s.}$$

for all $t \geq 0$. If (\mathcal{F}_t) satisfies the usual conditions, then \tilde{X} is a cadlag version of X .

This is not unexpected; we saw that martingales tended to converge in the discrete time sense, both forwards and backwards. We can use this idea in the continuous setting.

Lemma 3.1. *Let $f : \mathbb{Q}_+ \rightarrow \mathbb{R}$ be a function. Suppose that for all $a < b$, $a, b \in \mathbb{Q}$ and $I \subseteq \mathbb{Q}_+$ bounded, then f is bounded on I , and define*

$$\begin{aligned} N([a, b], I, f) &= \text{number of upcrossings that } f \text{ makes across } [a, b] \text{ in } I \\ &= \sup\{n \geq 0 \mid 0 \leq s_1 < t_2 < \cdots < s_n < t_n, s_i, t_i \in I, f(s_i) < a, f(t_i) > b\}. \end{aligned}$$

Suppose that $N([a, b], I, f) < \infty$ for all a, b and I . Then for all $t \in \mathbb{R}_+$,

$$\lim_{s \uparrow t | s \in \mathbb{Q}} f(s), \quad \lim_{s \downarrow t | s \in \mathbb{Q}} f(s)$$

exist and are finite.

Proof: Suppose that (s_n) is a sequence in \mathbb{Q}_+ which decreases to t . Then $\lim_n f(s_n)$ exists, since the number of upcrossings is finite.

This limit does not depend on (s_n) , since if (s'_n) is another sequence, we can

consider $(s_1, s'_1, s_2, s'_2, \dots)$.

We can use the same argument for (s_n) in \mathbb{Q}_+ which increases to t .

Now we are ready to prove martingale regularization.

Proof: Our goal is to set

$$\tilde{X}_t = \lim_{\substack{s \downarrow t \\ s \in \mathbb{Q}_+}} X_s$$

on a set of probability 1, and show that the conclusions hold for \tilde{X}_t .

The first step is to show that X is bounded on bounded subsets of \mathbb{Q}_+ . Fix $I \subseteq \mathbb{Q}_+$ bounded, and let $J = \{j_1, \dots, j_n\} \subseteq I$, with j_i in increasing order.

Then (X_{j_i}) is a discrete time martingale. From Doob's maximal inequality,

$$\lambda \mathbb{P}(\max_{j \in J} |X_{j_i}| > \lambda) \leq \mathbb{E}[|X_{j_n}|] \leq \mathbb{E}[|X_k|],$$

where $k = \sup I < \infty$. Take a monotone limit over $J \subseteq I$ finite, which increases to I , to get

$$\lambda \mathbb{P}(\sup_{t \in I} |X_t| > \lambda) \leq \mathbb{E}[|X_k|] < \infty.$$

Here we use the fact that I is a subset of \mathbb{Q}_+ , which is countable. Hence $\mathbb{P}(\sup_{t \in I} |X_t| < \infty) = 1$.

For step 2, fix $a < b$. We want to show that

$$N([a, b], I, X) = \sup_{J \subseteq I \text{ finite}} N([a, b], J, X) < \infty.$$

Fix $J = \{a_1, \dots, a_n\}$ in increasing order. Then (X_{a_i}) is a martingale in discrete time, so we can apply Doob's upcrossing inequality, to get

$$(b - a) \mathbb{E}[N([a, b], J, X)] \leq \mathbb{E}[(X_{a_n} - a)^-] \leq \mathbb{E}[(X_k - a)^-].$$

From monotone convergence theorem, let $I_M = \mathbb{Q}_+ \cap [0, M]$, then

$$N([a, b], I_M, X) < \infty.$$

Now we start defining \tilde{X} . First of all we want to define the null set for which we do not care about \tilde{X} . Define

$$\Omega_0 = \bigcap_{\substack{a < b \\ a, b \in \mathbb{Q} \\ M \in \mathbb{N}}} (\{N([a, b], I_M, X) < \infty\} \cap \{\sup_{t \in I_M} |X_t| < \infty\}).$$

Then from the above, this is a full set, i.e. $\mathbb{P}(\Omega_0) = 1$.

From the lemma,

$$X_{t+} = \lim_{\substack{s \downarrow t \\ s \in \mathbb{Q}}} X_s, \quad X_{t-} = \lim_{\substack{s \uparrow t \\ s \in \mathbb{Q}}} X_s$$

exist on Ω_0 . Set

$$\tilde{X}_t = \begin{cases} X_{t+} & \text{on } \Omega_0, \\ 0 & \text{on } \Omega \setminus \Omega_0. \end{cases}$$

Then (\tilde{X}_t) is $(\tilde{\mathcal{F}}_t)$ -adapted, since $(\tilde{\mathcal{F}}_t)$ contains all \mathbb{P} -null sets for all $t \geq 0$.

Our third step is to show that

$$X_t = \mathbb{E}[\tilde{X}_t | \mathcal{F}_t] \quad \text{a.s.}$$

for all $t \geq 0$. Let (t_n) be a sequence in \mathbb{Q} with $t_n \downarrow t$ as $n \rightarrow \infty$. So

$$\tilde{X}_t = \lim_{n \rightarrow \infty} X_{t_n}.$$

The process (X_{t_n}) is a backwards martingale, hence converges almost surely and in \mathcal{L}^1 , so

$$\mathbb{E}[X_{t_n} | \mathcal{F}_t] \rightarrow \mathbb{E}[\tilde{X}_t | \mathcal{F}_t]$$

in \mathcal{L}^1 . But since $\mathbb{E}[X_{t_n} | \mathcal{F}_t] = X_t$ almost surely for all n , by the martingale property of X , we get

$$X_t = \mathbb{E}[\tilde{X}_t | \mathcal{F}_t] \quad \text{a.s.}$$

Next (step four) we show that (\tilde{X}_t) is an $(\tilde{\mathcal{F}}_t)$ martingale. Fix $s < t$, and let (s_n) be a sequence in \mathbb{Q} with $s_n \downarrow s$ and $s_0 < t$. Then,

$$\tilde{X}_s = \lim_{n \rightarrow \infty} X_{s_n} = \lim_{n \rightarrow \infty} \mathbb{E}[X_t | \mathcal{F}_{s_n}].$$

Since $(\mathbb{E}[X_t | \mathcal{F}_{s_n}])$ is a backwards martingale, it converges almost surely and in \mathcal{L}^1 , to $\mathbb{E}[X_t | \mathcal{F}_{s+}]$. Hence

$$\tilde{X}_s = \mathbb{E}[X_t | \mathcal{F}_{s+}].$$

Two things to fix: first change X_t to \tilde{X}_t , and \mathcal{F}_{s+} to $\tilde{\mathcal{F}}_s$. For the first, note for any $T > t > s$,

$$\begin{aligned}\tilde{X}_s &= \mathbb{E}[X_T | \mathcal{F}_{s+}] = \mathbb{E}[\mathbb{E}[X_T | \mathcal{F}_{t+}] | \mathcal{F}_{s+}] \\ &= \mathbb{E}[\tilde{X}_t | \mathcal{F}_{s+}].\end{aligned}$$

Note that if \mathcal{G} is a σ -algebra, and $X \in \mathcal{L}^1$, then

$$\mathbb{E}[X | \sigma(\mathcal{G}, \mathcal{N})] = \mathbb{E}[X | \mathcal{G}] \quad \text{a.s.}$$

Applying this gives

$$\tilde{X}_s = \mathbb{E}[\tilde{X}_t | \tilde{\mathcal{F}}_s] \quad \text{a.s.}$$

Finally, for step five we show that \tilde{X}_t has the cadlag property.

Suppose that $\omega \in \Omega_0$ is such that $\tilde{X}(\omega)$ is not right-continuous. Then there exists t and (s_n) with $s_n \downarrow t$ as $n \rightarrow \infty$ so that

$$|\tilde{X}_{s_n}(\omega) - \tilde{X}_t(\omega)| > \varepsilon$$

for some $\varepsilon > 0$, and all n . By the definition of \tilde{X} , there exists $(s'_n) \in \mathbb{Q}$ with $s'_n > s_n$ and $s'_n \downarrow t$, with

$$|\tilde{X}_{s_n}(\omega) - X_{s'_n}(\omega)| \leq \varepsilon/2,$$

which gives

$$|X_{s'_n}(\omega) - \tilde{X}_t(\omega)| \geq \varepsilon/2.$$

This contradicts $X_{s'_n}(\omega) = \tilde{X}_t(\omega)$, so \tilde{X} is right-continuous on Ω_0 .

To show \tilde{X} has left limits is an exercise.

Example 3.3.

Let ξ, η be independent random variables which are ± 1 with probability $1/2$. Set

$$X_t = \begin{cases} 0 & t < 1, \\ \xi & t = 1, \\ \xi + \eta & t > 1. \end{cases}$$

Let $\mathcal{F}_t = \sigma(X_s \mid s \leq t)$. We can see that (X_t) is a (\mathcal{F}_t) -martingale, but this is not right-continuous at $t = 1$.

It is also easy to see that $\mathcal{F}_1 = \sigma(\xi)$, but $\mathcal{F}_{1+} = \sigma(\xi, \eta)$. To make a cadlag process, we set

$$\tilde{X}_t = \begin{cases} 0 & t < 1, \\ \xi + \eta & t \geq 1. \end{cases}$$

Then $X_t = \mathbb{E}[\tilde{X}_t | \mathcal{F}_t]$ almost-surely for all $t \geq 0$. Moreover \tilde{X} is a martingale with respect to (\mathcal{F}_{t+}) , and is cadlag.

But \tilde{X} is not a version of X , since $X_1 \neq \tilde{X}_1$ almost-surely.

From now on, we always consider the cadlag version of a martingale when the filtration satisfies the usual conditions.

3.2 Convergence and Doob's Inequalities

Theorem 3.2 (Martingale Convergence). *Let (X_t) be a \mathcal{L}^1 bounded cadlag martingale. Then $X_t \rightarrow X_\infty$ almost-surely for $X_\infty \in \mathcal{L}^1(\mathcal{F}_\infty)$, where $\mathcal{F}_\infty = \sigma(\mathcal{F}_t \mid t \geq 0)$.*

Proof: Let $N([a, b], I_m, X)$ be the number of upcrossings across $[a, b]$ in the interval $I_m = [0, m] \cap \mathbb{Q}$.

As in martingale regularization,

$$(b - a)\mathbb{E}[N([a, b], I_m, X)] \leq |a| + \sup_{t \geq 0} \mathbb{E}[|X_t|] < \infty,$$

since X is \mathcal{L}^1 -bounded. From MCT,

$$(b - a)\mathbb{E}[N([a, b], \mathbb{Q}_+, X)] \leq |a| + \sup_{t \geq 0} \mathbb{E}[|X_t|] < \infty,$$

hence $N([a, b], \mathbb{Q}, X) < \infty$ almost-surely. Again let

$$\Omega_0 = \bigcap_{\substack{a < b \\ a, b \in \mathbb{Q}}} \{N([a, b], \mathbb{Q}_+, X) < \infty\}.$$

Then $\mathbb{P}(\Omega_0) = 1$. On Ω_0 , $X_q \rightarrow X_\infty$ as $q \rightarrow \infty$, for $q \in \mathbb{Q}_+$. We now use the cadlag property to show that $X_t \rightarrow X_\infty$ as $t \rightarrow \infty$. Since $X_q \rightarrow X_\infty$, for all $\varepsilon > 0$ there is q_0 with

$$|X_q - X_\infty| \leq \varepsilon/2,$$

for all $q \geq q_0$. From right-continuity, if $t > q_0$, then there is $q \in \mathbb{Q}_+$ with, for all $q > t$, $|X_t - X_q| < \varepsilon/2$. Combining these inequalities gives

$$|X_t - X_\infty| < \varepsilon$$

for all $t > q_0$, showing $X_t \rightarrow X_\infty$.

Theorem 3.3 (Doob's Maximal Inequality). *Let (X_t) be a cadlag martingale, and define*

$$X_t^* = \sup_{0 \leq s \leq t} |X_s|.$$

Then for all $\lambda \geq 0$ and $t \geq 0$,

$$\lambda \mathbb{P}(X_t^* \geq \lambda) \leq \mathbb{E}[|X_t| \mathbb{1}(X_t^* \geq \lambda)] \leq \mathbb{E}[|X_t|].$$

Proof: From the cadlag property,

$$\sup_{s \leq t} |X_s| = \sup_{s \in \{t\} \cup ([0, t) \cap \mathbb{Q})} |X_s|.$$

Then we can argue as before.

We will list the following statements which we have proven for discrete time martingales and can be easily extended to continuous time.

Theorem 3.4 (Doob's \mathcal{L}^p Inequality). *Let (X_t) be a cadlag martingale, and $p > 1$. Let $X_t^* = \sup_{s \leq t} |X_s|$. Then*

$$\|X_t^*\|_p \leq \frac{p}{p-1} \|X_t\|_p.$$

Theorem 3.5. *Let (X_t) be a cadlag martingale, and $p > 1$. Then the following are equivalent:*

- (i) X is \mathcal{L}^p bounded.
- (ii) X converges almost-surely and in \mathcal{L}^1 to X_∞ .
- (iii) There exists $Z \in \mathcal{L}^p$ with $X_t = \mathbb{E}[Z | \mathcal{F}_t]$ for all $t \geq 0$.

Theorem 3.6. *Let (X_t) be a cadlag martingale. Then the following are equivalent:*

- (i) X is uniformly integrable.
- (ii) X converges almost-surely and in \mathcal{L}^1 to X_∞ .
- (iii) There exists $Z \in \mathcal{L}^1$ with $X_t = \mathbb{E}[Z | \mathcal{F}_t]$ almost-surely, for all $t \geq 0$.

Theorem 3.7 (Optional Stopping Theorem). *Let X be a cadlag, UI martingale. Then for all $S \leq T$ stopping times,*

$$\mathbb{E}[X_T | \mathcal{F}_S] = X_S \quad a.s.$$

Proof: Let $A \in \mathcal{F}_S$, then we need to show that

$$\mathbb{E}[X_T \mathbb{1}_A] = \mathbb{E}[X_S \mathbb{1}_A].$$

Let $T_n = 2^{-n} \lceil 2^n T \rceil$, and $S_n = 2^{-n} \lceil 2^n S \rceil$. Then $S_n \downarrow S$, $T_n \downarrow T$ as $n \rightarrow \infty$. By right continuity, $X_{S_n} \rightarrow X_S$, $X_{T_n} \rightarrow X_T$ as $n \rightarrow \infty$.

From optional stopping theorem in the discrete case, we have $X_{T_n} = \mathbb{E}[X_\infty | \mathcal{F}_{T_n}]$, hence X_{T_n} is UI (as a conditional expectation of the same variable), so convergence $X_{T_n} \rightarrow X_T$ happens in \mathcal{L}^1 . Similarly $X_{S_n} \rightarrow X_S$ in \mathcal{L}^1 .

From optional stopping theorem in the discrete sense, $X_{S_n} = \mathbb{E}[X_{T_n} | \mathcal{F}_{S_n}]$. Since $A \in \mathcal{F}_S \subseteq \mathcal{F}_{S_n}$ as $S_n \geq S$, we have

$$\mathbb{E}[X_{T_n} \mathbb{1}_A] = \mathbb{E}[X_{S_n} \mathbb{1}_A].$$

Take $n \rightarrow \infty$. From \mathcal{L}^1 convergence,

$$\mathbb{E}[X_T \mathbb{1}_A] = \mathbb{E}[X_S \mathbb{1}_A].$$

3.3 Kolmogorov Continuity Criterion

Let $\mathcal{D}_n = \{k2^{-n} \mid 0 \leq k \leq 2^n\}$, and $\mathcal{D} = \bigcup \mathcal{D}_n$.

Theorem 3.8. *Let $(X_t)_{t \in \mathcal{D}}$ be a \mathbb{R} -valued process. Suppose that there exists $p, s > 0$ so that*

$$\mathbb{E}[|X_s - X_t|^p] \leq C|t - s|^{1+\varepsilon},$$

for all $s, t \in \mathcal{D}$ for some $C < \infty$. Then for $\alpha \in (0, \varepsilon/p)$, the process $(X_t)_{t \in \mathcal{D}}$ is α -Hölder continuous, i.e. there exists $K_\alpha < \infty$ such that

$$|X_t - X_s| \leq K_\alpha |t - s|^\alpha,$$

for all $s, t \in \mathcal{D}$.

Proof: From Markov's inequality,

$$\mathbb{P}(|X_{k2^{-n}} - X_{(k+1)2^{-n}}| \geq 2^{-n\alpha}) \leq C2^{n\alpha p} \cdot 2^{-n(1+\varepsilon)},$$

from the given condition. The union bound gives

$$\mathbb{P}\left(\max_{1 \leq k \leq 2^n} |X_{k2^{-n}} - X_{(k+1)2^{-n}}| \geq 2^{-n\alpha}\right) \leq C2^{-n(\varepsilon - \alpha p)}.$$

Since $\alpha \in (0, \varepsilon/p)$, from Borel-Cantelli, since the RHS is summable,

$$\max_{0 \leq k \leq 2^n} |X_{k2^{-n}} - X_{(k+1)2^{-n}}| \leq 2^{-n\alpha},$$

for all n sufficiently large. Hence there exists M such that

$$\sup_{n \geq 0} \max_{0 \leq k \leq 2^n} \frac{|X_{k2^{-n}} - X_{(k+1)2^{-n}}|}{2^{-n\alpha}} \leq M < \infty.$$

Now we need to transform this M into M' such that

$$|X_t - X_s| \leq M'|t - s|^\alpha,$$

for all $t, s \in \mathcal{D}$. Fix $s, t \in \mathcal{D}$, and fix $r \in \mathbb{Z}$ such that $2^{-r-1} < t - s \leq 2^{-r}$. So there exists k such that

$$s < k2^{-(r+1)} < t.$$

Let β be this dividing dyadic. We know that $0 < t - \beta < 2^{-r}$, so write

$$t - \beta = \sum_{j=r+1}^{\infty} \frac{x_j}{2^j}$$

for $x_j \in \{0, 1\}$. Similarly we can write

$$\beta - s = \sum_{j=r+1}^{\infty} \frac{y_j}{2^j},$$

for $y_j \in \{0, 1\}$. This partitions $[s, t)$ into a disjoint union of dyadic intervals of length 2^{-n} where $n \geq r + 1$. Importantly, there are at most 2 intervals of each length. Hence

$$|X_t - X_s| \leq 2 \sum_{n \geq r+1} M2^{-n\alpha} = 2M \frac{2^{-(r+1)\alpha}}{1 - 2^{-\alpha}}.$$

Set $M' = 2M/(1 - 2^{-\alpha})$. Then we get

$$|X_t - X_s| \leq M'2^{-(r+1)\alpha} \leq M'|t - s|^\alpha.$$

4 Weak Convergence

Let (M, d) be a metric space equipped with the Borel σ -algebra.

Definition 4.1. Let (μ_n) be a sequence of probability measures on M . We say that $\mu_n \rightarrow \mu$ *weakly* if $\mu_n(f) \rightarrow \mu(f)$ for all f bounded, continuous.

Example 4.1.

1. Let (x_n) be the sequence in M with $x_n \rightarrow x$ as $n \rightarrow \infty$. Then $\delta_{x_n} \rightarrow \delta_x$ as $n \rightarrow \infty$. Indeed,

$$\delta_{x_n}(f) = f(x_n) \rightarrow f(x) = \delta_x(f)$$

for all f continuous.

2. Let $M = [0, 1]$ with the Euclidean metric, and let

$$\mu_n = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{k/n}.$$

Then we have

$$\mu_n(f) = \frac{1}{n} \sum_{k=0}^{n-1} f(k/n).$$

This is a Riemann sum, which converges to

$$\int_0^1 f(x) dx$$

for f continuous. Hence $\mu_n \Rightarrow \mu$, the Lebesgue measure on $[0, 1]$.

Remark. If A is a Borel set, it is not true that $\mu_n \Rightarrow \mu$ implies that $\mu_n(A) = \mu(A)$ as $n \rightarrow \infty$.

For example, let $x_n = \frac{1}{n}$, and $\mu_n = \delta_{x_n}$, then $\mu_n \Rightarrow \delta_0$.

Take $A = (0, 1)$, then $\mu_n(A) = 1$ for all $n \geq 2$, but $\delta_0(A) = 0$.

Theorem 4.1. Suppose that μ_n is a sequence of probability measures. The following are equivalent:

- (i) $\mu_n \Rightarrow \mu$.
- (ii) $\liminf \mu_n(G) \geq \mu(G)$ for all G open.

- (iii) $\limsup \mu_n(A) \leq \mu(A)$ for all A closed.
 (iv) $\lim \mu_n(A) = \mu(A)$ for all A with $\mu(\partial A) = 0$.

Proof:

(i) \implies (ii): Let G be open with $G^c \neq \emptyset$. For each $M > 0$, set

$$f_M(x) = 1 \wedge (Md(x, G^c)).$$

Then f_M is bounded, continuous and $f_M \uparrow \mathbb{1}_G$ as $M \rightarrow \infty$. Maybe we can use Tietze in the non-metric case. Hence

$$\mu_n(G) \geq \mu_n(f_M) \rightarrow \mu(f_M)$$

as $n \rightarrow \infty$. Then

$$\liminf_{n \rightarrow \infty} \mu_n(G) \geq \liminf_{n \rightarrow \infty} \mu_n(f_M) = \mu(f_M).$$

Now take $f_M \uparrow \mathbb{1}_G$ and MCT to find

$$\liminf_{n \rightarrow \infty} \mu_n(G) \geq \mu(G).$$

(ii) \implies (iii): This is easy to see as the complement of an open set is closed. Similarly (iii) \implies (ii).

Index

- π -system, 10
- adapted, 14, 34
- backwards martingale, 28
- cadlag, 34
- closed martingale, 24
- conditional density function, 13
- conditional distribution, 13
- conditional expectation, 6
- filtration, 14, 34
- independent, 7
- integrable process, 14
- law, 38
- martingale, 14, 39
- natural filtration, 14
- orthogonal projection, 5
- Radon-Nikodym derivative, 32
- right-continuous filtration, 37
- stochastic process, 14, 34
- stopped σ -algebra, 16
- stopped process, 16
- stopping time, 15
- submartingale, 14
- supermartingale, 14
- uniformly integrable, 25
- uniformly integrable martingale, 26
- usual conditions, 39
- version, 38
- weak convergence, 47