Part III Advanced Probability

Based on lectures by P. Sousi

Notes taken by Pantelis Tassopoulos

Michaelmas 2023

Contents

1 Conditional Expectation

Lecture 1

1.1 Basic definitions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Remember the following definitions

Definition 1.1 (Sigma algebra). $\mathcal{F}$ is a sigma algebra if and only if: $(\mathcal{F} \in \mathcal{P}\Omega)$

1. $\Omega \in \mathcal{F}$
2. $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
3. $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{F} \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$

Definition 1.2 (Probability measure). $\mathbb{P}$ is a probability measure if

1. $\mathbb{P} : \mathcal{F} \to [0,1]$ (i.e. a set function)
2. $\mathbb{P}(\Omega) = 1$, and $\mathbb{P}(\varnothing) = 0$
3. $(A_n)_{n\in\mathbb{N}}$ pairwise disjoint $\implies \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$.

Definition 1.3 (Random Variable). $X : \Omega \to \mathbb{R}$ is a random variable if for all B open in $\mathbb{R}$, $X^{-1}(B) \in \mathcal{F}$.

Remark. Observe that the sigma algebra $\mathcal{G} = \{B \subseteq \mathbb{R} : X(B) \in \mathcal{F}\} \supseteq \mathcal{O} \implies \mathcal{G} \supseteq \mathcal{B}(\mathbb{R})$, the former being the collection of open sets in $\mathbb{R}$ and the latter the Borel sigma algebra on $\mathbb{R}$ with the usual topology, namely, $\sigma(\mathcal{O})$ (see below for the notation).

Let $\mathcal{A}$ be a collection of subsets of Ω. We define

$$\begin{aligned} \sigma(\mathcal{A}) &= \text{smallest sigma algebra containing } \mathcal{A} \\ &= \bigcap \{\mathcal{T} : \mathcal{T} \text{ sigma algebra containing } \mathcal{A}\}. \end{aligned}$$

Definition 1.4 (Borel sigma algebra on $\mathbb{R}$). Let $\mathcal{O} = \{\text{open sets}\mathbb{R}\}$. Then, the Borel sigma algebra $\mathcal{B}(\mathbb{R}) (:= \mathcal{B})$ is defined as above, namely,

$$\mathcal{B}(\mathbb{R}) := \sigma(\mathcal{O}).$$

Let $(X_i)_{i\in I}$ be a family of random variables, then $\sigma(X_i : i \in I)$ = the smallest sigma algebra that makes them all measurable. We also have the characterisation $\sigma(X_i : i \in I) = \sigma(\{\underbrace{\{\omega \in \Omega : X_i(\omega) \in B\}}_{X_i^{-1}(B)}, i \in I, B \in \mathcal{B}(\mathbb{R})\})$.

1.2 Expectation

Note we use the following for the indicator function on some event A

$$\mathbf{1}(A)(x) = \mathbf{1}(x \in A) := \left.\begin{matrix} 1, & x \in A \\ 0, & x \notin A \end{matrix}\right\}, \quad A \in \mathcal{F}.$$

We now begin the construction of the expectation of generic random variables.

<u>Positive simple random variables:</u> $X = \sum_{i=1}^{n} \mathbf{1}(A_i), c_i \geqslant 0, A_i \in \mathcal{F}.$.

$$\mathbb{E}[X] := \sum_{i=1}^{n} c_i \mathbb{P}(A_i).$$

Non-negative random variables: ($X \geqslant 0$). We proceed by approximation. Namely, let $X_n(\omega) := 2^{-n}\lfloor 2^{-n} \cdot X(\omega)\rfloor \wedge n \uparrow X(\omega), n \to \infty$. Now, by monotone convergence,

$$\mathbb{E}[X] :=\uparrow \lim_{n\to\infty} \mathbb{E}[X_n] = \sup \mathbb{E}[X].$$

General random variables: Have the decomposition $X = X^+ - X^-$, where $X^+ = X \vee 0$, $X^- = -X \wedge 0$. If one of $\mathbb{E}[X^+], \mathbb{E}[X^-] < \infty$ then set

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

Definition 1.5. X is called integrable if $\mathbb{E}[|X|] < \infty$.

Definition 1.6. Let $B \in \mathcal{F}$ with $\mathbb{P}(B) > 0$. Then for all $A \in \mathcal{F}$, set

$$\mathbb{P}(A|B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Now for an integer-valued random variable X, we set:

$$\mathbb{E}[X|B] := \frac{\mathbb{E}[X \cdot \mathbf{1}_B]}{\mathbb{P}(B)}$$

1.3 Conditional expectation with respect to countably generated sigma algebras

Lecture 2 We now extend the definition of the conditional expectation for a countably generated sigma algebra. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We call the sigma algebra $\mathcal{G}$ countably generated if there exists a collection $(B_n)_{n\in\mathbb{N}}$ of pairwise disjoint events such that $\bigcup_{n\in I} B_n = \Omega$ with (I countable) and $\mathcal{G} = \sigma(B_i : i \in I)$.

Let X be an integrable random variable. We want to define $\mathbb{E}[X|\mathcal{G}]$.

Define $X'(\omega) = \mathbb{E}[X|B_i]$, whenever $w \in B_i$, i.e.

$$X' = \sum_{i\in I} \mathbf{1}(B_i) \cdot \mathbb{E}[X|B_i].$$

We make the convention that $\mathbb{E}[X|B_i] = 0$ if $\mathbb{P}(B_i) = 0$. It is easy to check that X' is $\mathcal{G}-$measurable. We also have that

$$\mathcal{G} = \left\{ \bigcup_{j\in} B_j : J \subseteq I \right\}$$

and X' satisfies for all $G \in \mathcal{G}$:$\mathbb{E}[X \cdot \mathbf{1}_G] = \mathbb{E}[X' \cdot \mathbf{1}_G]$ and

$$\begin{aligned}
\mathbb{E}[|X'|] &\leqslant \mathbb{E}\left[\sum_{i\in I} |\mathbb{E}[X|B_i]\mathbf{1}(B_i)\right] \\
&= \sum_{i\in I} \mathbb{P}(B_i) \cdot |\mathbb{E}[X|B_i]| \\
&\leqslant \sum_{i\in I} \mathbb{P}(B_i) \cdot \underbrace{\mathbb{E}[X \cdot \mathbf{1}(B_i)]}_{\mathbb{P}(B_i)} \\
&= \mathbb{E}[|X|] < \infty.
\end{aligned}$$

1.4 General case

We say $A \in \mathcal{F}$ happens a.s. if $\mathbb{P}(A) = 1$. Recall (from measure theory and basic functional analysis):

Theorem 1.1 (Monotone Convergence Theorem (MCT)). Let $(X_n)_{n\in\mathbb{N}}$ be such that $X_n \geqslant 0, X$ be random variables such that $X_n \uparrow X$ as $n \to \infty$. Then, $\mathbb{E}[X_n] \uparrow \mathbb{E}[X]$ as $n \to \infty$.

Theorem 1.2 (Dominanted Convergenec Theorem (DCT)). Let $(X_n)_{n\in\mathbb{N}}$ be random variables such that $X_n \to X$ a.s. as $n \to \infty$ and $|X_n| \leqslant Y$ a.s. for all $n \in \mathbb{N}$, where Y is integrable, then $\mathbb{E}[X_n] \to \mathbb{E}[X]$, as $n \to \infty$.

Let $1 \leqslant p < \infty$ and f a measurable function, then set $\|f\|_p := (\mathbb{E}[\|f\|^p])^{\frac{1}{p}}$. If $p = \infty$, then set $\|f\|_\infty := \inf\{\lambda : |f| \leqslant \lambda \text{ a.s.}\}$. Recall for all p, the Lebesgue spaces, $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) = \{f : \|f\|_p < \infty\}$.

Theorem 1.3. $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ is a Hilbert space, with inner product $\langle u, v\rangle_2 = \mathbb{E}[u \cdot v]$. Furthermore, for any closed subspace $\mathcal{H}$, if $f \in \mathcal{L}^2$, there exists a unique $g \in \mathcal{H}$ s.t. $\|f - g\|_{\mathcal{L}^2} = \inf_{h\in\mathcal{H}} \|f - h\|_{\mathcal{L}^2}$ and $\langle f - g, h\rangle = 0$, for all $h \in \mathcal{H}$. We say that g is the <u>orthogonal projection</u> of f in $\mathcal{H}$.

We now construct the conditional expectation in the general case, for any integrably random variable with respect to an arbitrary sigma algebras.

Theorem 1.4 (Conditional Expectation). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G} \subseteq \mathcal{F}$ a sub-sigma algebra, $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Then there exists an integrable random variable Y satisfying:

1. Y is $\mathcal{G}-$measurable
2. for all $G \in \mathcal{G}, \mathbb{E}[X \cdot \mathbf{1}(G)] = \mathbb{E}[Y \cdot \mathbf{1}(G)]$.

Moreover, Y unique in the sense that if Y' also satisfies the above $1), 2)$, then $Y = Y'$ a.s.. We call Y a version of the conditional expectation of X given G. We write $Y = \mathbb{E}[X\mathcal{G}]$ a.s. If $\mathcal{G} = \sigma(Z)$, where Z is a random variable, then we write $\mathbb{E}[Z] = \mathbb{E}[X|\mathcal{G}]$.

Remark. 2) could be replaced by $\mathbb{E}[X \cdot Z] = \mathbb{E}[Y \cdot Z]$ for all Z bounded $\mathcal{G}-$measurable random variables.

We now state and prove the main theorem of this section:

Proof. (Theorem 1.4) <u>Uniqueness:</u> Let Y, Y' satisfy $1), 2)$. Let $A = \{Y > Y'\} \in \mathcal{G}$. Then 2)

$$\begin{aligned}
&\implies \mathbb{E}[Y \cdot \mathbf{1}(A)] = \mathbb{E}[Y' \cdot \mathbf{1}(A)] = \mathbb{E}[X \cdot \mathbf{1}(A)] \\
&\implies \mathbb{E}[(Y - Y') \cdot \mathbf{1}(A)] = 0 \\
&\implies \mathbb{P}(A) = \mathbb{P}(Y > Y') = 0 \\
&\implies Y \leqslant Y' \text{ a.s..}
\end{aligned}$$

We similarly obtain $Y \geqslant Y'$ a.s., hence we deduce that $Y = Y'$ a.s.

<u>Existence:</u> three steps.

1. Assume that $X \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Observe that $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ is a closed subspace of $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Hence, Theorem 1.3, we have the decomposition $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}) = \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P}) \oplus \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})^\perp$. Then, we have the corresponding decomposition $X = Y + Z$, where $Y \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ and $Z \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ respectively. Define $\mathbb{E}[X\mathcal{G}] := Y$, Y is $\mathcal{G}-$measurable and for all $A \in \mathcal{G}$, $\mathbb{E}[X \cdot \mathbf{1}(A)]\mathbb{E}[Y \cdot \mathbf{1}(A)] = \mathbb{E}[Z \cdot \mathbf{1}(A)]$ since $Z \in \mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})^\perp$.

   <u>Claim:</u> If $X \geqslant 0$, a.s. then $Y \geqslant 0$ a.s. Indeed, let $A = \{Y < 0\} \in \mathcal{G}$. Then observe that $0 \leqslant \mathbb{E}[X \cdot \mathbf{1}(A)] = \mathbb{E}[Y \cdot \mathbf{1}(A)] \leqslant 0$. Hence $\mathbb{E}[Y \cdot \mathbf{1}(A)] = 0$ and so $\mathbb{P}(A) = 0$, gibing $Y = 0$ a.s.

2. Assume $X \geqslant 0$.
   Define $X_n = X \wedge n \leqslant n$, meaning X_n is bounded for all $n \in \mathbb{N}$. So $X_n \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Let $Y_n = \mathbb{E}[X_n]$ a.s.. $(X_n)_{n\in\mathbb{N}}$ is an increasing sequence. By the claim above, so is $(Y_n)_{n\in\mathbb{N}}$ a.s. Define $Y = \limsup_n Y_n$ meaning Y is $\mathcal{G}-$measurable and $Y =\uparrow \lim_{n\to\infty} Y_n$ a.s. Now, we have that for all $A \in \mathcal{G}$, $\mathbb{E}[X_n \cdot \mathbf{1}(A)] = \mathbb{E}[Y_n \cdot \mathbf{1}(A)]$. Thus, by theorem 1.1 (MCT), $\mathbb{E}[X \cdot \mathbf{1}(A)] = \lim_{n\to\infty} \mathbb{E}[X_n \cdot \mathbf{1}(A)] = \lim_{n\to\infty} \mathbb{E}[Y_n \cdot \mathbf{1}(A)] = \mathbb{E}[Y \cdot \mathbf{1}(A)]$.

3. X general in $\mathcal{L}^1$.
   Decompose as before $X = X^+ - X^-$. Define, $\mathbb{E}[X\mathcal{G}] = \mathbb{E}[X^+|\mathcal{G}] - \mathbb{E}[X^-|\mathcal{G}]$.

$\square$

Lecture 3

Remark. From the second step of the proof of Theorem 1.4 we see that we can define $\mathbb{E}[X|\mathcal{G}]$ for all $X \geqslant 0$, not necessarily integrable. It satisfies all conditions $1), 2)$ except for the integrability one.

Definition 1.7. $\underbrace{\mathcal{G}_1, \mathcal{G}_2, \dots}_{\text{sigma algebras}} \subset \mathcal{F}$. We call them independent if whenever $G_i \in \mathcal{G}_i$ and $i_1 < \dots i_k$ for some $k \in \mathbb{N}$, then $\mathbb{P}(G_{i_1} \cap \dots \cap G_{i_k}) = \prod_{j=1}^{k} \mathbb{P}(G_{i_j})$.

Moreover, let X be a random variable and $\mathcal{G}$ a sigma algebra, then they are said to be int if $\sigma(X)$ is independent of $\mathcal{G}$.

Properties of conditional expectations: Fix $X, y \in \mathcal{L}^1$, $G \in \mathcal{F}$.

1. $\mathbb{E}[\mathbb{E}[X\mathcal{G}]] = \mathbb{E}[X]$ (take $A = \Omega$)
2. If X is $\mathcal{G}-$measurable, then $\mathbb{E}[X\mathcal{G}] = X$ a.s.
3. If X is independent of $\mathcal{G}$, then $\mathbb{E}[X\mathcal{G}] = \mathbb{E}[X]$
4. If $X \geqslant 0$ a.s., then $\mathbb{E}[X\mathcal{G}] \geqslant 0$ a.s.
5. For $\alpha, \beta \in \mathbb{R}$ $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha\mathbb{E}[X] + \beta\mathbb{E}[Y]$
6. $\mathbb{E}[X|\mathcal{G}]| \leqslant \mathbb{E}[|X||\mathcal{G}]$ a.s.

Below we proved:we expansions of useful measure theoretic results for the expectation to their corresponding conditional counterparts. First recall:

Lemma 1.1 (Fatou's Lemma). Let $X_n \geqslant 0$ for all $n \in \mathbb{N}$. Then

$$\mathbb{E}[\liminf_n X_n] \leqslant \liminf_n \mathbb{E}[X_n] \quad \text{a.s}$$

Theorem 1.5 (Jensen's Inequality). If X is integrable and $\phi : \mathbb{R} \to \mathbb{R}$ is a convex function, then

$$\phi(\mathbb{E}[X]) \leqslant \mathbb{E}[\phi(X)] \quad \text{a.s.}$$

Now the results themselves:

Theorem 1.6 (Conditional Monotone Convergence theorem (MCT)). Let $\mathcal{G} \subset \mathcal{F}$ be sigma algebras, $X_n \geqslant 0$ a.a. and $X_n \uparrow X$, as $n \to \infty$ a.s. Then

$$\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}] \quad \text{a.s.}$$

Proof. Theorem 1.6 Set $Y_n = \mathbb{E}[X_n\mathcal{G}]$ a.s. Observe that Y_n is a.s. increasing. Set $Y = \limsup_n Y_n$. Y_n is $\mathcal{G}-$measurable, hence, so is Y (as a lim sup of $\mathcal{G}-$measurable random variables) is also $\mathcal{G}-$measurable. Also, $Y = \lim_{n\to\infty} Y_n$ a.s.

Need to show: $\mathbb{E}[Y \cdot \mathbf{1}(A)]\mathbb{E}[X \cdot \mathbf{1}(A)]$ for all $A \in \mathcal{G}$. Indeed,

$$\begin{aligned}
\mathbb{E}[Y \cdot \mathbf{1}(A)] &= \mathbb{E}[\lim_{n\to\infty} Y_n \cdot \mathbf{1}(A)] \overset{\text{MCT}}{=} \lim_{n\to\infty} \mathbb{E}[Y_n \cdot \mathbf{1}(A)] \\
&= \lim_{n\to\infty} \mathbb{E}[X_n \cdot \mathbf{1}(A)] = \mathbb{E}[X \cdot \mathbf{1}(A)].
\end{aligned}$$

$\square$

Proof. Theorem 1.1 $\liminf_n X_n = \lim_{n\to\infty}\left(\inf_{k\geqslant n} X_k\right)$, the limit of an increasing sequence. By Theorem 1.1, we have

$$\lim_{n\to\infty} \mathbb{E}[\inf_{k\geqslant n} X_n|\mathcal{G}] = \mathbb{E}[\liminf_n X_n|\mathcal{G}]$$

and

$$\mathbb{E}[\inf_{k\geqslant n} X_k|\mathcal{G}] \overset{\text{a.s.}}{\leqslant} \inf_{k\geqslant n} \mathbb{E}[X_k|\mathcal{G}]^1$$

which gives the result

$$\mathbb{E}[\liminf_n X_n] \leqslant \liminf_n \mathbb{E}[X_n] \quad \text{a.s.}$$

□

Theorem 1.7 (Conditional Dominated Convergence Theorem). SUppose $X_n \to X$ a.s. $n \to \infty$ and $|X_n| \leqslant Y$ a.s. for all $n \in \mathbb{N}$ with Y integrable. Then $\mathbb{E}[X_n\mathcal{G}] \to \mathbb{E}[X\mathcal{G}]$ a.s. as $n \to \infty$.

Proof. From $-Y \leqslant X_n \leqslant Y$, we have $X_n + Y \geqslant 0$ for all $n \in \mathbb{N}$ and $Y - X_n \geqslant 0$a.s. By Theorem 1.1,

$$\begin{aligned}\mathbb{E}[X + Y\mathcal{G}] &= \mathbb{E}[\liminf_n (X_n + Y)|\mathcal{G}] \\ &\leqslant \liminf_n \mathbb{E}[X_n + Y|\mathcal{G}] = \liminf_n \mathbb{E}[X_n\mathcal{G}] + \mathbb{E}[X]\end{aligned}$$

Thus,

$$\begin{aligned}\mathbb{E}[|X - Y||\mathcal{G}] &= \mathbb{E}[Y - \liminf_n X_n|\mathcal{G}] \\ &\leqslant \mathbb{E}[Y] + \liminf_n \mathbb{E}[X_n|\mathcal{G}]\end{aligned}$$

Hence,

$$\limsup_n \mathbb{E}[X_n|\mathcal{G}] \leqslant \mathbb{E}[X|\mathcal{G}]$$

and

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

a.s., concluding the proof.

□

Theorem 1.8 (Conditional Jensen). Let $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function s.t. $\phi(X)$ is integrable or $\phi(X) \geqslant 0$

$$\phi(\mathbb{E}[X|\mathcal{G}]) \leqslant \mathbb{E}[\phi(X)|\mathcal{G}] \quad \text{a.s.}$$

Proof. Claim: (true for any convex function, no proof given) $\phi(x) = \sup_{i\in\mathbb{N}}(a_i x + b_i)$, $a_i b_i \in \mathbb{R}$. Thus,

$$\mathbb{E}[\phi(X)|\mathcal{G}] \geqslant a_i\mathbb{E}[X|\mathcal{G}] + b_i \quad \text{for all } i \in \mathbb{N}.$$

Taking the supremum gives [2]

$$\begin{aligned}\mathbb{E}[\phi(X)|\mathcal{G}] &\geqslant \sup_{i\in\mathbb{N}} (a_i\mathbb{E}[X|\mathcal{G}] + b_i) \\ &= \phi(\mathbb{E}[X|\mathcal{G}]) \quad \text{a.s.}\end{aligned}$$

□

Corollary 1.8.1. For all $1 \leqslant p < \infty$ $\|\mathbb{E}[X|\mathcal{G}]\|_p \leqslant \|X\|_p$.

Proof. Apply conditional Jensen. □

[1] can take the infinum due to countability that preserves a.s.

[2] can take the supremum due to countability which again preserves a.s.

Proposition 1.1 (Tower Property). Let X be integrable and $\mathcal{H} \subseteq \mathcal{G}$ sigma algebras. Then

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

Proof. (a) $\mathbb{E}[X|\mathcal{H}]$ is $\mathcal{H}-$measurable.

(b) For all $A \in \mathcal{H}$ NTS:

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot \mathbf{1}(A)] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}] \cdot \mathbf{1}(A)]$$

Indeed, both terms above are equal to $\mathbb{E}[X \cdot \mathbf{1}(A)]$ since $A \in \mathcal{G} \subseteq \mathcal{H}$.

□

Proposition 1.2. Let $X \in \mathcal{L}^1$, $\mathcal{G} \subseteq \mathcal{F}$, Y bounded $\mathcal{G}-$measurable. Then

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

Proof. (a) RHS is clearly $\mathcal{G}-$measurable.

(b) For all $A \in \mathcal{G}$:

$$\begin{aligned} \mathbb{E}[X \cdot Y \cdot \mathbf{1}(A)] &= \mathbb{E}[Y \cdot \mathbb{E}[X\mathcal{G}] \cdot \mathbf{1}(A)] \\ \mathbb{E}[X \cdot (\underbrace{Y \cdot \mathbf{1}(A)}_{\mathcal{G}\text{-meas. and bounded}})] &= \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot Y \cdot \mathbf{1}(A)] = RHS. \end{aligned}$$

(Also observe that by a monotone class argument, we have for any integrable function $f : \Omega \to \mathbb{R}$, $\mathbb{E}[X \cdot f] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot f]$) □

Lecture 4 We are building towards the Theorem

Theorem 1.9. $X \in \mathcal{L}^1, \mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$. Assume $\sigma(\mathcal{G}, \mathcal{H}) \perp \mathcal{H}$, Then

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

We begin with a definition

Definition 1.8. Let $\mathcal{A}$ be a collection of sts. It is called a $\pi-$system if for all $A, B \in \mathcal{A}$, we also have $A \cap B \in \mathcal{A}$.

Theorem 1.10 (Uniquenes of extension). Let (E, ξ)be a measurable space and let $\mathcal{A}$ be a $\pi-$system generating the sigma algebra ξ. Let μ, ν be two measures on (E, ξ) with $\mu(E) = \nu(E) < \infty$. If $\mu = \nu$ on $\mathcal{A}$, then $\mu = \nu$ on ξ.

Proof. (Theorem 1.9) NTS: for all $F \in \sigma(\mathcal{G}, \mathcal{H})$

$$\mathbb{E}[X \cdot \mathbf{1}_F] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot \mathbf{1}_F]$$

Now, set $\mathcal{A} = \{A \cap B : A \in \mathcal{G}, B \in \mathcal{H}\}$. It is easy to check that $\mathcal{A}$ is a $\pi-$system generating $\sigma(\mathcal{G}, \mathcal{H})$. If $F = A \cap B$ for some $A \in \mathcal{G}$ and $B \in \mathcal{H}$, Then

$$\begin{aligned} \mathbb{E}[X \cdot \mathbf{1}(A \cap B)] &= \mathbb{E}[X \cdot \mathbf{1}(A) \cdot \mathbf{1}(B)] \\ &= \mathbb{E}[X \cdot \mathbf{1}(A)] \cdot \mathbb{E}[\mathbf{1}(B)] \overset{H \perp \sigma(\mathcal{G},\mathcal{H})}{=} \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \cdot \mathbf{1}(A \cap B)]. \end{aligned}$$

Now assume $X \geqslant 0$; in the general case, decompose $X = X^+ - X^-$ and apply same argument to both $X^\pm$. Define the measures $\mu(F) = \mathbb{E}[X \cdot \mathbf{1}(F)]$ and $\nu(F) = \mathbb{E}[X \cdot \mathbf{1}(F)]$ for all $F \in \sigma(\mathcal{G}, \mathcal{H})$. Observe that $\mu(\Omega) = \nu(\Omega) = \mathbb{E}[X] < \infty$ and we have shown that $\mu = \nu$ on $\mathcal{A}$. Thus, $\mu = \nu$ on $\sigma(\mathcal{G}, \mathcal{H})$ which finally implies the result

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

□

Examples:

Definition 1.9 (Gaussian). $(X_1, X_2, \cdots, X_n) \in \mathbb{R}^n$ has the Gaussian distribution if and only if for all scalars $a_1, a_2, \cdots, a_n \in \mathbb{R}$, $a_1X_1 + \cdots a_nX_n$ has the Gaussian distrubition in $\mathbb{R}$.

A stochastic process (more on that later) $(X_t)_{t\geqslant 0}$ is a Gaussian process if for all $t_1 < t_2 < \cdots t_n$ the vector $(X_{t_1}, X_{t_2}, \cdots, X_{t_n})$ is Gaussian.

Let (X, Y) be a Gaussian vector in $\mathbb{R}^2$. We compute $\mathbb{E}[X|Y]$.
Let $X' = \mathbb{E}[X|Y]$. Looking for f a Borel measurable function s.t. $\mathbb{E}[X|Y] = f(Y)$ a.s. Let $f(y) = ay + b$ for some $a, b \in \mathbb{R}$ to be determined. Now, $X' = aY + b$, $\mathbb{E}[X'] = \mathbb{E}[X] = a\mathbb{E}[Y] + b$ and $\mathbb{E}[X' \cdot Y] = \mathbb{E}[X \cdot Y] \implies \mathbb{E}[(X - X') \cdot Y] = 0$. Thus $\mathrm{Cov}(X - X', Y) = 0 \implies \mathrm{Cov}(X, Y) = a^2\mathrm{Var}(Y)$.

Need to check: that for all Z bounded $\sigma(Y)-$measurable, $\mathbb{E}[(X - X') \cdot Z] = 0$.
Indeed, observe that $(X - X', Y)$ is a Gaussian vector and since $\mathrm{Cov}(X - X', Y) = 0 \implies X - X' \perp Y \implies (X - X') \perp Z$.

2. Let (X, Y) be a random vector with density in $\mathbb{R}^2$ with joint density function $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$. Let $h : \mathbb{R} \to \mathbb{R}$ be a Borel function such that $h(X)$ is integrable. We now compute $\mathbb{E}[h(X)|Y]$.
We have for all g bounded $\sigma Y-$measurable functions.

$$\begin{aligned}\int_{\mathbb{R}^2} h(x)g(y)f_{X,Y}(x, y)\,\mathrm{d}x\,\mathrm{d}y &= \mathbb{E}[h(X)g(Y)]\\ &= \mathbb{E}[\mathbb{E}[h(X)|Y]g(Y)] = \mathbb{E}[\phi(Y)g(Y)]\\ &= \int_{\mathbb{R}^2} \phi(y)g(y)f_{Y(y)}\,\mathrm{d}y\end{aligned}$$

where $f_Y(y) = \int_{\mathbb{R}} f_{X,Y}(x, y)\,\mathrm{d}x$ and $\phi : \mathbb{R} \to \mathbb{R}$ is some Borel measurable function. Hence,

$$\phi(y) = \begin{cases}\displaystyle\int_{\mathbb{R}} h(x)\frac{f_{X,Y}(x, y)}{f_Y(y)}\,\mathrm{d}\mathrm{d}x, & f_Y(y) > 0\\ 0, \quad \text{otherwise}\end{cases}$$

can be seen to work. Thus, we obtain

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

2 Discrete Time Martingales

Definition 2.1 (Filtration). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A filtration is a sequences of increasing sigma sub-algebras of $\mathcal{F}$, $(\mathcal{F}_n)_{n\in\mathbb{N}}$, $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for all $n \in \mathbb{N}$. We call $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}})$ a filtered probability space.

Let $X = (X_n)_{n\in\mathbb{N}}$ be a sequence of random variables/a stochastic process. Then it induces $(\mathcal{F}_n^X)_{n\in N}$, where $\mathcal{F}_n^X := \sigma(X_{:k\leqslant n})$ for all $n \in \mathbb{N}$: the canonical filtration associated to X. We call X adapted to a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ if $X$is $\mathcal{F}_n-$measurable for all $n \in \mathbb{N}.X$ is called integrable if X_n is integrable for all $n \in \mathbb{N}$.

Definition 2.2 (Martingale discrete time). Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ be a filtered probability space. Let $X = (X_n)_{n\in\mathbb{N}}$be an integrable and adapted process.

- X is called a martingale if $\mathbb{E}[X_n|\mathcal{F}_m] = X_m$ a.s. for all $n \geqslant m$.

- X is called a super-martingale if $\mathbb{E}[X_n|\mathcal{F}_m] \leqslant X_m$ a.s. for all $n \geqslant m$.
- X is called a sub-martingale if $\mathbb{E}[X_n|\mathcal{F}_m] \geqslant X_m$ a.s. for all $n \geqslant m$.

Remark. If X is a (super/sub)martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$, then it is also a martingale with respect to $(\mathcal{F}_n^X)_{n\in\mathbb{N}}$. To see this, one has to use the tower property 1.1: $\mathcal{F}_n^X \subseteq \mathcal{F}_n$ for all $n \in \mathbb{N}$ implies $\mathbb{E}[X_n|\mathcal{F}_m^X] = \mathbb{E}[\mathbb{E}[X_n|\mathcal{F}_m]|\mathcal{F}_m^X]$ (since $\mathbb{E}[X_n|\mathcal{F}_m]$ a.s.).

Examples:

1. Let $(\xi_i)_{i\in\mathbb{N}}$ be iid. s.t. $\mathbb{E}[\xi_i] = 0$ for all $i \in \mathbb{N}$ and define $X = (X_n)_{n\in\mathbb{N}}$ by $X_n = \xi_1 + \cdots + \xi_n$ for all $n \in \mathbb{N}$, $X_0 = 0$. X is a martingales with respect to $(\mathcal{F}_n^\xi)_{n\in\mathbb{N}}$.
2. Let $(\xi_i)_{i\in\mathbb{N}}$ be iid. s.t. $\mathbb{E}[\xi_i] = 1$ for all $i \in \mathbb{N}$ and define $X = (X_n)_{n\in\mathbb{N}}$ by $X_n = \prod_{i=1}^{n} \xi_i$ for all $n \in \mathbb{N}$, $X_0 = 1$. X is again a martingales with respect to $(\mathcal{F}_n^\xi)_{n\in\mathbb{N}}$.

Lecture 5

Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ be a filtered probability space.

Definition 2.3 (Stopping time discrete time). A stopping time T is a random variable $T : \Omega \to \mathbb{Z}_+ \cup \{\infty\}$ s.t. $\{T \leqslant n\} \in \mathcal{F}_n$ for all $n \in \mathbb{N}$. Equivalently, if $\{f = n\} \in \mathcal{F}_n$ for all $n \in \mathbb{N}$ since

$$\{T = n\} = \underbrace{\{T \leqslant n\}}_{\mathcal{F}_n} \setminus \underbrace{\{T \leqslant n-1\}}_{\mathcal{F}_{n-1} \subset \mathcal{F}_n} \in \mathcal{F}_n.$$

and

$$\{T \leqslant n\} = \bigcup_{k=1}^{n} \{T = k\} \in \mathcal{F}_k \subset \mathcal{F}_n.$$

Examples:

1. Constant time are trivially stopping times.
2. Let $X = (X_n)_{n\in\mathbb{N}}$ be a stochastic process taking values in $\mathbb{R}$ and $A \in \mathcal{B}(\mathbb{R})$ (X adapted). Define
$$T_A = \{n \geqslant 0 : X_{n\in A}\}.$$
Then $\{T_A \leqslant n\} = \bigcup_{k=0}^{n} \{X_{k\in A}\} \in \mathcal{F}_n$ for all $n \in \mathbb{N}$ (with convention $\inf \varnothing = \infty$).
3. $L_A = \sup\{n \geqslant 0 : X_{n\in A}\}$ is NOT a stopping time.

Properties: $S, T, (T_n)_{n\in\mathbb{N}}$ stopping times. Then $S \wedge T, S \vee T, \inf_n T_n, \sup_n T_n, \liminf_n T_n, \limsup_n T_n$ are also stopping times.

Definition 2.4 (Stopping time sigma algerbra). It $T$is a stopping time, define

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leqslant t\} \in \mathcal{F}_t\}$$

Let $(X_n)_{n\geqslant 0}$ be a process. Write $X_T(\omega) = X_{T(\omega)}(\omega)$ for $\omega \in \Omega$ whenever $T(\omega) < \infty$. Define the stopped process: $X_t^T := X_{T\wedge t}$.

Proposition 2.1. Let S and T be stopping times, and let X be an adapted process, then:

1. If $S \leqslant T$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$.
2. $X_T \cdot$ is $\mathcal{F}_T-$measurable.
3. X^T is adapted.

4. If X is integrable, then the stopped process iss integrable.

Proof. 1. Immediate from definition.

2. Let $A \in \mathcal{B}(\mathbb{R})$. Need to show:

$$\{X_T \mathbf{1}(T < \infty)\} \cap \{T \leqslant t\} \in A, \quad \text{for all } t \geqslant 0.$$

Indeed, we have that

$$\{X_T \mathbf{1}(T < \infty)\} = \bigcup_{s=0}^{t} \underbrace{\{X_s \in A\}}_{\mathcal{F}_s \subseteq \mathcal{F}_t} \cap \underbrace{\{T = s\}}_{\mathcal{F}_s} \in \mathcal{F}_t.$$

3. $X_t^T = X_{T \wedge t}$, this being $\mathcal{F}_{T \wedge t}-$measurable $\subseteq \mathcal{F}_t-$measurable by 1), so we conclude it is $\mathcal{F}_t-$measurable.

4.
$$\begin{aligned}
\mathbb{E}[|X_t^T|] &= \mathbb{E}[|X_{T \wedge t}|] \\
&= \sum_{s=0}^{t-1} \mathbb{E}[|X_s| \cdot \mathbf{1}(T = s)] + \mathbb{E}[|X_t| \cdot \mathbf{1}(T \geqslant t)] \\
&\leqslant \sum_{s=0}^{t} \mathbb{E}[|X_s|] \quad \underbrace{< \infty}_{X_t \text{ is integrable}} .
\end{aligned}$$

□

We now state and prove a fundamental theorem in Martingale theory:

Theorem 2.1 (Optional Stopping Theorem discrete time). Let $(X_n$ be a martingale.

1. If T is a stopping time, then the stopped process X^T is also a martingale. In particular for all $t \geqslant 0$:
$$\mathbb{E}[X_{T \wedge t}] = \mathbb{E}[X_0].$$

2. It $S \leqslant T$ are bounded stopping times, then
$$\mathbb{E}[X_T | \mathcal{F}_S] = X_T, \quad \text{a.s.}$$
and hence $\mathbb{E}[X_T]\mathbb{E}[X_S]$.

3. It there exists an integrable random variable Y such that $|X_n \leqslant Y|$ for all $n \geqslant 0$ and T is finite, then $\mathbb{E}[X_T] = \mathbb{E}[X_0]$.

4. If there exists $M \geqslant 0$ such that $|X_{n+1} - X_n| \leqslant M$ for all $n \in \mathbb{N}$ and T is a stopping time with $\mathbb{E}[T] < \infty$, then $\mathbb{E}[X_T] = \mathbb{E}[X_0]$.

Proof. 1. NTS: for all $t \geqslant 0$, $\mathbb{E}[X_{T \wedge t} | \mathcal{F}_{t-1}] = X_{T \wedge t}$ a.s. Indeed,

$$\begin{aligned}
\mathbb{E}[X_{T \wedge t} | \mathcal{F}_{t-1}] &= \sum_{s=0}^{t-1} \mathbb{E}[X_s \cdot \mathbf{1}(T = s) | \mathcal{F}_{t-1}] \mathbb{E}[X - -t] \cdot \mathbf{1}(T \geqslant t) | \mathcal{F}_{t-1}] \\
&= \sum_{s=0}^{t-1} \mathbf{1}(T = s) \cdot X_s + X_{t-1} \cdot \mathbf{1}(T \geqslant t) \quad \text{a.s.} \\
&= \sum_{s=0}^{t-2} \mathbf{1}(T = s) \cdot X_s + X_{t-1} \cdot \mathbf{1}(T \geqslant t-1) \quad \text{a.s.} \\
&= X_{T \wedge t-1} \quad \text{a.s.}
\end{aligned}$$

2. $S \leqslant T \leqslant n, n \in \mathbb{N}$ fixed. Let $A \in \mathcal{F}_S$. <u>NTS:</u> $\mathbb{E}[X_T \cdot \mathbf{1}(A)] = [X_s \cdot \mathbf{1}(A)]$. We compute

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

Thus,

$$\mathbb{E}[X_T \cdot \mathbf{1}(A)] \overset{(A\in\mathcal{F}_S)}{=} \mathbb{E}[X_S \cdot \mathbf{1}(A)] + \sum_{k=0}^{n-1} \mathbb{E}[(X_{k+1}-X_k)\cdot \mathbf{1}(S \leqslant k < T)\cdot \mathbf{1}(A)]$$

Have, $A \cap \{S \leqslant k\} \in \mathcal{F}_k$ and $A \cap \{T > k\} \in F_k$. Thus, $\mathbf{1}(S \leqslant k < T)\cdot \mathbf{1}(A)$ is $\mathcal{F}_k-$measurable. Using $\mathbb{E}[X_{k+1}|\mathcal{F}_k] = X_k$ a.s. we deduce that

$$\begin{aligned}\mathbb{E}[(X_{k+1}-X_k)\cdot \mathbf{1}(S \leqslant k < T]\cdot \mathbf{1}(A)] &= \mathbb{E}[\overbrace{\mathbb{E}[(X_{k+1}-X_k)|\mathcal{F}_k]}^{0}\cdot \mathbf{1}(S \leqslant k < T]\cdot \mathbf{1}(A)] \\ &= 0\end{aligned}$$

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

3. By the Optional Stopping Theorem applied to $(X_{T\wedge n})_{n\geqslant 0}$, we have

$$\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0] \quad \text{for all } n \geqslant 0.$$

Now, T being finite a.s. implies that $X_T = \lim_{n\to\infty} X_{T\wedge n}$ a.s. By assumption, have $|X_{T\wedge n}| \leqslant Y$ a.s. for all $n \in \mathbb{N}$ and so can apply DCT to conclude.

4. Observe that for all $n \geqslant 1$

$$X_{T\wedge n} - X_0 = \sum_{k=0}^{n-1}(X_k - X_0)\cdot \mathbf{1}(T=k) + (X_n - X_0)\mathbf{1}(T \geqslant n)$$

Hence, we have the bound (using that $|X_{k+1}-X_k| \leqslant M$ a.s. for all $k \geqslant 0$)

$$\begin{aligned}|X_{T\wedge n} - X_0| &\leqslant M\sum_{k=0}^{n-1} k\mathbf{1}(T=k) + n\mathbf{1}(T\geqslant n) \\ &\leqslant \mathbb{E}[T] < \infty \quad \text{a.s.}\end{aligned}$$

Now, $\mathbb{E}[T] < \infty$ gives $T < \infty$ a.s. and so can deduce as before that $X_T = \lim_{n\to\infty} X_{T\wedge n}$ and use the DCT to conclude the argument.

□

Corollary 2.1.1. Let X be a positive superartingale, T a stopping time such that $T < \infty$ a.s., then

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

Proof. Use Fatou 1.1: $\mathbb{E}[\liminf_{t\uparrow\infty} X_{T\wedge t}] \leqslant \liminf_{t\uparrow\infty} \mathbb{E}[X_{T\wedge t}] \leqslant \mathbb{E}[X_0]$. □

Simple random walk on $\mathbb{Z}$

Let $(\xi_i)_{i\geqslant 0}$ be iid Bernoulli random variables with success probability 1/2. Define the process $(X_n)_{n\geqslant 0}$ by setting $X_n = \xi_1 + \cdots + \xi_n$ for all $n \geqslant 1$ and $X_0 = 0$. Furthermore, let $T = \inf\{n \geqslant 0 : X_n = 1\}$. Using the analysis below, we will see that $\mathbb{P}(T < \infty) = 1$. The Optional Stopping Theorem gives $\mathbb{E}[X_{T\wedge t}] = 0$ for all $t \geqslant 0$. Yet, $1 = \mathbb{E}[X)_T] \neq 0$. We thus see that the condition $\mathbb{E}[T] < \infty$ in 4) is necessary, since $\mathbb{E}[T] = \infty$.

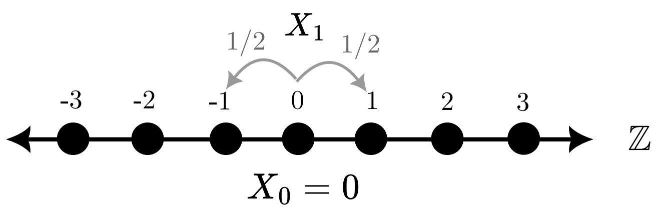

Figure 1: Illustration of simple random walk (first step) on $\mathbb{Z}$.

Lecture 6

We consider again the example of the simple random walk 2 $(X_n)_{n\in\mathbb{N}}$ and define the stopping times

$$T_c = \inf n \geqslant 0 : X_{n=c}, \quad c \in \mathbb{Z}$$

Set $T = T_{-a} \wedge T_b$ for $ab \in \mathbb{Z}$. We now ask what is $\mathbb{P}(T_{-a} \wedge T_b)$?

To answer this, note first that $X_n^T = X_{T\wedge n}$ is a martingale by the Optional Stopping Theorem and we also have the bounded differences $|X_{n+1} - X_n| \leqslant 1$ for all $n \geqslant 1$.

Claim: $\mathbb{E}[T] < \infty$.
To show this, we will *stochastically dominate* T be a geometric random variable, which automatically has a finite expectation and then conclude using the non-negativity of T. Now we have that for the sequence $\xi_1, \xi_2, \cdots, \xi_{a+b}$ the probability that they all are either $+1$ or -1 is $2 \cdot 2^{-(a+b)}$ by independence, call this event A_1. The same is true for the shifted sequence $\xi_{k(a+b)+1} \cdots \xi_{(k+1)(a+b)}$ for all $k \in \mathbb{N}$, where we call the corresponding event A_k.

Thus, we can bound T by the random variable

$$Z(\omega) = \inf\{n \geqslant 0 : \omega \in A_n\}$$

which has the distribution $Z \sim Geom(2 \cdot 2^{-(a+b)})$. Thus, $\mathbb{E}[T] < \mathbb{E}[Z] \leqslant (a+b) \cdot 2^{a+b-1} < \infty$. Thus, we conclude by the OST that $\mathbb{E}[X_T] = \mathbb{E}[X_0] = 0$. Hence, $-a\mathbb{P}(T_a < T_b) + b\mathbb{P}(T_b < T_{-a}) = 0$ and so a quick computation yields that $\mathbb{P}(T_{-a} < T_b) = \frac{b}{a+b}$.

3 Martingale Convergence Theorem

Theorem 3.1 (Almost sure martingale convergence theorem). Let X be a supermartingale bounded in $\mathcal{L}^1$, i.e. satisfying $\sup_n \mathbb{E}[|X_n|] < \infty$. Then, there exists $X_\infty \in \mathcal{L}^1(\mathcal{F}_\infty), \mathcal{F}_\infty = \sigma(\mathcal{F}_n : n \geqslant 0)$ such that $X_n \overset{n\to\infty}{\longrightarrow} X_\infty$, a.s.

Before we embark on the proof of this theorem, we need so me supporting results. First we have a result from analysis and we set up some notation. Let $x - (x_{n n\in\mathbb{N}})$ be a real sequence and let $a < b$ be reals. We proceed to define the *number of upcrossings of the sequence* before time $n \in \mathbb{N}$. Wec constructrecursively the sequence of times:

$$\begin{aligned} T_0(x) &= 0 \\ S_{k+1}(x) &= \inf\{n \geqslant T_k(x) : x_n \leqslant a\} \\ T_{k+1}(x) &= \inf\{n \geqslant S_{k+1}(x) : x_n \geqslant b\} \end{aligned}$$

and

$$N_n([a,b], X) = \sup\{k \geqslant 0 : T_k(x) \leqslant n\}$$

Observe that as $n \to \infty$, $N_n([a,b],x) \uparrow N([a,b],x) = \sup\{kgeq0 : T_k(x) < \infty\}$ (see figure 2 for an illustration).

Lemma 3.1. Let $X = (X_n)$ be a real sequence. Then X converges in $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ if and on ly if for all $a < b$, $a, b \in \mathbb{Q}$, $N([a,b], X) < \infty$.

Proof. $\Longrightarrow$: Suppose x converges, if $a < b$ such that $N([a,b],x) = \infty$, then $\liminf_n x_n \leqslant a < b \leqslant \limsup_n x_n$, a contradiction.
$\Longleftarrow$: if not, then $\liminf_n x_n < \limsup_n x_n$ which implies that there exists $a < b$ in $\mathbb{Q}$ with $\liminf_n x_n < a < b < \limsup_n x_n$, and hence $N([a,n],x) = \infty$, a contradiction. □

Now we state it Doob's upcrossing Inequality

Lemma 3.2 (Doob's upcrossing inequality). Let X be a supermartingale, then for all $n \in \mathbb{N}$:

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

Proof. It is not hard to check that the sequences of times in 3 are stopping times. Now we have:

$$\begin{aligned}&\sum_{k=1}^{n}(X_{T_k\wedge n}-X_{S_k\wedge n})\\&=\underbrace{\sum_{k=1}^{N_n}(X_{T_k}-X_{S_k})}_{\geqslant N_n\cdot(b-a)}+(X_n-X_{S_{N_n+1}})\mathbf{1}(S_{N_n+1}\leqslant n)\end{aligned}$$

Since $T_{k\wedge n}\geqslant S_{k\wedge n}$, the OST gives $\mathbb{E}[X_{T_k\wedge n}]\leqslant \mathbb{E}[X_{S_k\wedge n}]$. Note:

$$\underbrace{X_n-X_{S_{N_n}+1}}_{\geqslant (X_n-a)\wedge 0=-(X_n-a)^-}\quad \mathbf{1}(S_{N_n+1}\leqslant n).$$

Thus, taking expectations on both sides gives:

$$0\geqslant (b-a)\cdot\mathbb{E}[N_n]-\mathbb{E}[(X_n-a)^-].$$

thus concluding the proof. □

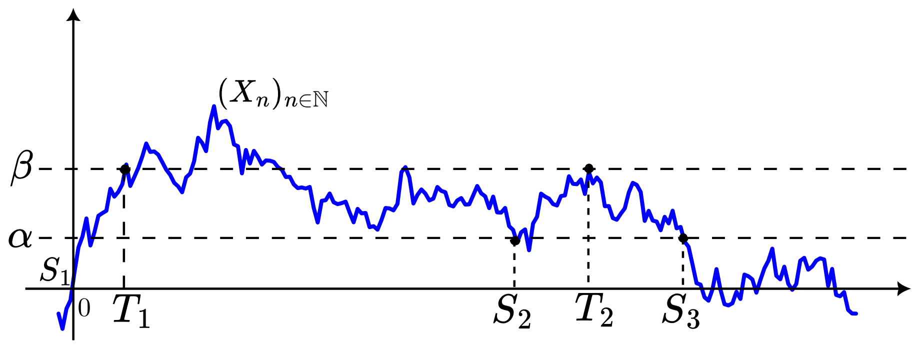

Figure 2: Illustration of upcrossings for the process $(X_n)_{n\in\mathbb{N}}$.

Now we proceed to the proof of the martingale convergence theorem:

Proof. (Theorem 3.1) Fix $a<b$, in $\mathbb{Q}$. Have

$$\begin{aligned}\mathbb{E}[N_{n([a,b],X)}] &\leqslant (b-a)^-\underbrace{\mathbb{E}[(X_n-a)^-]}_{\leqslant \mathbb{E}[|X_n|+a]}\\&\leqslant (b-a)^-\left(\sup_{n\geqslant 0}\underbrace{\mathbb{E}[|X_n|]}_{<\infty}+a\right)\end{aligned}$$

Also have $N_n([a,b],X)\uparrow N([a,b],X)$ as $n\to\infty$. By monotone convergence: $\mathbb{E}[N([a,b],X)]<\infty$. Set

$$\Omega_0=\bigcap_{a<ba,b,\in\mathbb{Q}}\{N([a,b],X)<\infty\}\in\mathcal{F}_\infty$$

and $\mathbb{P}(\Omega_0)=1$. On Ω_0, X converges. set

$$X_\infty=\begin{cases}\lim\limits_{n\to\infty}X_n & \text{on } \Omega_0\\ 0, & \text{on } \Omega\backslash\Omega_0.\end{cases}$$

So, X_∞ is $\mathcal{F}_\infty-$measurable, $X_n\overset{n\to\infty}{\longrightarrow}X_\infty$ a.s. and

$$\mathbb{E}[|X_\infty|]=\mathbb{E}[\liminf_n |X_n|]\leqslant \liminf_{\mathbb{E}[X_n]}<\infty.$$

□

Corollary 3.1.1. Let B be a upermaartingale. Then, X converges a.s.

Proof. $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \leqslant \mathbb{E}[X_0]$. Apply the martingale convergence theorem to conclude. □

Lecture 7

4 Doob's inequalities

Theorem 4.1 (Doob's maximal inequality). Let X be a non-negative submartingale and set $X_n^* = \sup_{0 \leqslant k \leqslant n} X_k$. Then for all $\lambda \geqslant 0$,

$$\begin{aligned} \lambda \cdot \mathbb{P}(X_n^* \geqslant \lambda) &\leqslant \mathbb{E}[X_n \cdot \mathbf{1}(X_n^* \geqslant \lambda)] \\ &\leqslant \mathbb{E}[X_n]. \end{aligned}$$

Proof. Let $T = \inf\{k \geqslant 0 : X_k \geqslant \lambda\}$ (it is a stopping time). Then $\{X_n^* \geqslant \lambda\} = \{T \leqslant n\}$. Also have that $X_{T\wedge n}$ is a submartingale by the OST. Then $\mathbb{E}[X_{T\wedge n}] \leqslant \mathbb{E}[X_n]$. Now,

$$\begin{aligned} \mathbb{E}[X_{T\wedge n}] &= \mathbb{E}[X_T \cdot \mathbf{1}(T \leqslant n)] \\ &+\mathbb{E}[X_n \cdot \mathbf{1}(T > n)] \\ &\geqslant \lambda \cdot \mathbb{P}(T \leqslant n) + \mathbb{E}[X_n \cdot \mathbf{1}(T > n)] \\ &\implies \lambda \cdot \mathbb{P}(T \leqslant n) \leqslant \mathbb{E}\left[X_n \cdot \mathbf{1}(\underbrace{T \leqslant n}_{=\{X_n^* \geqslant \lambda\}}) \right] \\ &\leqslant \mathbb{E}\left[X_n\right] \end{aligned}$$

□

Theorem 4.2 (Doob's $\mathcal{L}^1$ inequality). Lte $p > 1$ and let X be a martingale or a non-negative submartingale. Set $X_n^* = \sup_{0 \leqslant k \leqslant n} |X_k|$. Then

$$\left\|X_n^8\right\|_p \leqslant \frac{p}{p-1} \left\|X_n\right\|_p .$$

Proof. By Jensen, it is enough to prove 4.2 for a non-negative submartingale. Now, observe that

$$\begin{aligned} & = b \\ (y \wedge k)^p &= \int_k^0 px^{p-1}\mathbf{1}(\mathbf{y} \geqslant \mathbf{x})\,\mathrm{d}x = \mathbb{E}[\int_0^k [x^{p-1}\mathbf{1}(X_n^8)\,\mathrm{d}x] \\ &\overset{\text{Fubini}}{=} \int_0^k px^{p-1} \underbrace{\mathbb{P}(X_n^* \geqslant x)}_{\leqslant \frac{1}{x}\mathbb{E}\left[X_n \cdot \mathbf{1}(X_n^* \geqslant x)\right]} \mathrm{d}x \\ &\leqslant \mathbb{E}\left[\int_0^k px^{p-2} \cdot \mathbf{1}(X_n^* \geqslant x)\,\mathrm{d}x \cdot X_n \right] \\ &= \mathbb{E}\left[\tfrac{p}{p-1}(X_n^* \wedge k)^{p-1} \cdot X_n \right] \\ &\overset{\text{Hölder}}{\leqslant} \tfrac{p}{p-1} \cdot \left\|X_n\right\|_p \cdot \left\|X_n^* \wedge k\right\|_p^{p-1} . \end{aligned}$$

So we proved $\left\|X_n^* \wedge k\right\|_p^p \leqslant \frac{p}{p-1} \left\|X_n\right\|_p \cdot \left\|X_n^* \wedge k\right\|_p^{p-1}$, which implies $\left\|X_n^* \wedge k\right\|_p \leqslant \frac{p}{p-1} \cdot \left\|X_n\right\|_p$. Now take $k \to \infty$ and use monotone convergence to conclude the argument. □

Theorem 4.3 ($\mathcal{L}^p$-convergence theorem). Let X be a martingale and $1 < p < \infty$, then the following are equivalent:

1. X is bounded in $\mathcal{L}\checkmark$, i.e. $\sup_{n \geqslant 0} \left\|X_n\right\|_p < \infty$.

2. X converges 'underlinealmost surely and in $\mathcal{L}^p$ to a limit $X_\infty \in \mathcal{L}^p$.

3. There exists $Z \in \mathcal{L}^p$ s.t. $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ a.s.

Proof. 1) $\implies$ 2): X bounded in $\mathcal{L}^p$ implies X is bounded in $\mathcal{L}^1$. So there exists X_∞ such that $X_n \overset{n\to\infty}{\longrightarrow} X_\infty$ a.s.
Also, $\mathbb{E}[|X_\infty|^p] = \mathbb{E}\left[\liminf_n |X_n|^p\right] \overset{\text{Fatou}}{\leqslant} \liminf_{\mathbb{E}[|X_n|^p]} < \infty$. Thus, $X_\infty \in \mathcal{L}^p$.

Now, let $X_n^* = \sup_{0\leqslant k\leqslant n} |X_k|$, $X_\infty^* = \sup_{k\in\mathbb{N}} |X_k|$. Thus,

$$|X_n - X_\infty| \leqslant 2X_\infty^*$$

for all $n \in \mathbb{N}$. Thus, it is enough to show by DCT that $X_\infty^* \in \mathcal{L}^p$. By Doob's $\mathcal{L}^p-$inequality, $\|X_n^*\|_p = \frac{p}{p-1} \cdot \sup_{n\in\mathbb{N}} \|X_n\|_p < \infty$ By MCT ($X_n^* \uparrow X_\infty^*$): $\|X_\infty^*\|_p \leqslant \frac{p}{p-1} \sup_{n\in\mathbb{N}} \|X_n\|_p < \infty$ Thus, $X_\infty^* \in \mathcal{L}^p$.

2) $\implies$ 3): $X_n \overset{n\to\infty}{\longrightarrow} X_\infty$ a.s. and in $\mathcal{L}^p$. Set $Z = X_\infty$. Need to show: $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$ for all $n \in \mathbb{N}$.

$$\begin{aligned}\|X_n - \mathbb{E}[X_{oo}|\mathcal{F}_n]\|_p &\overset{m\geqslant n}{=} \|\mathbb{E}[X_m - X_\infty|\mathcal{F}_n]\|_p \\ &\overset{\text{contraction (Jensen)}}{\leqslant} \|X_m - X_\infty\|_p \to 0, \quad m\to\infty.\end{aligned}$$

3) $\implies$ 1): By conditional Jensen, we can conclude. $\square$

Definition 4.1. A martingale of the form $X_n = \mathbb{E}[Z|\mathcal{F}_n]$, $Z \in \mathcal{L}^p$ is called a martingale closed in $\mathcal{L}^p$.

Corollary 4.3.1. Let $Z \in \mathcal{L}^p$, $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ a.s. Then $X_n \overset{n\to\infty}{\longrightarrow} \mathbb{E}[Z|\mathcal{F}_\infty]$ a.s. and in $\mathcal{L}^p$ where $F_\infty = \sigma(X_n, n \geqslant 0)$.

Proof. By theorem 4.3, we have $X_n \overset{n\to\infty}{\longrightarrow} X_\infty$ a.s. And in $\mathcal{L}^p$. Now, we need to show:

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

Now, we have that X_∞ is $\mathcal{F}_\infty-$measurable (being the point wise limit of $X_n, n \geqslant 0$) and for all $A \in \mathcal{F}_\infty$, $\mathbb{E}[Z \cdot \mathbf{1}(A)] = \mathbb{E}[X_\infty \cdot \mathbf{1}(A)]$. Fix $A \in \bigcup_{n\geqslant 0} \mathcal{F}_n$, a $\pi-$system generating $\mathcal{F}_\infty$. There exists $N \in \mathbb{N}$ such that $A \in \mathcal{F}_N$. Let $n \geqslant N$, now

$$\mathbb{E}[Z \cdot \mathbf{1}(A)] = \mathbb{E}[X_n \cdot \mathbf{1}(A)] \overset{n\to\infty}{\longrightarrow} \mathbb{E}[X_\infty \cdot \mathbf{1}(A)].$$

$\square$

Definition 4.2 (Uniform integrability). A collection of variables $(X_i)_{i\in I}$ is called uniformly integrable (UI) if

$$\sup_{i\in I} \mathbb{E}[|X_i| \cdot \mathbf{1}(|X_i| > M)] \overset{M\to\infty}{\longrightarrow} 0.$$

Equivalently, $(X_i)_{i\in I}$ is UI if (X_i) is bounded in $\mathcal{L}^1$ and for all $\epsilon > 0$, there exists $\delta > 0$ such that for all $A \in \mathcal{F}$ with $\mathbb{P}(A) < \delta$,

$$\sup_{i\in I} \mathbb{E}[|X_i| \cdot \mathbf{1}(A_i)] < \epsilon.$$

- A UI family is bounded in $\mathcal{L}^1$.
- If a family (X_i) is bounded in $\mathcal{L}^p$, $p > 1$, then it is also UI.

Lemma 4.1. Let $(X_n)_{n\in\mathbb{N}}$, X be in $\mathcal{L}^1$ and $X_n \xrightarrow{n\to\infty} X$ a.s. Then $X_n \xrightarrow{n\to\infty}$ in $\mathcal{L}^1$ if and only if $(X_n)_{n\in\mathbb{N}}$ is UI.

Theorem 4.4. Let $X \in \mathcal{L}^1$. The family $\{\mathbb{E}[X|\mathcal{G} : \mathcal{G} \subset \mathcal{F}]\}$ is uniformly integrable (UI).

Proof. Need to show for all $\epsilon > 0$, there exists λ large enough such that for all $\mathcal{G} \subset \mathcal{F}$

$$\begin{aligned}&\mathbb{E}[||\mathbb{E}[X\mathcal{G}] \cdot \mathbf{1}(|\mathbb{E}[X\mathcal{G}]| > \lambda)] < \epsilon \\ &\leqslant \mathbb{E}[\mathbb{E}[|X||\mathcal{G}] \cdot \mathbf{1}(|\underbrace{\mathbb{E}[X|\mathcal{G}|]}_{\mathcal{G}-\text{measurable}}| > \lambda)].\end{aligned}$$

Since $X \in \mathcal{L}^1$, for all $\epsilon > 0$, there exists $m\delta > 0$ such that if $A \in \mathcal{F}$, $\mathbb{P}(A) < \delta$, then $\mathbb{E}[|X| \cdot \mathbf{1}(A)] < \epsilon$. Now,

$$\begin{aligned}\mathbb{P}(|\mathbb{E}[X\mathcal{G}]| > \lambda) &\overset{\text{Markov}}{\leqslant} \frac{\mathbb{E}[|\mathbb{E}[X\mathcal{G}]|]}{\lambda} \\ &\leqslant \frac{\mathbb{E}[\mathbb{E}[|X|\mathcal{G}]]}{\lambda} = \frac{\mathbb{E}[|X|]}{\lambda}.\end{aligned}$$

Take $\lambda = \frac{\mathbb{E}[|X|]}{\lambda}$, then we are done. ☐

Definition 4.3. $X = (X_n)_{n\in\mathbb{N}}$ is called UI (super/sub) martingale if it is a (super/sub) martingale and $(X_n)_{n\geqslant 0}$ is UI.

Examples:
Let $X_1, X_2, \cdots$ be an iid sequence with $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 2) = 1/2$. Set $Y_n = X_{12} \cdot \cdots \cdot X_n$, which can be seen to be a martingale. Also have $\mathbb{E}[Y_n] = 1$ for all $n \in \mathbb{N}$ and $Y_n \xrightarrow{n\in\mathbb{N}} Y_\infty = 0$ a.s. by the martingale convergence theorem, not <u>not</u> in $\mathcal{L}^1$ (because it is not UI).

Theorem 4.5. Let X be a martingale. Then the following are equivalent:

1. X is UI.
2. X converges a.s. and in $\mathcal{L}^1$ to X_∞ as $n \to \infty$.
3. There exists $Z \in \mathcal{L}^1$ such that $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ for all $n \geqslant 0$.

Proof. <u>1) $\implies$ 2)</u>: X is bounded in $\mathcal{L}^1$ implies (by the martingale convergence theorem), $X_n \to$ a.s. Since X_n is UI, then $X_n \to X_\infty$ in $\mathcal{L}^1$.

<u>2) $\implies$ 3)</u>: Set $Z = X_\infty$. Need to show: $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$ a.s. Indeed,

$$\begin{aligned}\|X_n - \mathbb{E}[X_\infty|\mathcal{F}_n]\|_1 &\overset{m\geqq n}{=} \|\mathbb{E}[X_m - X_\infty|\mathcal{F}_n]\|_1 \\ &\leqslant \|X_m - X_\infty\|_1 \xrightarrow{m\to\infty} 0.\end{aligned}$$

<u>3) $\implies$ 1)</u>: The tower property implies $(X_n)_{n\in\mathbb{N}}$ is a martingale and the previous theorem implies that $(X_{n n\in\mathbb{N}})$ is UI. ☐

Remark. 1. We get as before, $X_\infty = \mathbb{E}[Z|\mathcal{F}]$ a.s., where $\mathcal{F}_\infty = \sigma(X_n : n \geqslant 0)$.

2. It X were a UI super/sub martingale, then we would get $\mathbb{E}[X_\infty|\mathcal{F}_n] \overset{\geqslant\text{sub}}{\leqslant} X_n$ (check!).

X is UI implies $X_n \to X_\infty$ in $\mathcal{L}^1$ and a.s. Now let T be a stopping time. We can then define

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

Theorem 4.6 (Optional stopping theorem for UI martingales). Let X be a UI martingale and let S, T be stopping times with $S \leqslant T$. Then

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

Proof. We know that $X_n = \mathbb{E}\left[X_\infty \mathcal{F}_n\right]$ a.s. since X is UI. It suffices to prove that for any stopping times T, $\mathbb{E}\left[X_\infty|\mathcal{F}_T\right] = X_T$ a.s. Indeed, $\mathbb{E}\left[X_T|\mathcal{F}_S\right] = \mathbb{E}\left[\mathbb{E}\left[X_\infty|\mathcal{F}_T\right]|\mathcal{F}_S\right]$ and since $S \leqslant T$we have $\mathcal{F}_S \subseteq \mathcal{F}_T$ and hence the tower property would give:

$$\mathbb{E}\left[X_T|\mathcal{F}_S\right] = \mathbb{E}\left[X_\infty|\mathcal{F}_S\right] = X_S$$

a.s. Thus, we need to show: for all T stopping times, $\mathbb{E}\left[X_\infty|\mathcal{F}_T\right] = X_T$ a.s.

1. NTS: $X_T \in \mathcal{L}^1$:

$$\begin{aligned}
\mathbb{E}\left[|X_T|\right] &= \sum_{n=0}^{\infty} \mathbb{E}\left[|X_n \cdot \mathbf{1}(T=n)|\right] + \mathbb{E}\left[|X_\infty| \cdot \mathbf{1}(T=\infty)\right] \\
&\overset{\text{have } X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]}{\leqslant} \sum_{n=0}^{\infty} \mathbb{E}\left[\mathbb{E}\left[|X_\infty \mathcal{F}_n|\right] \cdot \overbrace{\in \mathcal{F}_n}^{\mathbf{1}(T=n)}\right] \\
&+\mathbb{E}\left[|X_\infty \cdot \mathbf{1}(T=\infty)|\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[|X_\infty| \cdot \mathbf{1}(T=n)\right] + \mathbb{E}\left[|X_\infty \cdot \mathbf{1}(T=\infty)|\right] \\
&= \mathbb{E}\left[|X_\infty|\right] < \infty
\end{aligned}$$

as $X_\infty \in \mathcal{L}^1$. It is also not hard to check that X_T is $\mathcal{F}_T-$measurable.

2. NTS: for all $B \in \mathcal{F}_T$: $\mathbb{E}\left[X_\infty \cdot \mathbf{1}(B)\right] = \mathbb{E}\left[X_T \cdot \mathbf{1}(B)\right]$

$$\begin{aligned}
\mathbb{E}\left[X_T \cdot \mathbf{1}(B)\right] &= \sum_{n=0}^{\infty} \mathbb{E}\left[X_n \cdot \underbrace{\mathbf{1}(T=n) \cdot \mathbf{1}(B)}_{\in \mathcal{F}_n}\right] \\
&+\mathbb{E}\left[X_\infty\right. \\
\cdot \mathbf{1}(T=\infty) \cdot \mathbf{1}(B) & \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[X_\infty \cdot \mathbf{1}(T=n) \cdot \mathbf{1}(B)\right] \\
&= \mathbb{E}\left[X_\infty \cdot \mathbf{1}(B)\right]
\end{aligned}$$

□

Definition 4.4 (Backwards martinagles). Let $\cdots \subseteq \mathcal{G}_{-2} \subseteq \mathcal{G}_{-1} \subseteq \mathcal{G}_0$ be a decreasing family of sub sigma algebras of $\mathcal{F}$. We call $X = (X_n)_{n \leqslant 0}$ a backwards martingale if $X_o \in \mathcal{L}^1$ and for all $n \leqslant -1$ $\mathbb{E}\left[X_{n+1}|\mathcal{G}_n\right] = X_n$ a.s. By the tower property, $\mathbb{E}\left[X_0|\mathcal{G}_n\right] = X_n$ for all $n \leqslant 0$. Since $X_0 \in \mathcal{L}^1$, a backwards martingale is automatically UI.

Theorem 4.7 ($\mathcal{L}^p$/a.s. backwards martingale convergence theorem). Let X be a backwards martingale with $X_0 \in \mathcal{L}^p$, $1 \leqslant p < \infty$. Then $X_n \to X_{-\infty}$ as $m \to -\infty$ a.s. and in $\mathcal{L}^p$ and $X_{-\infty} = \mathbb{E}\left[X_o|\mathcal{G}_{-\infty}\right]$ a.s., where $\mathcal{G}_{-\infty} = \bigcap_{n \leqslant 0} \mathcal{G}_n$.

Proof. Set $\mathcal{F}_k = \mathcal{G}_{-n+k}$, $0 \leqslant k \leqslant n$. This is an increasing filtration and $(X_{-n+k})_{0 \leqslant k \leqslant n}$ is $\mathcal{F}_k-$martingale. Let $N_{-n}([a,b], X)$ be the number of upcrossings of the interval $[a,b]$ between $-n$ and 0. Doob's upcrossing inequality gives:

$$(b-a) \cdot \mathbb{E}\left[N_{-n}([a,b], X)\right] \leqslant \mathbb{E}\left[(X_n - a)^-\right].$$

As before, we get that $X_n \to X_{-\infty}$ as $n \to -\infty$ a.s. We also have $X_{-\infty}$is $\mathcal{G}_{-\infty}-$measurable. Also observe that n$X_o \in \mathcal{L}^p$ implies $X_n \in \mathcal{L}^p$ for all $n \leqslant 0$.

Lecture 9 $X_n = \mathbb{E}[X_n|\mathcal{G}_n]$ a.s. (backwards martingale). If $X_n \in \mathcal{L}^p$, $p \in [1,\infty)$ $X_{n \to X_{-\infty}}$ a.s. $n \to -\infty$ a.s. and $X_{-\infty}$ is $\mathcal{G}_{-\infty} = \bigcap_{n \leqslant 0} \mathcal{G}_n-$measurable.

□

Observe we have that $X_n \in \mathcal{L}^p$ by conditional Jensen and using Fatou, we obtain $X_{-\infty} \in \mathcal{L}^p$. Now we need to show that $X_n \to X_{-\infty}$ in $\mathcal{L}^p$. Indeed,

$$\begin{aligned} |X_n - X_{-\infty}|^p &= |\mathbb{E}[X_0|\mathcal{G}_n] - \mathbb{E}\left[X_{-\infty}|\mathcal{G}_\backslash\right]|^p \\ &= |\mathbb{E}\left[X_{0]-X_{-\infty}|\mathcal{G}_n}\right]|^p \\ &\overset{\text{Jensen}}{\leqslant} \underbrace{\mathbb{E}[|X_0 - X_{-\infty}|^p|\mathcal{G}_n]}_{\text{UI family}}. \end{aligned}$$

Hence, $(|X_n - X_{-\infty}|^p)_{n \leqslant 0}$ is UI, hence giving $\mathcal{L}^1$ convergence.

<u>NTS:</u> $X_{-\infty} = \mathbb{E}[X_o|\mathcal{G}_{-\infty}]$ a.s.

Let $A \in \mathcal{G}_{-\infty} = \bigcap_{n \geqslant 0} \mathcal{G}_n$ implies that $A \in \mathcal{G}_n$ for all $n \leqslant 0$. Hence, $\mathbb{E}[X_n \cdot \mathbf{1}(A)] = \mathbb{E}[X_0 \cdot \mathbf{1}(A)]$, for all $n \leqslant 0$. Take $n \to -\infty$ and use $\mathcal{L}^1$ convergence to get $\mathbb{E}[X_{-\infty} \cdot \mathbf{1}(A)] = \mathbb{E}[X_0 \cdot \mathbf{1}(A)]$ to conclude.

5 Applications of martingales

sec: applications of mgs

Theorem 5.1 (Kolmogorov's $0-1$ law). Let (X_i) be iid and for all $n \in \mathbb{N}$, $\mathcal{F}_n = \sigma(X_k : k \geqslant n)$, $\mathcal{F}_\infty = \bigcap_{n \geqslant 0} \mathcal{F}_n$. Then, $\mathcal{F}_\infty$ is trivial, i.e. for all $A'in'F_\infty$, $\mathbb{P}(A) \in \{0,1\}$.

Proof. Let $A \in \mathcal{F}_\infty$. Define $\mathcal{G}_\backslash = \sigma(\mathcal{X}_\backslash : \| \leqslant \backslash)$ and $\mathcal{G}_\infty = \sigma(\mathcal{G}_n, n \geqslant)$. Now, we have that $\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_n]$ is a martingale and

$$\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_n] \overset{n\to\infty}{\longrightarrow} \mathbb{E}[\mathbf{1}(A)|\mathcal{G}_\infty] \quad \text{a.s.}$$

Now, $A \in \mathcal{F}_\infty$ implies that $A \in \mathcal{F}_{n+1}$ and also have $\mathcal{G}_n \perp \mathcal{F}_{n+1}$ and $\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_n] = \mathbb{P}(A)$ a.s., $\mathbb{E}[\mathbf{1}(A)|\mathcal{G}_\infty] = \mathbf{1}(A)$ a.s. since $\mathcal{F}_\infty \subseteq \mathcal{G}_\infty$ implies that $A \in \mathcal{G}_\infty$. So $\mathbb{P}(A) = \mathbf{1}(A)$ a.s. finally giving $\mathbb{P}(A) \in \{0,1\}$. □

Theorem 5.2 (Strong law of large numbers). Let $(X_i)_{i \in I}$ be an iid sequence in $\mathcal{L}^1$ with $\mathbb{E}[X_1]$. Define $S_n = X_1 + \cdots X_n$. Then $\frac{S_n}{n}$ converges a.s. and in $\mathcal{L}^1$ to μ as $n \to \infty$ a.s.

Proof. Define $\mathcal{G} = \sigma(S_n, S_{n+1}\cdots) = \sigma(S_n, X_{n+1}, \cdots)$. For $n \leqslant -1$, $M^n = \frac{S_{-n}}{-n}$. We will show that $(M_n)_{n \leqslant -1}$ is a backwards martingale with respect to $(\mathcal{G}_{-nn \leqslant -1})$. Indeed,

$$\begin{aligned} \mathbb{E}[M_{m+1}|\mathcal{G}_{-m}] &= M_{-m} \quad \text{a.s. for } m \leqslant -1 \\ &= \mathbb{E}\left[\tfrac{S_{-m-1}}{-m-1}|\mathcal{G}_{-m}\right] \overset{setn=-m}{=} \mathbb{E}\left[\tfrac{S_{n-1}}{n-1}|\mathcal{G}_n\right] \\ &= \mathbb{E}\left[\tfrac{S_{n-1}}{n-1}|S_{n-1}, X_{n+1}\cdots\right] \\ &= \mathbb{E}\left[\tfrac{S_n - X_n}{n-1} S_n\right] \\ &= \tfrac{S_n}{n-1} - \mathbb{E}\left[\tfrac{X_n}{n-1}|S_n\right]. \end{aligned}$$

Now since $S_n = X_1 + \overset{\text{iid}}{\cdots} + X_n$, we have that $\mathbb{E}[X_k|S_n] = \mathbb{E}[X_1|]\, S_n$ and so $\frac{S_n}{n-1} - \frac{1}{n-1}\left(\frac{S_n}{n}\right) = \frac{S_n}{n-1}\left(\frac{n-1}{n}\right) = \frac{S_n}{n}$. Hence $\frac{S_n}{n} \overset{n\to\infty}{\longrightarrow} Y$ a.s. and in $\mathcal{L}^1$measurable for all $k \geqslant 0$. Thus Y is

$\underbrace{\bigcap_k \sigma(X_{k+1}, \cdots)}_{\text{Kolmogorov 0-1 law} \implies \text{trivial}}$ −measurable. So there exists $c \in \mathbb{R}$ such that $\mathbb{P}(Y = c) = 1$. So $\frac{S_n}{n} \xrightarrow{n\to\infty}$ in $\mathcal{L}^1$ and hence $c = \mathbb{E}[Y] = \lim_{i\to\infty} \mathbb{E}\left[\frac{S_n}{n}\right] = \mu$ and so finally $c = \mu$. □

Theorem 5.3 (Radon-Nikodym Theorem). Let P and Q be two probability measures on the space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose that $\mathcal{F}$ is countable generated, i.e. there exists a sequence $(F_n)_{n\in\mathbb{N}}$ such that $\mathcal{F} = \sigma(F_n : n \in \mathbb{N})$. Then the following are equivalent:

1. For all $A \in \mathcal{F}$, $\mathbb{P}(A) = 0$ implies $Q(A) = 0$. ($Q << P$).
2. For all $\epsilon > 0$, there exists $\delta > 0$ such that if $A \in \mathcal{F}$ with $\mathbb{P}(A) < \delta$, then $Q(A) < \epsilon$.
3. There exists a non-negative random variable X such that $Q(A) = \mathbb{E}[X \cdot \mathbf{1}(A)]$, for all $A \in \mathcal{F}$.

Remark. X is called a version of the Radon-Nikodym derivative of Q with respect to P, or $X = \frac{\mathrm{d}Q}{\mathrm{d}P}$ on $\mathcal{F}$ a.s.

Proof. 1) $\implies$ 2) : Suppose 2) does not hold, then there exists an $\epsilon > 0$ such that for all $n \in \mathbb{N}$, there exist A_n with $P(A_n) \leqslant \frac{1}{n^2}$ and $Q(A_n) \geqslant \epsilon$. Now, since $\sum_{n=1}^{\infty} P(A_n) < \infty$ Borel-Cantelli implies $P(A_n \text{ i.o}) = 0$ and so $Q(A_n) = 0$. However,

$$\begin{aligned} \{A_n \text{i.o.}\} &= \textstyle\bigcap_n \bigcup_{k\geqslant n} A_k \implies Q(A_n \text{ i.o}) \\ &= \lim_{n\to\infty} Q\left(\textstyle\bigcup_{k\geqslant n} A_n\right) \\ &\geqslant \lim_{n\to\infty} Q(A_n) \geqslant \epsilon \end{aligned}$$

a contradiction.

3) $\implies$ 1) : trivial.

2) $\implies$ 3) : Let $\mathcal{A}_n = \{H_1 \cap \cdots \cap H_n : H_i = F_i \text{ or } F_i^c \text{ for all } i\}$. In other words $\mathcal{A}_n = \left\{F_1, F_2, \cdots, F_n, \bigcup_{k\geqslant n} F_k\right\}$. Let $\mathcal{F}_N = \sigma(\mathcal{A}_n)$, so $\mathcal{F}_n$ is a filtration.

Now defined

$$X_n(\omega) = \sum_{A\in\mathcal{A}_\setminus} \frac{Q(A)}{P(A)} \cdot \mathbf{1}(\omega \in A).$$

Thus, for all $A|in\mathcal{F}_n$, $\mathbb{E}[X_n \cdot \mathbf{1}(A)] = Q(A) \stackrel{F_n\subseteq\mathcal{F}_{n+1}}{=} \mathbb{E}[X_{n+1} \cdot \mathbf{1}(A)]$. So $(X_n)_{n\in\mathbb{N}}$ is indeed a martingale. Furthermore $\mathbb{E}[X_n] = Q(\Omega) = 1$ (and since $X_n \geqslant 0$ for all $n \geqslant 0$), we have that X_n is an $\mathcal{L}^1$ bounded martingale. Thus, $X_n \xrightarrow{n\to\infty} X_\infty$ a.s.

Now we show that $(X_n)_{n\in\mathbb{N}}$ is UI:

$$\begin{aligned} \mathbb{P}(X_n \geqslant \lambda) &\leqslant 1/\lambda < \infty \\ &\leqslant \delta \end{aligned}$$

using Markov's inequality and taking $\lambda = 1/\delta$. Thus, $\mathbb{E}[X_n \cdot \mathbf{1}(X_n \geqslant \lambda)] = Q(X_n \geqslant \lambda) < epsilon$. Thus $(X_n)_{n\in\mathbb{N}}$ is UI and so $X_n \to X_\infty$ in $\mathcal{L}^1$.

Now define $\tilde{Q}(A) = \mathbb{E}[X_\infty \cdot \mathbf{1}(A)]$. Want to show: $\tilde{(Q)}(A) = Q(A)$ for all $A \in \mathcal{F}$. Indeed, we have $X_n = X_\infty|\mathcal{F}_n$. Now if we let for a moment $A \in \bigcup_{n\geqslant 0} \mathcal{F}_n$, there exists some $N \in \mathbb{N}$ such that $A \in \mathcal{F}_N$. Thus,

$$\underbrace{\mathbb{E}[X_N \cdot \mathbf{1}(A)]}_{=Q(A)} = \underbrace{\mathbb{E}[X_\infty \cdot \mathbf{1}(A)]}_{=\tilde{Q}(A)}.$$

Hence, $Q = \tilde{Q}$ on a $\pi-$system, $(\bigcup_n \mathcal{F}_n)$, that generates $\mathcal{F}$, and by the extension theorem we have that $Q \equiv \tilde{Q}$ everywhere. □

Lecture 10

6 Continuous Time processes

Let $X = (X_n)_{n\in\mathbb{N}}$ be a process, that is for all $n \in \mathbb{N}$ X_n is a random variable on some underlying probability space $(\Omega, \mathcal{F}, \mathbb{P}))$. X can also be viewed as the map

$$X : (\omega, n) \mapsto X_n(\omega).$$

and observe that this map is $\mathcal{F}\otimes\mathcal{P}(\mathbb{N}) = \sigma(\{A\times\{k\} : A \in \mathcal{F}, k \in \mathbb{N}\})$ as long as X_n is $\mathcal{F}-$measurable for all $n \in \mathbb{N}$. Now we consider random variables taking values in the spaces $\mathbb{R}^d$, $d \geqslant 1$.

Definition 6.1 (Stochastic process). The family $(X_t)_{t\in\mathbb{R}_+}$ is called a stochastic process if for all t positive X_t is a random variable.

Remark. The map $X : (\omega, t) \mapsto X_t(\omega)$ need not be $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)-$measurable.

Claim: If for all $\omega \in \Omega$, $\mapsto X_t(\omega)$ is a continuous function for $t \in (0, 1]$, then the map $X : (\omega, t) \mapsto X_t(\omega)$ is $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)-$measurable.

Indeed, by continuity we can write

$$X_t(\omega) = \lim_{n\to\infty} \overbrace{\sum_{k=0}^{2^n-1} \mathbf{1}(t \in (k \cdot 2^{-n}, (k+1) \cdot 2^{-n}])X_{k\cdot 2^{-n}}(\omega)}^{\text{for all } n \text{this sum is} \mathcal{F}\otimes\mathcal{B}((0,1])-\text{meas.}}$$

Thus X is measurable with as a limit of measurable functions.

From now onwards, we will always (unless otherwise stated) assume that X is right-continuous and admits left limits, almost everywhere. We call such processes cadlag.

We now revisit some of the earlier definition we have made in the discrete setting and extend the to the continuous case. A filtration is an increasing family of sigma algebras $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ whenever $t \leqslant t'$. We say X is adapted to the filtration above if X_t if $\mathcal{F}_t-$measurable for all $t \in \mathbb{R}_+$. A random variable $T : \Omega \to [0, \infty]$ is called a stopping time if for all t, $\{T \leqslant t\} \in \mathcal{F}_t$. Define $\mathcal{F}_T = \{A|in\mathcal{F} : A \cap \{T \leqslant t\} \in \mathcal{F}_t \text{ for all } t\}$ and $A|in\mathcal{B}(\mathbb{R})$. Furthermore, $T_A = \inf_{t\geqslant 0: X_t \in A}$ is not always a stopping time.

$$\{T_A \leqslant t\} = \bigcup_{s\leqslant t}\{X_s \in A\}$$

an uncountable union so not immediately clear whether it in $\mathcal{F}_t$.

Examples:
Let $J = \begin{cases} 1, & \text{with probability} \frac{1}{2} \\ -1, & \text{with probability} \frac{1}{2} \end{cases}$ and

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

Let $(\mathcal{F}_t)_{t\geqslant 0} = (\mathcal{F}_t^X)_{t\geqslant 0}$ and fix $A \in (1, 2)$. Then $\{T_A \leqslant 1\}$ | $in\mathcal{F}_1 = \{\varnothing, \Omega\}$, since $\{T_A \leqslant 1\} = \{J = 1\}$.

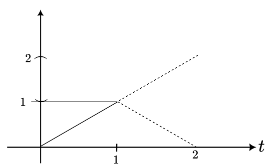

Figure 3: Illustration of X.

Again, we say $X_t^T = X_{T\wedge t}$, $X_T(\omega) = X_{T(\omega)}(\omega)$ whenever $T(\omega) < \infty$.

Proposition 6.1. Let S, T be stopping times and X a cadlag adapted process. Then

1. If $S \leqslant T$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$.

2. $S \wedge T$ is a stopping time.

3. $X_T \cdot \mathbf{1}(T < \infty)$ is $\mathcal{F}_T-$measurable.

4. X^T is adapted.

Proof. $1), 2)$ are clear (check!) and 4) is immediate from 3), since $X_{T\wedge t}$ if $\mathcal{F}_{T\wedge t}-$measurable and $\mathcal{F}_{T\wedge t} \subseteq \mathcal{F}_t$.

proof of 3): Claim: a random variable $Z$is $\mathcal{F}_{T\wedge t}-$measurable if and only if $Z \cdot \mathbf{1}(T \leqslant t)$ is $\mathcal{F}_t-$measurable for all $t \geqslant 0$. Indeed,

$\Longleftarrow$): is true by definition.

$\Longrightarrow$) : if $Z = c \cdot \mathbf{1}(A)$, $A \in \mathcal{F}$, then $A \in \mathcal{F}_T$ which means that Z is $\mathcal{F}_T-$measurable. Now, if $Z = \sum_i c_i \cdot \mathbf{1}(A_i)$, a finite sum with $c_i > 0$, $A_i \in \mathcal{F}$, then Z is $\mathcal{F}_T-$measurable.

Z general ($\geqslant 0$): let $Z_n \uparrow Z$, where

$$Z_n = 2^{-n}\lfloor 2^n Z\rfloor \wedge n, \quad \text{for all } n \in \mathbb{N}.$$

Observe that Z_n are simple for all n and so by the previous steps Z_n is $\mathcal{F}_T-$measurable and hence so is Z, being an a.s. pointwise limit of measurable functions.

The case for completely general Z follows by decomposing $Z = Z^+ - Z^-$, $Z^+ = Z\vee, Z^- = (-Z) \vee 0$ and apply the previous case to Z^+, Z^-.

Now, by the above claim, it suffice to show: $X_T \cdot \mathbf{1}(T \leqslant t)$ if F_t measurable for all t. We have $X_T\mathbf{1}(T \leqslant t) = X_T \cdot \mathbf{1}(T < t) + X_t \cdot \mathbf{1}(T = t)$. Hence, it suffices to show that $X_T \cdot \mathbf{1}(T < t)$ if F_t measurable for all t.

Define $T_n = 2^{-n}\lceil 2^n T\rceil$, stopping times since

$$\begin{aligned}\{T_n \leqslant t\} &= \{\lceil 2^n T\rceil \leqslant 2^n t\} \\ &= \{2^n T \leqslant \lfloor 2^n t\rfloor\} = \{T \leqslant 2^{-n}\lfloor 2^n T\rfloor\} \\ &\in \mathcal{F}_{2^{-n}\lfloor 2^n T\rfloor} \subseteq \mathcal{F}_t.\end{aligned}$$

Also, $T_n \downarrow T$, as $n \to \infty$. Now by the cadlag property of X,
$X_T \cdot \mathbf{1}(T < t) = \lim_{n\to\infty} X_{T_n\wedge t} \cdot \mathbf{1}(T < t)$.

Furthermore, T_n takes values in $\mathcal{D}_n = \{k \cdot 2^{-n}, k \in \mathbb{N}\}$. Now,

$$\begin{aligned}X_{T_n\wedge t} \cdot \mathbf{1}(T < t) &= \sum_{d\in\mathcal{D}_n, d\leqslant t} \overbrace{X_d \cdot \mathbf{1}(T_n = d) \cdot \mathbf{1}(T < t)}^{\mathcal{F}_t-\text{meas.}} \\ &+ X_t \cdot \mathbf{1}(T_n = t) \cdot \mathbf{1}(T < t).\end{aligned}$$

Hence, $X_T \cdot \mathbf{1}(T < \infty)$ is $\mathcal{F}_t-$measurable as a limit of $\mathcal{F}_t-$measurable functions. $\square$

Proposition 6.2. Let X be a continuous and adapted process and let A be a closed set. Then $T_A = \{t \geqslant 0 : X_t \in A\}$ is a stopping time.

Proof. Need to show: $\{T_A \leqslant t\} = \left\{\inf_{s\in\mathbb{Q}, s\leqslant t} d(X_s, A) = 0\right\}$.

$(\subseteq)$: $d(x, A)$ =distance of x from A. Let $T_A = s \leqslant t$, then there exists a sequence $s_n \downarrow s$, such that $X_{S_n} \in A$. Since A is closed, we have $d(X_s, A) = 0$ and $X_{s_n} \to X_s$, as $n \to \infty$. Again A being closed implies that $d(X_s, A) = 0$. The continuity of X and $d(\cdot, A)$ means that there exists another sequence $(q_n)_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ such that $q_n \uparrow s$ such that $d(X_{q_n}, A) \to 0$ hence $\inf_{s\in\mathbb{Q}, s\leqslant t} d(X_s, A) = 0$.

$(\supseteq)$: If $\inf_{s\in\mathbb{Q}, s\leqslant t} d(X_s, A) = 0$, then there exists a sequence $(s_n)_{n\in\mathbb{N}}$ such that $s_n \leqslant t$ for all n and $d(X_{s_n,A} \to 0)$ as $n \to \infty$. Then by compactness, there exists a convergent subsequence of $s_n \to s$ (without relabelling), such that $s \leqslant t$ and $d(X_{s_n,A}) \to 0$ as $n \to \infty$ and by continuity we obtain $d(X_s, A) = 0$, hence $X_s \in A$ and so $T_A \leqslant t$. $\square$

Definition 6.2. Given a filtration $(\mathcal{F}_t)_{t\geqslant 0}$, we define $\mathcal{F}_{t^+} = \bigcap_{s>t} \mathcal{F}_s$, for all $t \geqslant 0$. Observe that $(\mathcal{F}_{t^+})_{t\geqslant 0}$ is a filtration. If for all $t \geqslant 0$, $\mathcal{F}_{t^+}$, we say $(\mathcal{F}_t)_{t\geqslant 0}$ is right-continuous.

Lecture 11

Proposition 6.3. Let X be a continuous process, and A be an open set. Then

$$T_A = \inf\{t \geqslant 0 : X_t \in A\}$$

is a stopping time with respect to the filtration $(\mathcal{F}_{t^+})_{t\geqslant 0}$.

Proof. Need to show: for all $t \geqslant 0$, $\{T_A \leqslant t\} \in \mathcal{F}_{t^+}$. Have,

$$\begin{aligned} \{T_A < s\} &= \bigcup_{q\in\mathbb{Q}, q<s} \underbrace{X_q \in A}_{\in \mathcal{F}_s} \in \mathcal{F}_s \\ \{T_A \leqslant t\} &= \bigcap_n \underbrace{\{T_A < t + \frac{1}{n}\}}_{\in \mathcal{F}_{t+\frac{1}{n}}} \in \mathcal{F}_{t^+}. \end{aligned}$$

$\square$

Let $(X_t)_{t\geqslant 0}$ be a stochastic process. It can be viewed, as a random element in the space of functions $\{f : \mathbb{R}_+ \to E\}$ endowed with the product sigma-algebra making all projections measurable. Further, let $\mathcal{C}(\mathbb{R}_+, E)$ be the space of all continuous functions and $\mathcal{D}(\mathbb{R}_+, E)$ the space of all cad lag functions. Endow the spaces $\mathcal{C}, \mathcal{D}$ with the sigma algebra that makes all projections $\pi_t : f \mapsto f_t$ measurable for all $t \geqslant 0$. This sigma algebra is generated by the cylinder sets

$$\left\{ \bigcap_{s\in J} \{f_s \in A_s : \text{for all } T \subseteq \mathbb{R}_+, \text{ finite}, A_s \in \mathcal{B}(E)\} \right\}.$$

For A in the product sigma algebra, we write $\mu(A) = \mathbb{P}(X \in A)$ and we call μ the law of X. ("$X_*\mathbb{P} = \mu$"). For every J finite subset of $\mathbb{R}_+$, write μ_J for the law of $(X_t)_{t\in J}$. The measures (μ_J) are called the finite dimensional marginals of X. The μ_J completely characterise the law of μ. This follows because the sets above form a $\pi-$system that generates the sigma fields previously mentioned.

Examples:
Let $X = 0$ for all $t \in [0,1]$ and $U \sim [0,1]$ (uniform) and $X_{t'} = \mathbf{1}(U = t)$ for $t \in [0,1]$. Both of them have the same finite dimensional distributions which are Dirac masses at zero, but the processes are not equal.

$$\begin{aligned} &\mathbb{P}(X_t = 0 \text{ for all } t \in [0,1])) &&= 1 \\ &\mathbb{P}(X'_t = 0 \text{ for all } t \leqslant 1) &&= 0. \quad \text{But,} \\ &\mathbb{P}(X_t = X'_t) &&= 1 \quad \text{for all } t \in [0,1]. \end{aligned}$$

Definition 6.3. Let X and X' be two processes on $(\Omega, \mathcal{F}, \mathbb{P})$, we say X' is a version of X if $(X_t = X'_t \text{ a.s.})$ for all t. That is

$$\text{For all} t \geqslant 0 : \mathbb{P}(X_t = X'_t) = 1.$$

Definition 6.4. Fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P})$. Set $\mathcal{N}$ to be the collection of sets of measure zero. Furthermore, set

$$\tilde{\mathcal{F}}_t = \sigma(\mathcal{F}_t, \mathcal{N})$$

for all $t \geqslant 0$. If for all t, $\mathcal{F}_t = \tilde{\mathcal{F}}_t$, we say that $(\mathcal{F}_t)_{t\geqslant 0}$ satisfies the usual conditions.

Theorem 6.1 (Martingale regularisation theorem). Let $(X_t)_{t \geqslant 0}$ be a martingale wrt $(\mathcal{F}_t)_{t \geqslant 0}$. Then, there exists a cadlag process $(\tilde{X}_t)_{t \geqslant 0}$ satisfying for all $t \geqslant 0$:

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

and X is a martingale with respect to the augmented filtration $(\tilde{\mathcal{F}}_t)_{t \geqslant 0}$. If $(\mathcal{F}_{t)_{t \geqslant 0}}$ satisfies the usual conditions, then $\tilde{X}$ is a version of X.

We start with a Lemma

Lemma 6.1. Let $f : \mathbb{Q}_+ \to \mathbb{R}$ such that for all $I \subseteq \mathbb{Q}_+$ bounded, f is bounded on I and for any $a < b, a, b, \in \mathbb{Q}_+$, for all I bounded and suppose

$$\begin{aligned}\mathcal{N}([a,b], I, f) = \sup\{n \geqslant 0 : &\text{ there exists } 0 < s_1 < t_1 < \cdots < s_n < t_n,\\ &s_i, t_i \in I \text{ s.t. } f(s_i) < a, f(t_i > b)\} < \infty.\end{aligned}$$

Then, for all $t \geqslant 0$, the limits

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

exist and are finite.

Proof. Let $s_n \downarrow t$, the sequence $(f(s_n))$ will converge by the finite upcrossing property (see lemma 3.1). Now suppose $t_n \downarrow t$ is another such sequence, then combining them (by selecting elements from each sequence in an alternating fashion exploiting convergence) we get a decreasing sequence converging to t to conclude $\lim_{n \to \infty} f(s_n) = \lim_{n \to \infty} f(t_n)$. Finally, f being bounded gives that both limits are equal and finite. $\square$

<u>Goal:</u> To define $\tilde{X}_t = \lim_{s \downarrow t, s \in \mathbb{Q}_+} X_s$ on a set of measure 1, and zero otherwise. We now outline below the main steps in the proof of Theorem 6.1.

<u>Steps:</u>

1. Show that the limit exists and is finite on a set of measure one.
2. Show that $\tilde{X}$ is $\tilde{\mathcal{F}}_t$−measurable and satisfies $\mathbb{E}\left[\tilde{X}_t | \mathcal{F}_t\right]$ a.s. for all $t \geqslant 0$.
3. $\tilde{X}$ is a $(\tilde{\mathcal{F}}_t)_{t \geqslant 0}$ martingale.
4. $\tilde{X}$ is cadlag.

Proof. (Theorem 6.1)

1. Let I be a bounded subset of $\mathbb{Q}_+$. Need to show that $\mathbb{P}\left(\sup_{t \in I} |X_t| < \infty\right) = 1$. Observe that

$$\sup_{t \in I} |X_t| = \sup_{J \subseteq I, J \text{ finite}} \sup_{t \in J} |X_t|.$$

Now, let $J = \{j_1, \cdots, j_n\} \subseteq I$ with $j_1 < \cdots j_n$ and $k > \sup I$. Then $(X_t)_{t \in J}$ is a discrete time martingale. Hence the maximal inequality in 4.1 gives

$$\lambda \cdot \mathbb{P}(\sup_{t \in J} |X_t| \geqslant \lambda) \leqslant \mathbb{E}\left[|X_{j_n}|\right] \leqslant \mathbb{E}\left[|X_k|\right]$$

by the martingale property and Jensen. Now taking the limit as $J \uparrow I$,

$$\lambda \cdot \mathbb{P}\left(\sup_{t \in I} |X_t| \geqslant \lambda\right) \leqslant \mathbb{E}\left[|X_{j_n}|\right] \leqslant \mathbb{E}\left[|X_k|\right]$$

So, $\mathbb{P}\left(\sup_{t \in I} |X_t| \geqslant \lambda\right) = 1$. Now for $M \in \mathbb{N}$ define $I_M = \mathbb{Q}_+ \cap [0, M]$, then by the above,

$$\mathbb{P}\left(\bigcap_{M \in \mathbb{N}} \left\{\sup_{t \in I_M} |X_t| < \infty\right\}\right) = 1$$

and on the above event, X_t is bounded on bounded intervals of $\mathbb{Q}_+$.

Lecture 12

Let $a < b$, $a, b \in \mathbb{Q}_+$, $I \subseteq \mathbb{Q}_+$, bounded. Observe that

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

Now, let $J = \{j_1, \cdots, j_n\} \subseteq I$ with $j_1 < \cdots j_n$ and $k > \sup I$. Then $(X_t)_{t \in J}$ is a discrete time martingale. Now, Doob's upcrossing inequality from 3.2 gives

$$\begin{aligned}(b-a) \cdot \mathbb{E}\left[\mathcal{N}([a,b], J, X)\right] &\leqslant \mathbb{E}\left[(X_{j_n} - a)^-\right] \\ &\leqslant \mathbb{E}\left[(X_k - a)^-\right].\end{aligned}$$

By monotone convergence, we get

$$(b-a) \cdot \mathbb{E}\left[\mathcal{N}([a,b], I, X)\right] < \infty.$$

Let $M \in \mathbb{N}$, $I_M = \mathbb{Q}_+ \cap [0, M]$ and

$$\Omega_0 = \bigcap_{m \in \mathbb{N}} \left(\bigcap_{a<b, a, b \in \mathbb{Q}} \{\mathcal{N}([a,b], I_M, X) < \infty\} \bigcup \left\{ \sup_{t \in I_m} |X_t| < \infty \right\} \right).$$

On Ω_0, from lemma 6.1, $\lim_{s \downarrow t X_s}$ exists and we have $\mathbb{P}(\Omega_0) = 1$. Now, define

$$\tilde{X}_t = \begin{cases} \lim_{s \downarrow t, s \in \mathbb{Q}_+} X_s, & \text{on } \Omega_0 \\ 0, & \text{otherwise.} \end{cases}$$

Recall $\tilde{\mathcal{F}_t} = \sigma(\mathcal{F}_t, \mathcal{N})$ for all $t \geqslant 0$. From the definition definition, we see that $\tilde{X}$ is $\tilde{\mathcal{F}}-$adapted.

It remains to show that $X_t = \mathbb{E}\left[\tilde{X}_t | \mathcal{F}_t\right]$ a.s. and $\tilde{X}$ is cadlag and a martingale.

2. Let $t_n \downarrow t$, $t_{n \in \mathbb{Q}_+}$, then
$$\tilde{X}_t = \lim_{n \to \infty} X_{t_n}$$
a.s. Observe that (X_{t_n}) is a backwards martingale with respect to the filtration $(\mathcal{F}_{t_n})_{n \in \mathbb{N}}$. SO (X_{t_n}) converges a.s. and in $\mathcal{L}^1$. In other words, $X_t = in\ \mathcal{L}^1$. So $X_t = \mathbb{E}\left[\tilde{X}_t | \mathcal{F}_t\right]$ a.s.

3. We now prove that $\tilde{X}$ is a martingale. Let $s < t$, we need to show that $\mathbb{E}\left[\tilde{X}_t | \tilde{\mathcal{F}_s}\right] =' tildeX_s$ a.s.

 Claim: $\mathbb{E}\left[X_t | \mathcal{F}_{t^+}\right] = \tilde{X}_s$ a.s. Indeed, first observe that for Y any random variable and $\mathcal{G}$ a sigma algebra it follows that
$$\mathbb{E}\left[Y | \sigma\mathcal{G}, \mathcal{N})\right] = \mathbb{E}\left[X | \mathcal{G}\right]$$
which is clear because the conditional expectation is defined almost surely and $\mathcal{N}$ only contains sets of measure zero.

 Now, fix $s < t$ and let $s_n \downarrow s$, $s_n \in \mathbb{Q}_+$, $s_0 < t$. We have by the tower property that $(\mathbb{E}\left[X_t | \mathcal{F}_{s_n}\right])_{n \in \mathbb{N}}$ is a backwards martingale and so it converges a.s. and in $\mathcal{L}^1$ to $\mathbb{E}\left[X_t | \mathcal{F}_{t^+}\right]$. But $\mathbb{E}\left[X_t | \mathcal{F}_{s_n}\right] = X_{s_n}$ a.s. and $X_{s_s} \to \tilde{X}_s$ a.s. as $n'to\infty$. So $\tilde{X}_s = \mathbb{E}\left[X_t | \mathcal{F}_{s^+}\right]$.

4. Finally, we show that $\tilde{X}$ is a cadlag. First we show that $\tilde{X}$ is right continuous. Suppose not. Then, there exists $\omega \in \Omega_0$ and some $t \geqslant 0$ such that $\tilde{X}(\omega)$ is not right continuous at t. That is there exists a sequence $s_n \downarrow t$ such that $|\tilde{X}_{s_n} - \tilde{X}_t| \geqslant \epsilon > 0$ (for some positive ϵ). By the definition of $\tilde{X}$, there exists another sequence $s'_n > s_n$, for all $n \in \mathbb{N}$ and $s'_n \downarrow t$, $s'_n \in \mathbb{Q}_+$ such that $|\tilde{X}_{s_n} - X_{s'_n}| \leqslant \frac{\epsilon}{2}$. So $|X_{s'_n} - \tilde{X}_t| \geqslant \frac{\epsilon}{2}$, a contradiction since $s'_n \downarrow t$, $s'_n \in \mathbb{Q}_+$. The argument for left continuity is entirely analogous.

□

Examples:
Let ξ, η be independent iid symmetric Bernoulli with success probability 1/2. Define

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

and let $\mathcal{F}_t = \sigma(X_s, s <\leqslant)$ for all $t \geqslant 0$. Observe that X is an $(\mathcal{F}_t)_{t\geqslant 0}$ martingale. Also, $\tilde{X}$ satisfies $X_t = \mathbb{E}\left[\tilde{X}_t|\mathcal{F}_t\right]$ where

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

Furthermore, $\mathcal{F}_1 = \sigma(\xi)$ and $\mathcal{F}_t = \sigma(\xi, \eta)$ for all $t > 1$, $\tilde{X}$ is cadlag with respect to $\tilde{F}$. Observe finally that $\mathcal{F}_{1^+} = \sigma(\xi, \eta)$ and so the filtration $\mathcal{F}$ is not right continuous and $\tilde{X}$ is not a version of X. We thus see that the right-continuity of $(\mathcal{F}_t)_{t\geqslant 0}$ is necessary in Theorem 6.1.

Theorem 6.2 (Almost sure martingale convergence theorem). Let X be a cadlag martingale bounded in $\mathcal{L}^1$. Then $X_t \to X_\infty$ a.s. with $X_\infty \in \mathcal{L}^1(\mathcal{F}_\infty)$.

Proof. Let $I_M = \mathbb{Q}_+ \cap [0, M]$. Then Doob's upcrossing inequality 3.2 from the discrete setting and a monotone convergence argument give for $a < b, a, b \in \mathbb{Q}_+$

$$(b - a) \cdot \mathbb{E}\left[\mathcal{N}([a, b], I_M, X)\right] \leqslant a + \sup_{t\geqslant 0} \mathbb{E}\left[|X_t|\right].$$

Taking $M \to \infty$ gives $\mathcal{N}([a, b], \mathbb{Q}_+, X) < \infty$ a.s. Hence, for the event

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

we have $\mathbb{P}(\Omega_0) = 1$ and on Ω_0, $\lim_{q\to\infty, q\in\mathbb{Q}_+} X_q$ exists and is finite. We thus have $X_\infty = \lim_{q\to\infty, q\in\mathbb{Q}_+} X_q$ on Ω_0. Now for all $\epsilon > 0$, there exists q_0 such that $|X_{q_0} - X_\infty| \leqslant \frac{\epsilon}{2}$ for all $q > q_0$, $q \in \mathbb{Q}_+$. Now let $t > q_0$. Then there exists some $q > t$, $q \in \mathbb{Q}_+$ such that $|X_t - X_q| \leqslant \frac{\epsilon}{2}$ by right continuity of X. So $|X_t - X_\infty| \leqslant \epsilon$.

□

Theorem 6.3 (Doob's maximal inequality). Let X be a cadlag martingale, $X_t^* = \sup_{s\leqslant t} |X_s|$. Then for all $\lambda > 0$,

$$\lambda \cdot \mathbb{P}(X_t^* \geqslant \lambda) \leqslant \mathbb{E}\left[|X_t| \cdot \mathbf{1}(X_t^* \geqslant \lambda)\right] \leqslant \mathbb{E}\left[|X_t|\right].$$

Proof. Have

$$\sup_{s\leqslant t} |X_s| = \sup_{s\in\{t\}\cup(\mathbb{Q}_+\cap[0,t])} |X_s|$$

and use the beginning of the proof of theorem 6.1. □

Theorem 6.4 (Optional stopping theorem for cadlag UI martingales). Let X be a cadlag UI martingale, then for all $S \leqslant T$ stopping times

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

Proof. Let $T_n = 2^{-n}]2^n T]$ and $S_n = 2^{-n}[2^n S]$. Both are stopping times and $T_n' downarrow T$, $S_n \downarrow S$ as $n \to \infty$. <u>need to show:</u> for $A' in \mathcal{F}_S$, then $\mathbb{E}\left[X_T \cdot \mathbf{1}(A)\right] = \mathbb{E}\left[X_S \cdot \mathbf{1}(A)\right]$. Indeed, $X_{T_n} \to X_T$ and $X_{S_n} \to X_S$ a.s. as $n \to \infty$ (X is right continuous).

Now, by the discrete optional stopping theorem applied to the martingale $(X_{k\cdot 2^{-n}})_{k\in\mathbb{N}}$ with respect to the filtration $(\mathcal{F}_{K\cdot 2^{-n}})_{k\in\mathbb{N}}$, $X_{T_n} = \mathbb{E}\left[X_\infty|\mathcal{F}_{T_n}\right]$, so X_{T_n} is UI (since T_n take values in

$2^{-n}\cdot\mathbb{N}$). Thus, $X_{T_n} \to X_T$ in $\mathcal{L}^1$, and the same holds for $X_{S_n} \to X_S$ using the exact same argument. By the discrete optional stopping theorem, we have that $\mathbb{E}[X_{T_n}|\mathcal{F}_{S_n}] = X_{S_n}$ a.s. Now for $A \in \mathcal{F}_S$, we have that $A \in \mathcal{F}_{S_n}$ for all $n \in \mathbb{N}$ since $S_n \geqslant S$. So $\mathbb{E}[X_{T_n} \cdot \mathbf{1}(A)] = \mathbb{E}[X_{S_n} \cdot \mathbf{1}(A)]$. □

Theorem 6.5 (Kolmogorov's continuity criterion). Let $\mathcal{D}_n = \{K \cdot 2^{-n} : 0 \leqslant k \leqslant 2^n\}$ and $\mathcal{D} = \bigcup_{n \geqslant 0} \mathcal{D}_n$. Let $(X_t)_{t \in \mathcal{D}}$ be a stochastic process taking real values. Suppose there exists some $\epsilon > 0$ $p > 0$, such that

$$\mathbb{E}[|X_t - X_s|^p] \leqslant c \cdot |t - s|^{1+\epsilon}, \quad \text{for all } s, t \in \mathcal{D}$$

where c is a positive constant. Then for all $\alpha \in (0, \epsilon/p)$, the process is $\alpha-$Hölder continuous, that is there exists a random variable $K_\alpha < \infty$ such that

$$|X_t - X_s| \leqslant K_\alpha \cdot |t - s|^\alpha, \quad \text{for all } s, t \in \mathcal{D}.$$

Proof.

$$\mathbb{P}\left(|X_{k\cdot 2^{-n}} - X_{(K+1)\cdot 2^{-n}}| \geqslant 2^{-n\alpha}\right) \overset{\text{Markov + assumption}}{\leqslant} c \cdot 2^{-n\alpha} p \cdot 2^{-n(1+\epsilon)}.$$

Thus,

$$\mathbb{P}\left(\max_{0 \leqslant k \leqslant 2^n} |X_{k\cdot 2^{-n}} - X_{(K+1)\cdot 2^{-n}}| \geqslant 2^{-n\alpha}\right) \overset{\text{union bound}}{\leqslant} c \cdot 2^{n\alpha p n\epsilon}, \quad (\alpha \in (0, \frac{\epsilon}{p}).$$

By Borel-Cantelli,

$$\max_{0 \leqslant k \leqslant 2^n} |X_{k\cdot 2^{-n}} - X_{(K+1)\cdot 2^{-n}}| \leqslant 2^{-n\alpha}$$

for all $n \in \mathbb{N}$ sufficiently large. Thus,

$$\sup_{n \geqslant 0} \max_{0 \leqslant k \leqslant 2^n} \frac{|X_{k\cdot 2^{-n}} - X_{(K+1)\cdot 2^{-n}}|}{2^{-n\alpha}} \leqslant 2^{-n\alpha} \leqslant M(\omega) < \infty$$

a.s. For some random variable M.

Need to show: there exists some M' such that $|X_t - X_s| \leqslant M' \cdot |t - s|^\alpha$ for all $s, t \in \mathcal{D}$.

Let $s < t$, $s, t \in \mathcal{D}$ and let r be the unique integer such that $2^{-(r+1)} < t - s \leqslant 2^{-r}$. Then there exists some $k \in \mathbb{N}$ such that $s < k \cdot 2^{-(r+1)} < t$. Now, observe that $t - \alpha \leqslant 2^{-r}$ so

$$t - \alpha = \sum_{j=r+1}^{\infty} \frac{x_j}{2^j}, \quad x_j \in \{0, 1\}$$

and

$$\alpha - s = \sum_{j=r+1}^{\infty} \frac{y_j}{2^j}, \quad y_j \in \{0, 1\}.$$

Observe that $[s, t)$ is a disjoint union of dyadic intervals each of them having length 2^{-n} with $n \geqslant r + 1$ and each interval of length will appear at most twice. Thus, we get the bound

$$\begin{aligned} |X_t - X_s| &\leqslant \overbrace{\sum_{d,n} \underbrace{|X_d - X_{d+2^{-n}}|}_{\leqslant 2^{-n\alpha\cdot M}}}^{d,n \text{ is the endpoint of a dyadic interval in the decomposition of } [s,t) \text{ of length } 2^{-n}} \\ &\leqslant 2 \cdot M \cdot \sum_{n=r+1}^{\infty} 2^{-n\alpha} = \frac{2M \cdot 2^{-(r+1)\alpha}}{1 - 2^{-\alpha}} < \frac{2M}{1 - 2^{-\alpha}} |t - s|^\alpha. \end{aligned}$$

□

7 Weak Convergence

We fix $(\mathcal{M}, d)$ a metric space endowed with its Borel sigma algebra.

Definition 7.1. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $\mathcal{M}$. We say $(\mu_n)_{n\in\mathbb{N}}$ converges weakly to μ and write $\mu_n \implies \mu$ as $n \to \infty$ if

$$\mu_n(f) := \int_{\mathcal{M}} f(x)\mu_n(\mathrm{d}x) \overset{n\to\infty}{\longrightarrow} \int_{\mathcal{M}} f(x)\mu(\mathrm{d}x) := \mu(f)$$

for any f continuous and bounded.

Examples:

1. Let $x_n \to x$ as $n \to \infty$ in $(\mathcal{M}, d)$ then $\delta_{x_n} \overset{n\to\infty}{\longrightarrow} \delta_x$,since $\delta_{x_n}(f) = f(x_n) \overset{n\to\infty}{\longrightarrow} f(x) = \delta_x(f)$.
2. Let $\mathcal{M} = [0,1]$, with the Euclidean metric and its Borel sigma algebra. Let $\mu_n = \frac{1}{n}\sum\limits_{0\leqslant k\leqslant n} \delta_{k/n}$. Then μ_n converges weakly to the Lebesgue measure. Indeed, $\mu_n(f) = \frac{1}{n}f(k/n) \overset{n\to\infty}{\longrightarrow} \int f(x)\,\mathrm{d}x$, being Riemann sums.
3. $\mu_n = \delta_{\frac{1}{n}} \implies \delta_0$, as $n \to \infty$. Notice however that for $A = (0,1)$, $\mu_n(A) =$ for all $n \geqslant 0$ and so $\nu_n(A) \not\to \delta_0(A) = 0$.

Theorem 7.1. Let $(\mu_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $(\mathcal{M}, d)$. Then the following are equivalent:

1. $\mu_n \implies \mu$.
2. For all G open, $\liminf\limits_n \mu_n(G) \geqslant \mu(G)$.
3. For all A closed, $\limsup\limits_n \mu_n(A) \leqslant \mu(A)$.
4. For all A with $\mu(\partial A) = 0$, then $\mu_n(A) \to \mu(A)$.

Proof. $\underline{1 \implies 2:}$ Let G be open with $G^c \neq \varnothing$. Let $M > 0$ and set $f_M(x) = \mathbf{1}(Md(x, G^c)) \leqslant \mathbf{1}(x \in G)$. Observe that $f_M(x) \uparrow \mathbf{1}(x \in G)$ as $M \to \infty$, f_M is bounded and continuous for all M. So $\mu_n(f_M) \to \mu(f_M)$ as $n \to \infty$ for all M. Thus,

$$\liminf_n \mu_n(G) \geqslant \liminf_n \mu_n(f_M) = \mu(f_M) \overset{\text{monotone convergence}}{\to} \mu(G).$$

$\underline{2 \implies 3:}$ follows from the previous case by taking complements. $\underline{2,3 \implies 4}$: $0 = \mu(\partial A) = \mu(A \backslash \int A)$, hence $\mu(\overline{A}) = \mu(A) = \mu(\int A)$. $\underline{2:}$ $\liminf\limits_n \mu(\int A) \geqslant \mu(\int A) = \mu(A)$. $\underline{3:}$ $\limsup\limits_n \mu_n(\overline{A}) \leqslant \mu(\overline{A}) = \mu(A)$.

$\underline{4 \implies 1:}$ Need to show for any f continuous and bounded, $\mu_n(f) \to \mu(f)$. We can assume further that $f \geqslant 0$. Fix $K > \sup f$. Have,

$$\begin{aligned}\int_{\mathcal{M}} f(x)\mu_n(\mathrm{d}x) &= \int_{\mathcal{M}} \left(\int_0^K \mathbf{1}(t \leqslant f(x))\,\mathrm{d}t\right)\mu_n(\mathrm{d}x) \\ &\overset{\text{Fubini}}{=} \int_0^K \mu_n(f \geqslant t)\,\mathrm{d}t.\end{aligned}$$

It suffices to show $\mu_n(f \geqslant t) \to \mu(f \geqslant t)$ as $n \to \infty$. Since then we can conclude using dominated convergence. Thus it suffices to show that $\mu(\partial\{f \geqslant t\}) = 0$. Indeed,

$$\partial\{f \geqslant t\} \subset \{f = t\}.$$

since f is continuous and $\{f > t\}$ is open and $\subset \mathring{\{f \geqslant t\}}$. Also observe thath there exists an at most countable number of t such that $\mu(f = t) > 0$. Thus,

$$\{t : \mu(f = t) > 0\} = \bigcup_n \underbrace{\{t : \mu(\{f = t\}) \geqslant \frac{1}{n}\}}_{\# \leqslant n}.$$

Thus, $\partial\{f \geqslant t\}$ is countable and has Lebesgue measure zero. $\square$