

1 Conditional Expectation

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $(X_i)_{i \in I}$ be a collection of random variables defined on this space. Then we define $\sigma(X_i : i \in I) \subseteq \mathcal{F}$ to be the smallest σ -algebra such that all of the X_i are measurable, i.e

$$\sigma(X_i : i \in I) = \sigma(X_i^{-1}(B) : i \in I, B \in \mathcal{B}(\mathbb{R})).$$

Definition. If $B \in \mathcal{F}$ has $\mathbb{P}(B) > 0$ then we define

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

for any $A \in \mathcal{F}$. Furthermore, if X is an integrable random variable we define

$$\mathbb{E}[X|B] = \frac{\mathbb{E}[X \mathbb{1}(B)]}{\mathbb{P}(B)}.$$

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. We say a σ -algebra \mathcal{G} is countably generated if there exist $(B_i)_{i \in I}$ pairwise disjoint (with I countable) such that $\bigcup_{i \in I} B_i = \Omega$ and $\mathcal{G} = \sigma(B_i : i \in I)$.

Let X be an integrable random variable and \mathcal{G} a countably generated σ -algebra. We want to define $X' = \mathbb{E}[X|\mathcal{G}]$. So define

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

Or equivalently,

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

where we use the convention that $\mathbb{E}[X|B_i] = 0$ if $\mathbb{P}(B_i) = 0$. Then X' is indeed \mathcal{G} -measurable (note \mathcal{G} is the set of $\bigcup_{j \in J} B_j$ for $J \subseteq I$).

Note that for any $G \in \mathcal{G}$ we have $\mathbb{E}[X \mathbb{1}(G)] = \mathbb{E}[X' \mathbb{1}(G)]$. Also

$$\mathbb{E}[|X'|] \leq \mathbb{E} \left[\sum_{i \in I} \mathbb{E}[|X||B_i] \mathbb{1}(B_i) \right] = \sum_{i \in I} \mathbb{E}[|X||B_i] \mathbb{P}(B_i) = \mathbb{E}[|X|] < \infty$$

so X' is integrable.

Theorem (Monotone convergence theorem). *Let $(X_n)_{n \geq 1}$ be a sequence of non-negative random variables with $X_n \uparrow X$ as $n \rightarrow \infty$ almost-surely. Then $\mathbb{E}X_n \uparrow \mathbb{E}X$ as $n \rightarrow \infty$.*

Proof. See Part II Probability & Measure. □

Theorem (Dominated convergence theorem). *Let $(X_n)_{n \geq 1}$ be a sequence of random variables with $X_n \rightarrow X$ as $n \rightarrow \infty$ almost-surely and $|X_n| \leq Y$ almost-surely for some Y integrable. Then $\mathbb{E}X_n \rightarrow \mathbb{E}X$ as $n \rightarrow \infty$.*

Proof. See Part II Probability & Measure. □

Definition (L^p). Let $p \in [1, \infty]$ and f be a measurable function. Define the L^p -norm

$$\|f\|_p = (\mathbb{E}[|f|^p])^{1/p} \text{ for } p \in [1, \infty)$$

$$\|f\|_\infty = \inf\{\lambda : |f| \leq \lambda \text{ a.e.}\}.$$

Furthermore write $f \sim g$ if $f = g$ almost everywhere. Then define the L^p -space $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) = \{f : \|f\|_p < \infty\} / \sim$.

Theorem (\mathcal{L}^2 is a Hilbert space). *$\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ is a Hilbert space with inner product $\langle U, V \rangle = \mathbb{E}[UV]$. For a closed subspace \mathcal{H} , if $f \in \mathcal{L}^2$ there exists a unique $g \in \mathcal{H}$ with $\|f - g\|_2 = \inf\{\|f - h\|_2 : h \in \mathcal{H}\}$ and $\langle f - g, h \rangle = 0$ for all $h \in \mathcal{H}$. g is called the orthogonal projection of f on \mathcal{H} .*

Proof. See Part II Probability & Measure. □

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

- (a) Y is \mathcal{G} -measurable;
- (b) for all $A \in \mathcal{G}$, $\mathbb{E}[X \mathbb{1}(A)] = \mathbb{E}[Y \mathbb{1}(A)]$.

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

Remark. (b) could be replaced by $\mathbb{E}[XZ] = \mathbb{E}[YZ]$ for all Z bounded and \mathcal{G} -measurable.

Proof. First we show uniqueness. Suppose Y and Y' both satisfy (a) and (b) and let $A = \{Y > Y'\} \in \mathcal{G}$. Then

$$\mathbb{E}[Y \mathbb{1}(A)] = \mathbb{E}[Y' \mathbb{1}(A)] \Rightarrow \mathbb{E}[(Y - Y') \mathbb{1}(A)] = 0 \Rightarrow \mathbb{P}(Y > Y') = 0 \Rightarrow Y \leq Y' \text{ a.s.}$$

and similarly $Y \geq Y'$ a.s.

Now we show existence. First assume $X \in \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$. Then $\mathcal{L}^2(\Omega, \mathcal{G}, \mathbb{P})$ is a closed subspace of $\mathcal{L}^2(\mathcal{F})$. Hence

$$\mathcal{L}^2(\mathcal{F}) = \mathcal{L}^2(\mathcal{G}) \oplus \mathcal{L}^2(\mathcal{G})^\perp$$

so we can write $X = Y + Z$ for $Y \in \mathcal{L}^2(\mathcal{G})$ and $Z \in \mathcal{L}^2(\mathcal{G})^\perp$. Define $\mathbb{E}[X|\mathcal{G}] = Y$, so Y is \mathcal{G} -measurable and for all $A \in \mathcal{G}$

$$\mathbb{E}[X \mathbb{1}(A)] = \mathbb{E}[Y \mathbb{1}(A)] + \underbrace{\mathbb{E}[Z \mathbb{1}(A)]}_{=0} = \mathbb{E}[Y \mathbb{1}(A)].$$

We claim that if $X \geq 0$ almost-surely, then $Y \geq 0$ almost-surely. Indeed, let $A = \{Y < 0\} \in \mathcal{G}$ so $0 \leq \mathbb{E}[X \mathbb{1}(Y < 0)] = \mathbb{E}[Y \mathbb{1}(Y < 0)] \leq 0$ which implies $\mathbb{P}(Y < 0) = 0$.

Assume now that $X \geq 0$ almost-surely. Define $X_n = X \wedge n \leq n$, so $X_n \in \mathcal{L}^2$ for all n . Let $Y_n = \mathbb{E}[X_n|\mathcal{G}]$. Then X_n is an increasing sequence and by the above claim, Y_n is also an increasing sequence almost-surely. Define $Y = \limsup_{n \rightarrow \infty} Y_n$, so Y is \mathcal{G} -measurable. Also $Y = \uparrow \lim_{n \rightarrow \infty} Y_n$ almost-surely. For any $A \in \mathcal{G}$ we have

$$\mathbb{E}[X \mathbb{1}(A)] = \lim_{n \rightarrow \infty} \mathbb{E}[X_n \mathbb{1}(A)] = \lim_{n \rightarrow \infty} \mathbb{E}[Y_n \mathbb{1}(A)] = \mathbb{E}[Y \mathbb{1}(A)]$$

by the Monotone Convergence Theorem.

Finally, for general X write $X = X^+ - X^-$ and define $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X^+|\mathcal{G}] - \mathbb{E}[X^-|\mathcal{G}]$. \square

Remark. From the last proof we can see that we can define $\mathbb{E}[X|\mathcal{G}]$ for $X \geq 0$ without assuming integrability of X . It satisfies all the conditions apart from integrability.

Definition. Let $(\mathcal{G}_n)_{n \geq 1}$ be sub σ -algebras of \mathcal{F} . We call them *independent* if whenever $G_i \in \mathcal{G}_i$ and $i_1 < i_2 < \dots < i_k$ we have

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

For a random variable X and a σ -algebra \mathcal{G} , we say they are *independent* if $\sigma(X)$ is independent of \mathcal{G} .

Properties of conditional expectation

Let $X, Y \in \mathcal{L}^1$, $\mathcal{G} \subseteq \mathcal{F}$ a sub σ -algebra. Then

1. $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$ (take $A = \Omega$);
2. If X is \mathcal{G} -measurable then $\mathbb{E}[X|\mathcal{G}] = X$ almost-surely (X clearly satisfies the conditions);
3. If X is independent of \mathcal{G} , then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ almost-surely;
4. If $X \geq 0$ almost-surely then $\mathbb{E}[X|\mathcal{G}] \geq 0$ almost-surely;
5. For $\alpha, \beta \in \mathbb{R}$, $\mathbb{E}[\alpha X + \beta Y|\mathcal{G}] = \alpha \mathbb{E}[X|\mathcal{G}] + \beta \mathbb{E}[Y|\mathcal{G}]$ almost-surely;
6. $|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X||\mathcal{G}]$ almost-surely.

Recall:

Theorem (Fatou's Lemma). *If $X_n \geq 0$ for all n almost-surely, then*

$$\mathbb{E}[\liminf_{n \geq 1} X_n] \leq \liminf_{n \geq 1} \mathbb{E}X_n.$$

Proof. See Part II Probability & Measure. □

Theorem (Jensen's Inequality). *If X is integrable, $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is convex, then*

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

We consider any analogues of our convergence theorems for conditional expectation.

Theorem (Conditional Monotone Convergence Theorem). *Suppose $X_n \geq 0$ for all n and $X_n \uparrow X$ almost-surely as $n \rightarrow \infty$. Let \mathcal{G} be a sub σ -algebra of \mathcal{F} . Then $\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}]$ almost-surely.*

Remark. Note that $\mathbb{E}[X_n|\mathcal{G}] \uparrow \mathbb{E}[X|\mathcal{G}]$ in the almost-sure sense, as these are random variables.

Proof. Let $Y_n = \mathbb{E}[X_n|\mathcal{G}]$ almost-surely. Then Y_n is increasing. Set $Y = \limsup_{n \geq 1} Y_n$. Since Y_n is \mathcal{G} -measurable, Y is \mathcal{G} -measurable. Also $Y = \uparrow \lim_{n \geq 1} Y_n$ almost-surely. We need to show $\mathbb{E}[Y\mathbb{1}(A)] = \mathbb{E}[X\mathbb{1}(A)]$ for all $A \in \mathcal{G}$. This follows from the usual Monotone Convergence Theorem as

$$\mathbb{E}[Y\mathbb{1}(A)] = \lim_{n \geq 1} \mathbb{E}[Y_n\mathbb{1}(A)] = \lim_{n \geq 1} \mathbb{E}[X_n\mathbb{1}(A)] = \mathbb{E}[X\mathbb{1}(A)].$$

□

Theorem (Conditional Fatou's Lemma). *Let $(X_n)_{n \geq 1}$ be a non-negative sequence of random variables. Then*

$$\mathbb{E}[\liminf_{n \rightarrow \infty} X_n | \mathcal{G}] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}] \text{ almost-surely.}$$

Proof. Note that $\inf_{k \geq n} X_k \uparrow \liminf_{n \rightarrow \infty} X_n$ so by the conditional MCT

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

We also have

$$\mathbb{E}[\inf_{k \geq n} X_k | \mathcal{G}] \leq \mathbb{E}[X_k | \mathcal{G}] \quad \forall k \geq n \text{ almost-surely.}$$

Which implies

$$\mathbb{E}[\inf_{k \geq n} X_k | \mathcal{G}] \leq \inf_{k \geq n} \mathbb{E}[X_k | \mathcal{G}] \quad \forall k \geq n \text{ almost-surely}$$

since k takes countable values (intersection of countable sets of full measure also has full measure). Now taking limits as $n \rightarrow \infty$ we are done. □

Theorem (Conditional Dominated Convergence Theorem). *Suppose $X_n \rightarrow X$ almost-surely, $|X_n| \leq Y$ almost-surely with Y integrable. Then $\mathbb{E}[X_n | \mathcal{G}] \rightarrow \mathbb{E}[X | \mathcal{G}]$ almost-surely.*

Proof. We apply the Conditional Fatou's Lemma. Indeed $-Y \leq X_n \leq Y$ so $X_n + Y \geq 0$ and $Y - X_n \geq 0$ for all n . By Conditional Fatou's Lemma

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

and

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

Hence $\limsup_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}] \leq \mathbb{E}[X | \mathcal{G}]$ and $\liminf_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}] \geq \mathbb{E}[X | \mathcal{G}]$ almost-surely. □

Theorem (Conditional Jensen's Inequality). *Let X be integrable, $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ a convex function such that $\varphi(X)$ is integrable or $\varphi(X) \geq 0$. Then $\mathbb{E}[\varphi(X) | \mathcal{G}] \geq \varphi(\mathbb{E}[X | \mathcal{G}])$ almost-surely.*

Proof. We claim that $\varphi(x) = \sup_{i \in \mathbb{N}}(a_i x + b_i)$, $a_i, b_i \in \mathbb{R}$.

Then $\varphi(X) = \sup_{i \in \mathbb{N}}(a_i X + b_i)$. So

$$\mathbb{E}[\varphi(X)|\mathcal{G}] \geq \sup_{n \geq 1} (a_i \mathbb{E}[X|\mathcal{G}] + b_i) \quad \forall i \in \mathbb{N} \text{ almost-surely.}$$

□

Note. We need the supremum in the claim to be over a countable set so we can preserve the almost-sure property of an inequality.

Corollary. For all $p \in [1, \infty)$ we have

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

Proof. Apply conditional Jensen ($x \mapsto x^p$ is convex). □

Theorem (Tower property). Let X be integrable and $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ sub σ -algebras. Then

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] \text{ almost-surely.}$$

Proof. $\mathbb{E}[X|\mathcal{H}]$ is certainly \mathcal{H} -measurable so it remains to check

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}(A)] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]\mathbb{1}(A)] \quad \forall A \in \mathcal{H}.$$

But since $A \in \mathcal{G}$ whenever $A \in \mathcal{H}$ we have

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

□

Proposition. Let $X \in \mathcal{L}^1$, $\mathcal{G} \subseteq \mathcal{F}$ a sub σ -algebra, Y bounded and \mathcal{G} -measurable. Then

$$\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}] \text{ almost-surely.}$$

Proof. $Y\mathbb{E}[X|\mathcal{G}]$ is certainly \mathcal{G} -measurable. Also for any $A \in \mathcal{G}$

$$\mathbb{E}[XY\mathbb{1}(A)] = \mathbb{E}[X \underbrace{(Y\mathbb{1}(A))}_{\substack{\text{bounded,} \\ \mathcal{G}\text{-measurable}}}] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}](Y\mathbb{1}(A))].$$

□

Definition. Let \mathcal{A} be a collection of sets. It is called a π -system if whenever $A, B \in \mathcal{A}$ we have $A \cap B \in \mathcal{A}$.

Recall

Theorem (Uniqueness of extension). *Let (E, \mathcal{E}) be a measurable space and let \mathcal{A} be a π -system generating \mathcal{E} . Let μ, ν be two measures on (E, \mathcal{E}) with $\mu(E) = \nu(E) < \infty$. If $\mu = \nu$ on \mathcal{A} , then $\mu = \nu$ on \mathcal{E} .*

Proof. See Part II Probability & Measure. \square

Theorem. *Let $X \in \mathcal{L}^1$, $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ sub σ -algebras. Assume $\sigma(X, \mathcal{G})$ is independent of \mathcal{H} . Then*

$$\mathbb{E}[X|\sigma(\mathcal{G}, \mathcal{H})] = \mathbb{E}[X|\mathcal{G}] \text{ almost-surely.}$$

Proof. We need to show $\mathbb{E}[X\mathbb{1}(F)] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}(F)]$ for all $F \in \sigma(\mathcal{G}, \mathcal{H})$. Define $\mathcal{A} = \{A \cap B : A \in \mathcal{G}, B \in \mathcal{H}\}$. This is a π -system generating $\sigma(\mathcal{G}, \mathcal{H})$. If $F = A \cap B$, $A \in \mathcal{G}, B \in \mathcal{H}$ then

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

Assume $X \geq 0$. Define $\mu(F) = \mathbb{E}[X\mathbb{1}(F)]$ and $\nu(F) = \mathbb{E}[\mathbb{E}[X|\mathcal{G}]\mathbb{1}(F)]$ for $F \in \sigma(\mathcal{G}, \mathcal{H})$. Then $\mu = \nu$ on \mathcal{A} by the above and $\mu(\Omega) = \nu(\Omega) < \infty$. Therefore $\mu = \nu$ on $\sigma(\mathcal{G}, \mathcal{H})$. \square

Definition. We say $(X_1, \dots, X_n) \in \mathbb{R}^n$ has the *Gaussian distribution* iff for all $a_1, \dots, a_n \in \mathbb{R}$

$$a_1X_1 + \dots + a_nX_n$$

has the Gaussian distribution in \mathbb{R} .

A process $(X_t)_{t \geq 0}$ is called a *Gaussian process* if $\forall t_1 < t_2 < \dots < t_n$, the vector $(X_{t_1}, \dots, X_{t_n})$ is a Gaussian random vector.

Example. Let (X, Y) be a Gaussian vector in \mathbb{R}^2 . We want to compute $\mathbb{E}[X|Y] = \mathbb{E}[X|\sigma(Y)]$. Let $X' = \mathbb{E}[X|Y]$. Since X' is $\sigma(Y)$ -measurable it follows X' is a measurable function of Y . So are looking for f Borel such that $\mathbb{E}[X|Y] = f(Y)$ almost-surely. Let $f(y) = ay + b$ for some $a, b \in \mathbb{R}$ to be determined.

Since $\mathbb{E}[X'] = \mathbb{E}[X]$ we have $a\mathbb{E}Y + b = \mathbb{E}X$. Also

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

so we have determined a, b . We need to check that for any Z bounded and $\sigma(Y)$ -measurable we have $\mathbb{E}[(X - X')Z] = 0$. Write $Z = g(Y)$ and note $\text{Cov}(X - X', Y) = 0$, implying $X - X'$ is independent of Y . Therefore $\mathbb{E}[(X - X')g(Y)] = \mathbb{E}[X - X']\mathbb{E}[g(Y)] = 0$.

Example. Let (X, Y) be a random vector in \mathbb{R}^2 with joint density function $f_{X,Y}(x, y)$. Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a Borel function such that $h(X)$ is integrable. We want to compute $\mathbb{E}[h(X)|Y]$. Note

$$\mathbb{E}[h(X)g(Y)] = \int_{\mathbb{R}^2} h(x)g(y)f_{X,Y}(x, y)dx dy$$

and write

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

for the density of Y . So (using the convention $0/0 = 0$)

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} h(x) \frac{f_{X,Y}(x, y)}{f_Y(y)} dx \right) g(y) f_Y(y) dy$$

define

$$\varphi(y) = \begin{cases} \int_{\mathbb{R}} h(x) \frac{f_{X,Y}(x, y)}{f_Y(y)} dx & \text{if } f_Y(y) > 0 \\ 0 & \text{otherwise} \end{cases}.$$

Then $\mathbb{E}[h(X)|Y] = \varphi(Y)$ almost-surely.

2 Martingales

2.1 Discrete-time Martingales

Definition. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A *filtration* is a sequence of increasing sub σ -algebras of \mathcal{F} , $(\mathcal{F}_n)_{n \geq 0}$, $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$. We call $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ a *filtered probability space*.

If $X = (X_n)_{n \geq 0}$ is a sequence of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, define $\mathcal{F}_n^X = \sigma(X_k : k \leq n)$, the *natural filtration* associated with X . We say X is *adapted* to a filtration (\mathcal{F}_n) if X_n is \mathcal{F}_n -measurable for all n . X is *integrable* if X_n is integrable for all n .

Definition. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ be a filtered probability space. We say an integrable adapted process $X = (X_n)_{n \geq 0}$ is called a

- *martingale* if

$$\mathbb{E}[X_n|\mathcal{F}_m] = X_m \text{ almost-surely } \forall n \geq m.$$

- *super-martingale* if

$$\mathbb{E}[X_n|\mathcal{F}_m] \leq X_m \text{ almost-surely } \forall n \geq m.$$

- *sub-martingale* if

$$\mathbb{E}[X_n|\mathcal{F}_m] \geq X_m \text{ almost-surely } \forall n \geq m.$$

Remark. If X is a martingale with respect to (\mathcal{F}_n) , then it is also a martingale with respect to the natural filtration (\mathcal{F}_n^X) .

Example. Let (ξ_i) be a sequence of iid random variables with $\mathbb{E}[\xi_1] = 0$. Let $X_n = \xi_1 + \dots + \xi_n$, $X_0 = 0$. This is a martingale. We have

$$\mathbb{E}[X_n|\mathcal{F}_{n-1}] = \xi_1 + \dots + \xi_{n-1} + \mathbb{E}[\xi_n|\mathcal{F}_{n-1}] = \xi_1 + \dots + \xi_{n-1}$$

by independence.

Example. Let (ξ_i) be a sequence of iid random variables with $\mathbb{E}[\xi_1] = 1$. Let $X_n = \prod_{i=1}^n \xi_i$, $X_0 = 1$. This is a martingale.

Definition. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$ be a filtered probability space. A *stopping time* T is a random variable $T : \Omega \rightarrow \mathbb{Z}_+ \cup \{\infty\}$ such that $\{T \leq n\} \in \mathcal{F}_n$ for all n .

Note. T being a stopping time is equivalent to $\{T = n\} \in \mathcal{F}_n$ for all n .

Examples.

- Constant times are trivial stopping times;
- Suppose $(X_n)_{n \geq 0}$ is an adapted process taking values in \mathbb{R} . For $A \in \mathcal{B}$ define $T_A = \inf\{n \geq 0 : X_n \in A\}$ (with the convention that $\inf \emptyset = \infty$). Then $\{T_A \leq n\} = \bigcup_{k \leq n} \{X_k \in A\} \in \mathcal{F}_n$, so T_A is a stopping time;
- In the setting above, let $L_A = \sup\{n \geq 0 : X_n \in A\}$. This is in general not a stopping time.

Proposition. Let $S, T, (T_n)$ be stopping times. Then $S \wedge T$, $S \vee T$, $\inf T_n$, $\sup T_n$, $\liminf T_n$ and $\limsup T_n$ are also stopping times.

Proof. Follows directly from the definition. \square

Definition. If T is a stopping time, we define

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t, \forall t\}.$$

If $(X_n)_{n \geq 0}$ is a process, write $X_T(\omega) = X_{T(\omega)}(\omega)$ whenever $T(\omega) < \infty$. We define the *stopped process* $X_t^T = X_{T \wedge t}$.

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

1. If $S \leq T$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$;
2. $X_T \mathbb{1}(T < \infty)$ is \mathcal{F}_T -measurable;
3. X^T is adapted;
4. If X is integrable, then X^T is also integrable.

Proof.

1. Immediate from the definition;
2. Let $A \in \mathcal{B}(\mathbb{R})$. We need to show $\{X_T \mathbb{1}(T < \infty) \in A\} \in \mathcal{F}_T$. Note that

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

3. $X_t^T = X_{T \wedge t}$ is $\mathcal{F}_{T \wedge t}$ -measurable so \mathcal{F}_t -measurable by (1).

4. We have

$$\begin{aligned}\mathbb{E}[|X_t^T|] &= \mathbb{E}[|X_{T \wedge t}|] = \sum_{s=0}^{t-1} \mathbb{E}[|X_s| \mathbb{1}(T = s)] + \mathbb{E}[|X_t| \mathbb{1}(T \geq t)] \\ &\leq \sum_{s=0}^t \mathbb{E}[|X_s|] < \infty.\end{aligned}$$

□

Theorem (Optional Stopping Theorem). *Let (X_n) be a martingale.*

1. *If T is a stopping time, then X^T is also a martingale. In particular $\mathbb{E}[X_{T \wedge t}] = \mathbb{E}[X_0]$ for all t ;*
2. *If $S \leq T$ are bounded stopping times then $\mathbb{E}[X_T | \mathcal{F}_S] = X_S$ almost-surely, and $\mathbb{E}[X_T] = \mathbb{E}[X_S]$;*
3. *If there exists an integrable random variable Y such that $|X_n| \leq Y$ for all n , and T is finite almost-surely then $\mathbb{E}[X_T] = \mathbb{E}[X_0]$;*
4. *If there exists $M > 0$ such that $|X_{n+1} - X_n| \leq M$ for all n , and T is a stopping time with $\mathbb{E}T < \infty$, then $\mathbb{E}[X_T] = \mathbb{E}[X_0]$.*

Proof.

1. We need to show that for all t we have

$$\mathbb{E}[X_{T \wedge t} | \mathcal{F}_{t-1}] = X_{T \wedge (t-1)}$$

almost-surely. Indeed

$$\begin{aligned}\mathbb{E}[X_{T \wedge t} | \mathcal{F}_{t-1}] &= \mathbb{E}\left[\sum_{s=0}^{t-1} X_s \mathbb{1}(T = s) | \mathcal{F}_{t-1}\right] + \mathbb{E}[X_t \mathbb{1}(T \geq t) | \mathcal{F}_{t-1}] \\ &= \sum_{s=0}^{t-1} X_s \mathbb{1}(T = s) + \mathbb{1}(T \geq t) X_{t-1} \\ &= X_{T \wedge (t-1)}\end{aligned}$$

using the fact that $\mathbb{1}(T \geq t)$ is \mathcal{F}_{t-1} -measurable;

2. Suppose $S \leq T \leq n$ and let $A \in \mathcal{F}_S$. We need to show $\mathbb{E}[X_T \mathbb{1}(A)] = \mathbb{E}[X_S \mathbb{1}(A)]$. Note

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

Hence

$$\begin{aligned}\mathbb{E}[X_T \mathbb{1}(A)] &= \mathbb{E}[X_S \mathbb{1}(A)] + \sum_{k=0}^n \mathbb{E}[(X_{k+1} - X_k) \underbrace{\mathbb{1}(S \leq k < T) \mathbb{1}(A)}_{\in \mathcal{F}_k}] \\ &= \mathbb{E}[X_S \mathbb{1}(A)]\end{aligned}$$

since $\mathbb{E}[X_{k+1} | \mathcal{F}_k] = X_k$ almost-surely. Taking expectations gives $\mathbb{E}[X_T] = \mathbb{E}[X_S]$;

3. Example Sheet;
4. Example Sheet.

□

Note. Analogous results follow if (X_n) is instead a sub/super-martingale.

Corollary. If X is a positive super-martingale, T is a stopping time, $T < \infty$ almost-surely, then $\mathbb{E}[X_T] \leq \mathbb{E}[X_0]$.

Proof. Fatou's lemma gives $\mathbb{E}[\liminf_t X_{T \wedge t}] \leq \liminf_t \mathbb{E}[X_{T \wedge t}] \leq \mathbb{E}[X_0]$. □

Example. Let $(\xi_i)_{i \geq 0}$ be iid with $\mathbb{P}(\xi_0 = 1) = \mathbb{P}(\xi_0 = -1) = 1/2$. Define $X_0 = 0$ and $X_n = \sum_{i=1}^n \xi_i$ for $n \geq 1$. Then $(X_n)_{n \geq 0}$ is a martingale. Define $T = \inf\{n \geq 0 : X_n = 1\}$. Then $\mathbb{P}(T < \infty) = 1$ and for all t we have $\mathbb{E}[X_{T \wedge t}] = 0$, while $\mathbb{E}[X_T] = 1$. Hence (4) from the previous theorem tells us $\mathbb{E}T = \infty$.

Example. Consider a SRW on \mathbb{Z} , $X_0 = 0$, $X_n = \sum_{i=1}^n \xi_i$ with $(\xi_i)_{i \geq 1}$ iid taking values ± 1 with equal probability. Define $T_c = \inf\{n \geq 0 : X_n = c\}$ and set $T = T_{-a} \wedge T_b$. What is $\mathbb{P}(T_{-a} < T_b)$?

We have that $X_n^T = X_{T \wedge n}$ is a martingale by the optional stopping theorem. Furthermore $|X_{n+1} - X_n| = 1$ for all n . Need to check $\mathbb{E}[T] < \infty$: consider blocks

- ξ_1, \dots, ξ_{a+b}
- $\xi_{a+b+1}, \dots, \xi_{2(a+b)}$
- $\xi_{2(a+b)+1}, \dots, \xi_{3(a+b)}$
- \vdots

note that the probability the ξ_i in one of these blocks are all equal to either 1 or -1 is $2 \cdot 2^{-(a+b)}$. Hence $T \leq (a+b)\text{Geo}(2 \cdot 2^{-(a+b)})$ and $\mathbb{E}T \leq (a+b)2^{a+b-1} < \infty$.

So applying the optional stopping theorem to T we have $\mathbb{E}[X_T] = \mathbb{E}[X_0] = 0$. Hence $-a\mathbb{P}(T_{-a} < T_b) + b\mathbb{P}(T_b < T_{-a}) = 0$ and $\mathbb{P}(T_{-a} < T_b) + \mathbb{P}(T_b < T_{-a})$, which gives $\mathbb{P}(T_{-a} < T_b) = \frac{b}{a+b}$.

Martingale convergence theorem

Theorem (Almost-sure martingale convergence theorem). *Let X be a supermartingale bounded in \mathcal{L}^1 , i.e $\sup_{n \geq 0} \mathbb{E}|X_n| < \infty$. Then there exists a random variable $X_\infty \in \mathcal{L}^1(\mathcal{F}_\infty)$ where $\mathcal{F}_\infty = \sigma(\mathcal{F}_n : n \geq 0)$ such that $X_n \rightarrow X_\infty$ almost-surely as $n \rightarrow \infty$.*

Before we can prove this we will need some preliminary results.

Doob's upcrossing inequality

For a real sequence $(x_n)_{n \geq 0}$, for an interval $[a, b]$ we want to count the number of times (x_n) crosses below a or above b . Define $T_0(x) = 0$ and define for $k \geq 0$

$$S_{k+1}(x) = \inf\{n \geq T_k(x) : x_n \leq a\} \text{ the } (k+1)\text{st downcrossing}$$

$$T_{k+1}(x) = \inf\{n \geq S_{k+1}(x) : x_n \geq b\} \text{ the } (k+1)\text{st upcrossing.}$$

Also let $N_n([a, b], x) = \sup\{k \geq 0 : T_k(x) \leq n\}$, the number of up crossings up to time N . Then as $n \rightarrow \infty$, $N_n([a, b], x) \uparrow N([a, b], x) = \sup\{k \geq 0 : T_k(x) < \infty\}$.

Lemma. *Let $x = (x_n)_{n \geq 0}$ be a real sequence. Then x converges in $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ if and only if for all $a < b$, $a, b \in \mathbb{Q}$ we have $N([a, b], x) < \infty$.*

Proof. If x converges then suppose there is $a < b$ with $N([a, b], x) = \infty$. Then

$$\liminf x_n \leq a < b \leq \limsup x_n$$

a contradiction.

Conversely, if x doesn't converge we have $\liminf x_n < \limsup x_n$ so there are $a < b$ (with $a, b \in \mathbb{Q}$) with $\liminf x_n < a < b < \limsup x_n$ and hence $N([a, b], x) = \infty$. \square

Now we can prove

Theorem (Doob's upcrossing inequality). *Let X be a supermartingale and $a < b$. Then for all n ,*

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

Proof. We have $(T_k)_{k \geq 0}, (S_k)_{k \geq 0}$ stopping times. Then

$$\sum_{k=1}^n (X_{T_k \wedge n} - X_{S_k \wedge n}) = \sum_{k=1}^{N_n([a, b], X)} \underbrace{(X_{T_k} - X_{S_k})}_{\geq b-a} + \underbrace{(X_n - X_{S_{N_n+1}})\mathbb{1}(S_{N_n+1} \leq n)}_{\geq (X_n - a) \vee 0 = -(X_n - a)^-}.$$

Note $T_k \wedge n, S_k \wedge n$ are stopping times with $T_k \wedge n \geq S_k \wedge n$. Then by the optional stopping theorem $\mathbb{E}[X_{T_k \wedge n}] \leq \mathbb{E}[X_{S_k \wedge n}]$. So taking expectations we have

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

\square

Now we are ready to prove

Theorem (Almost-sure martingale convergence theorem). *Let X be a supermartingale bounded in \mathcal{L}^1 , i.e. $\sup_{n \geq 0} \mathbb{E}[X_n] < \infty$. Then there exists a random variable $X_\infty \in \mathcal{L}^1(\mathcal{F}_\infty)$ where $\mathcal{F}_\infty = \sigma(\mathcal{F}_n : n \geq 0)$ such that $X_n \rightarrow X_\infty$ almost-surely as $n \rightarrow \infty$.*

Proof. Let $a, b \in \mathbb{Q}$ be such that $a < b$. Then

$$\begin{aligned} \mathbb{E}[N_n([a, b], X)] &\leq (b - a)^{-1} \mathbb{E}[(X_n - a)^-] \\ &\leq (b - a)^{-1} \mathbb{E}[|X_n| + a] \\ &\leq (b - a)^{-1} \left(\sup_{n \geq 0} \mathbb{E}[|X_n|] + 1 \right). \end{aligned}$$

We know $N_n([a, b], X) \uparrow N([a, b], X)$ as $n \rightarrow \infty$, so by monotone convergence, $\mathbb{E}[N([a, b], X)] < \infty$. Set

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

so $\mathbb{P}(\Omega_0) = 1$ as the intersection of almost-sure events. On Ω_0 , X converges by a previous lemma. Set

$$X_\infty = \begin{cases} \lim_{n \rightarrow \infty} X_n & \text{on } \Omega_0 \\ 0 & \text{on } \Omega \setminus \Omega_0 \end{cases}.$$

So X_∞ is \mathcal{F}_∞ -measurable, and $X_n \rightarrow X_\infty$ almost surely. Also

$$\mathbb{E}[|X_\infty|] = \mathbb{E}[\liminf_n |X_n|] \leq \liminf_n \mathbb{E}[|X_n|] < \infty$$

by Fatou. □

Corollary. *Let X be a positive super-martingale. Then X converges almost-surely.*

Proof. $\mathbb{E}[|X_n|] = \mathbb{E}[X_n] \leq \mathbb{E}[X_0]$. So apply the previous. □

Doob's inequalities

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

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

Proof. Let $T = \inf\{k \geq 0 : X_k \geq \lambda\}$. Then T is a stopping time and $\{X_n^* \geq \lambda\} = \{T \leq n\}$. By the optional stopping theorem we have $\mathbb{E}[X_{T \wedge n}] \leq \mathbb{E}[X_n]$ and note

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

Therefore

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

□

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

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

Proof. By Jensen's inequality it is enough to prove for X a non-negative submartingale. Let $k > 0$ and note

$$(y \wedge k)^p = \int_0^k p x^{p-1} \mathbb{1}(y \geq x) dx$$

so

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

Therefore $\|X_n^* \wedge k\|_p \leq \frac{p}{p-1} \|X_n\|_p$. Taking $k \rightarrow \infty$ gives the result by monotone convergence. \square

Theorem (\mathcal{L}^p -convergence theorems). *Let X be a martingale, $p > 1$. The following are equivalent*

1. X is bounded in \mathcal{L}^p , i.e. $\sup_{n \geq 0} \|X_n\|_p < \infty$.
2. X converges almost-surely and in \mathcal{L}^p to a limit $X_\infty \in \mathcal{L}^p$.
3. There exists $Z \in \mathcal{L}^p$ such that $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ almost-surely.

Proof. (1 \Rightarrow 2) If X is bounded in \mathcal{L}^p then it is bounded in \mathcal{L}^1 . Hence there exists X_∞ such that $X_n \rightarrow X_\infty$ almost-surely as $n \rightarrow \infty$. Furthermore

$$\mathbb{E}|X_\infty|^p = \mathbb{E}[\liminf_n |X_n|^p] \leq \liminf_n \mathbb{E}[|X_n|^p] < \infty \quad (\text{Fatou})$$

so $X_\infty \in \mathcal{L}^p$. Define $X_n^* = \sup_{0 \leq k \leq n} |X_k|$, $X_\infty^* = \sup_{k \geq 0} |X_k|$. Then $|X_n - X_\infty| \leq 2X_\infty^*$ for all n . By dominated convergence it is enough to show $X_\infty^* \in \mathcal{L}^p$. Doob's \mathcal{L}^p inequality gives

$$\|X_n^*\|_p \leq \frac{p}{p-1} \|X_n\|_p \leq \frac{p}{p-1} \sup_{n \geq 0} \|X_n\|_p.$$

So by monotone convergence $\|X_\infty^*\|_p < \infty$.

(2 \Rightarrow 3) Set $Z = X_\infty$. Need to show $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$ almost-surely. We have for $m \geq n$ that

$$\begin{aligned} \|X_n - \mathbb{E}[X_\infty|\mathcal{F}_n]\|_p &= \|\mathbb{E}[X_m|\mathcal{F}_n] - \mathbb{E}[X_\infty|\mathcal{F}_n]\|_p \\ &\leq \|X_m - X_\infty\|_p \quad (\text{conditional Jensen}) \\ &\rightarrow 0 \text{ as } m \rightarrow \infty. \end{aligned}$$

(3 \Rightarrow 1) By conditional Jensen. \square

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

Corollary. *If $Z \in \mathcal{L}^p$, $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ almost-surely then $X_n \rightarrow \mathbb{E}[Z|\mathcal{F}_\infty]$ almost-surely and in \mathcal{L}^p , where $\mathcal{F}_\infty = \sigma(\mathcal{F}_n : n \geq 0)$.*

Proof. By the theorem we have $X_n \rightarrow X_\infty$ almost-surely and in \mathcal{L}^p . We need to show $X_\infty = \mathbb{E}[Z|\mathcal{F}_\infty]$ almost-surely.

- X_∞ is certainly \mathcal{F}_∞ -measurable.

- So we check that for all $A \in \mathcal{F}_\infty$ we have $\mathbb{E}[Z\mathbb{1}(A)] = \mathbb{E}[X_\infty\mathbb{1}(A)]$. Note that $\bigcup_{n \geq 0} \mathcal{F}_n$ is a π -system generating \mathcal{F}_∞ so it suffices to check for A in this π -system. Indeed for such A , there exists $N \geq 0$ such that $A \in \mathcal{F}_N$. Now let $n \geq N$ so

$$\begin{aligned}\mathbb{E}[Z\mathbb{1}(A)] &= \mathbb{E}[\mathbb{E}[Z|\mathcal{F}_N]\mathbb{1}(A)] \\ &= \mathbb{E}[X_N\mathbb{1}(A)] \rightarrow \mathbb{E}[X_\infty\mathbb{1}(A)] \text{ as } n \rightarrow \infty.\end{aligned}$$

□

Uniform integrability

Recall that a collection $(X_i)_{i \in I}$ of random variables is said to be *uniformly integrable* if

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

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

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

Remark. If $(X_i)_{i \in I}$ is bounded in \mathcal{L}^p for $p > 1$ then it is uniformly integrable.

Lemma. Let $(X_n)_{n \geq 1}, X$ be in \mathcal{L}^1 and $X_n \rightarrow X$ almost-surely as $n \rightarrow \infty$. Then $X_n \rightarrow X$ in \mathcal{L}^1 if and only if $(X_n)_{n \geq 1}$ is uniformly integrable.

Proof. See Part II Probability & Measure. □

Theorem. Let $X \in \mathcal{L}^1$. The family $\{\mathbb{E}[X|\mathcal{G}] : \mathcal{G} \subseteq \mathcal{F} \text{ a sub-}\sigma\text{-algebra}\}$ is uniformly integrable.

Proof. We need to show that for all $\varepsilon > 0$, there exists λ large enough such that for any sub- σ -algebra $\mathcal{G} \subseteq \mathcal{F}$ we have

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}] \mathbb{1}(|\mathbb{E}[X|\mathcal{G}]| > \lambda)] < \varepsilon.$$

Indeed

$$\begin{aligned} \mathbb{E}[\mathbb{E}[X|\mathcal{G}] \mathbb{1}(|\mathbb{E}[X|\mathcal{G}]| > \lambda)] &\leq \mathbb{E}[\underbrace{\mathbb{E}[|X||\mathcal{G}]}_{\mathcal{G}\text{-meas}} \mathbb{1}(|\mathbb{E}[X|\mathcal{G}]| > \lambda)] \\ &= \mathbb{E}[|X| \mathbb{1}(|\mathbb{E}[X|\mathcal{G}]| > \lambda)]. \end{aligned}$$

Since $X \in \mathcal{L}^1$, there exists $\delta > 0$ such that if $A \in \mathcal{F}$ has $\mathbb{P}(A) < \delta$, then $\mathbb{E}[|X| \mathbb{1}(A)] < \varepsilon$. Then

$$\mathbb{P}(|\mathbb{E}[X|\mathcal{G}]| > \lambda) \leq \frac{\mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|]}{\lambda} \leq \frac{\mathbb{E}|X|}{\lambda}.$$

So taking $\lambda = \mathbb{E}|X|/\delta$, we are done. \square

Definition. $X = (X_n)_{n \geq 0}$ is called a *UI [sub/super] martingale* if it is a [sub/super] martingale and $(X_n)_{n \geq 1}$ is uniformly integrable.

Example. Let X_1, X_2, \dots be iid with $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 2) = 1/2$. Set $Y_0 = 1$ and $Y_n = X_1 X_2 \dots X_n$ for $n \geq 1$, so $(Y_n)_{n \geq 0}$ is a martingale and $\mathbb{E}[Y_n] = 1$ for all n . But $Y_n \rightarrow 0$ almost surely.

Theorem. Let X be a martingale. The following are equivalent

- X is UI;
- X converges almost surely in \mathcal{L}^1 to X_∞ as $n \rightarrow \infty$;
- There exists $Z \in \mathcal{L}^1$ such that $X_n = \mathbb{E}[Z|\mathcal{F}_n]$ for all n almost-surely.

Proof. (1 \Rightarrow 2) X is bounded in \mathcal{L}^1 , so by the martingale convergence theorem X converges almost-surely to X_∞ . Since X is also UI, $X_n \rightarrow X_\infty$ in \mathcal{L}^1 too.

(2 \Rightarrow 3) Set $Z = X_\infty$. We need to show $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$ almost surely. Then for $m \geq n$

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

(3 \Rightarrow 1) The previous theorem implies X is UI. \square

Remark. As before we get $X_\infty = \mathbb{E}[Z|\mathcal{F}_\infty]$ almost-surely since $\mathcal{F}_\infty = \sigma(\mathcal{F}_n : n \geq 0)$.

Remark. If X were a UI super/sub-martingale, then we would get $\mathbb{E}[X_\infty|\mathcal{F}_n] \leq X_n$ or $\geq X_n$ respectively.

If X is UI with $X_n \rightarrow X_\infty$, and T is a stopping time then

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

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

$$\mathbb{E}[X_T|\mathcal{F}_S] = X_S \text{ almost-surely.}$$

Proof. We know $X_n = \mathbb{E}[X_\infty|\mathcal{F}_n]$ almost-surely since X is UI. It suffices to prove that for any stopping time T , $\mathbb{E}[X_\infty|\mathcal{F}_T] = X_T$ almost-surely. Indeed, then we will have

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

by the tower property since $\mathcal{F}_S \subseteq \mathcal{F}_T$.

So we just establish $\mathbb{E}[X_\infty|\mathcal{F}_T] = X_T$ almost-surely. First we show $X_T \in \mathcal{L}^1$. We have

$$\begin{aligned} \mathbb{E}[|X_T|] &= \sum_{n \geq 0} \mathbb{E}[|X_n| \mathbb{1}(T = n)] + \mathbb{E}[|X_\infty| \mathbb{1}(T = \infty)] \\ &\leq \sum_{n \geq 0} \mathbb{E}[\mathbb{E}[|X_\infty||\mathcal{F}_n] \mathbb{1}(T = n)] + \mathbb{E}[|X_\infty| \mathbb{1}(T = \infty)] \quad (\text{Jensen}) \\ &= \sum_{n \geq 0} \mathbb{E}[|X_\infty| \mathbb{1}(T = n)] + \mathbb{E}[|X_\infty| \mathbb{1}(T = \infty)] \\ &= \mathbb{E}[|X_\infty|] < \infty. \end{aligned}$$

We have that X_T is \mathcal{F}_T -measurable so we need to show that for all $B \in \mathcal{F}_T$, $\mathbb{E}[X_\infty \mathbb{1}(B)] = \mathbb{E}[X_T \mathbb{1}(B)]$. Indeed

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

□

Backwards martingales

Let $\mathcal{F} \supseteq \mathcal{G}_0 \supseteq \mathcal{G}_{-1} \supseteq \dots$ be a decreasing family of sub- σ -algebras of \mathcal{F} . We call $X = (X_n)_{n \geq 0}$ a *backwards martingale* if $X_0 \in \mathcal{L}^1$ and for all $n \leq -1$,

$\mathbb{E}[X_{n+1}|\mathcal{G}_n] = X_n$ almost-surely.

By the tower property, $\mathbb{E}[X_0|\mathcal{G}_n] = X_n$ for all $n \leq 0$ almost-surely. Since $X_0 \in \mathcal{L}^1$, a backwards martingale is automatically UI.

Theorem. *Let X be a backwards martingale with $X_0 \in \mathcal{L}^p$ for $p \in [1, \infty)$. Then $X_n \rightarrow X_{-\infty}$ almost-surely and in \mathcal{L}^p , where $X_{-\infty} = \mathbb{E}[X_0|\mathcal{G}_{-\infty}]$ for $\mathcal{G}_{-\infty} = \bigcap_{n \geq 0} \mathcal{G}_{-n}$.*

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

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

As before, we get $X_n \rightarrow X_{-\infty}$ as $n \rightarrow -\infty$ almost-surely. $X_{-\infty}$ is $\mathcal{G}_{-\infty}$ -measurable (since it's \mathcal{G}_{-n} -measurable for all $n \geq 0$, so measurable by the intersection). Since $X_0 \in \mathcal{L}^p$, we have $X_n \in \mathcal{L}^p$ for all $n \leq 0$ by Jensen. Also $X_{-\infty} \in \mathcal{L}^p$ by Fatou.

Now we need to show $X_n \rightarrow X_{-\infty}$ in \mathcal{L}^p . We have

$$\begin{aligned} |X_n - X_{-\infty}|^p &= |\mathbb{E}[X_0|\mathcal{G}_n] - \mathbb{E}[X_{-\infty}|\mathcal{G}_n]|^p \\ &\leq \mathbb{E}[|X_0 - X_{-\infty}|^p|\mathcal{G}_n] \end{aligned}$$

hence by a previous result, $(|X_n - X_{-\infty}|^p)_n$ is a UI family. Since $X_n \rightarrow X_{-\infty}$ almost-surely, we have \mathcal{L}^1 convergence of $|X_n - X_{-\infty}|^p$, i.e \mathcal{L}^p convergence of the X_n .

Finally we need to show $X_{-\infty} = \mathbb{E}[X_0|\mathcal{G}_{-\infty}]$ almost-surely. Let $A \in \mathcal{G}_{-\infty} = \bigcap_{n \leq 0} \mathcal{G}_n$ so $A \in \mathcal{G}_n$ for all $n \leq 0$. Then $\mathbb{E}[X_n \mathbb{1}(A)] = \mathbb{E}[X_0 \mathbb{1}(A)]$ for all n . Since $X_n \rightarrow X_{-\infty}$ in \mathcal{L}^1 we have $\mathbb{E}[X_{-\infty} \mathbb{1}(A)] = \mathbb{E}[X_0 \mathbb{1}(A)]$ and so $X_{-\infty} = \mathbb{E}[X_0|\mathcal{G}_{-\infty}]$. \square

Applications of martingales

Theorem (Kolmogorov's 0-1 Law). *Let $(X_n)_{n \geq 0}$ be iid and $\mathcal{F}_n = \sigma(X_k : k \geq n)$ be the tail σ -algebra. Take $\mathcal{F}_\infty = \bigcap_{n \geq 0} \mathcal{F}_n$. Then \mathcal{F}_∞ is trivial, i.e. for all $A \in \mathcal{F}_\infty$ we have $\mathbb{P}(A) \in \{0, 1\}$.*

Proof. Let $A \in \mathcal{F}_\infty$ and let $\mathcal{G}_n = \sigma(X_k : k \leq n)$ be the natural filtration of the X_n , and $\mathcal{G}_\infty = \sigma(\mathcal{G}_n : n \geq 0)$. Note $(\mathbb{E}[\mathbb{1}(A)|\mathcal{G}_n])_{n \geq 0}$ is a martingale and $\mathbb{E}[\mathbb{1}(A)|\mathcal{G}_n] \rightarrow \mathbb{E}[\mathbb{1}(A)|\mathcal{G}_\infty]$ almost-surely. Since $A \in \mathcal{F}_\infty$, we have $A \in \mathcal{F}_{n+1}$ and \mathcal{G}_n is independent of \mathcal{F}_{n+1} by independence of the X_n . So $\mathbb{E}[\mathbb{1}(A)|\mathcal{G}_n] = \mathbb{P}(A)$ almost-surely. Since $\mathcal{F}_\infty \subseteq \mathcal{G}_\infty$ we have $A \in \mathcal{G}_\infty$, we have $\mathbb{E}[\mathbb{1}(A)|\mathcal{G}_\infty] = \mathbb{1}(A)$ almost-surely. Therefore $\mathbb{P}(A) = \mathbb{1}(A)$ almost-surely, so $\mathbb{P}(A) \in \{0, 1\}$. \square

Theorem (Strong Law of Large Numbers). *Let (X_i) be an iid sequence in \mathcal{L}^1 with $\mu = \mathbb{E}[X_1]$. Define $S_n = X_1 + \dots + X_n$. Then $\frac{S_n}{n}$ converges almost-surely and in \mathcal{L}^1 to μ as $n \rightarrow \infty$.*

Proof. Define $\mathcal{G}_n = \sigma(S_n, S_{n+1}, \dots) = \sigma(S_n, X_{n+1}, \dots)$. For $n \leq -1$ let $M_n = \frac{S_{-n}}{-n}$. We will show $(M_n)_{n \leq -1}$ is a backwards martingale with respect to $(\mathcal{G}_{-n})_{n \leq -1}$. We have

$$\mathbb{E}[M_{m+1}|\mathcal{G}_{-m}] = \mathbb{E}\left[\frac{S_{-m-1}}{-m-1}|\mathcal{G}_{-m}\right].$$

Take $n = -m$ so this becomes

$$\begin{aligned} \mathbb{E}\left[\frac{S_{n-1}}{n-1}|\mathcal{G}_n\right] &= \mathbb{E}\left[\frac{S_{n-1}}{n-1}|S_n, X_{n+1}, \dots\right] \\ &= \mathbb{E}\left[\frac{S_n - X_n}{n-1}|S_n\right] && \text{(independence)} \\ &= S_n - \frac{\mathbb{E}[X_n|S_n]}{n-1} \\ &= S_n - \frac{S_n}{n-1} \\ &= \frac{S_n}{n} \\ &= M_m \end{aligned}$$

where we used the fact $\mathbb{E}[X_k|S_n] = \mathbb{E}[X_1|S_n]$ for all $k \in [n]$. Hence we have a backwards martingale, so $\frac{S_n}{n} \rightarrow Y$ almost-surely and in \mathcal{L}^1 for some Y by the Backwards Martingale Theorem.

To finish, we need to show $Y = \mu$ almost-surely. We have

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

Hence Y is $\sigma(X_{k+1}, \dots)$ measurable for all k . Hence Y is $\bigcap_{k \geq 0} \sigma(X_{k+1}, \dots)$ -measurable, so by Kolmogorov's 0-1 law Y is almost-surely constant. Since S_n/n converges to Y in \mathcal{L}^1 , $\lim_{n \rightarrow \infty} \mathbb{E}[S_n/n] = \mu = \mathbb{E}Y = Y$. \square

Theorem (Radon-Nikodym Theorem). *Let \mathbb{P} and Q be two probability measures on the space (Ω, \mathcal{F}) . Suppose \mathcal{F} is countably generated, i.e. there exist $(F_n)_{n \geq 1}$ such that $\mathcal{F} = \sigma(F_n : n \geq 1)$. The following are equivalent*

- $Q \ll \mathbb{P}$, i.e. for all $A \in \mathcal{F}$, $\mathbb{P}(A) = 0$ implies $Q(A) = 0$. We say Q is absolutely continuous with respect to \mathbb{P} ;
- For all $\varepsilon > 0$, there exists $\delta > 0$ such that if $A \in \mathcal{F}$ with $\mathbb{P}(A) < \delta$ then $Q(A) < \varepsilon$;
- There exists a non-negative random variable X such that $Q(A) = \mathbb{E}[X \mathbb{1}(A)]$ for all $A \in \mathcal{F}$.

Note. The general case where \mathcal{F} is not necessarily countably generated follows from this (see Williams).

Remark. X as in (3) is called a *version of the Radon-Nikodym derivative of Q with respect to \mathbb{P}* . We write $X = \frac{dQ}{d\mathbb{P}}$ on \mathcal{F} almost-surely.

Proof. (1 \Rightarrow 2) If 2 doesn't hold, there exists $\varepsilon > 0$ such that for all n there exists $A_n \in \mathcal{F}$ with $\mathbb{P}(A_n) \leq 1/n^2$ and $Q(A_n) \geq \varepsilon$. Then $\sum_{n \geq 1} \mathbb{P}(A_n) < \infty$, so by Borel-Cantelli we see $\mathbb{P}(A_n \text{ i.o.}) = 0$. Hence by (1), $Q(A_n \text{ i.o.}) = 0$. Since

$$\{A_n \text{ i.o.}\} = \bigcap_{n \geq 1} \bigcup_{k \geq n} A_k \implies Q(A_n \text{ i.o.}) = \lim_{n \rightarrow \infty} Q\left(\bigcup_{k \geq n} A_k\right) \geq \varepsilon$$

we have a contradiction.

(2 \Rightarrow 3) Define

$$\mathcal{A}_n = \{H_1 \cap \dots \cap H_n : H_i = F_i \text{ or } H_i = F_i^c \forall i\}$$

and $\mathcal{F}_n = \sigma(\mathcal{A}_n)$. Note the elements of \mathcal{A}_n are disjoint and define $X_n(\omega) = \sum_{A \in \mathcal{A}_n} \frac{Q(A)}{\mathbb{P}(A)} \mathbb{1}(\omega \in A)$. If $A \in \mathcal{F}_n$ we have $\mathbb{E}[X_n \mathbb{1}(A)] = Q(A) = \mathbb{E}[X_{n+1} \mathbb{1}(A)]$. Hence (X_n) is an (\mathcal{F}_n) -martingale.

We have $\mathbb{E}[X_n] = Q(\Omega) = 1$, so (X_n) is an \mathcal{L}^1 -bounded martingale and $X_n \rightarrow X_\infty$ almost-surely as $n \rightarrow \infty$. Furthermore, $\mathbb{P}(X_n \geq \lambda) \leq \frac{1}{\lambda}$ by Markov's inequality, so for any $\varepsilon > 0$, taking $\delta > 0$ as in (2) and setting $\lambda = 1/\delta$ we have

$$\mathbb{E}[X_n \mathbb{1}(X_n \geq \lambda)] = Q(X_n \geq \lambda) < \varepsilon$$

and so (X_n) is UI. Hence $X_n \rightarrow X_\infty$ in \mathcal{L}^1 . Define $\tilde{Q}(A) = \mathbb{E}[X_\infty \mathbb{1}(A)]$ for all $A \in \mathcal{F}$. Then if $A \in \bigcup_{n \geq 0} \mathcal{F}_n$, $A \in \mathcal{F}_n$ for some n and

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

Since $\bigcup_{n \geq 0} \mathcal{F}_n$ is a π -system generating \mathcal{F} , $Q = \tilde{Q}$ on \mathcal{F} .

(3 \Rightarrow 1) Trivial. □