Throughout, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

1 Conditional expectation

Theorem 1.1 (Existence and uniqueness of conditional expectation)**.** *Let $X \in L^1$, and $\mathcal{G} \subseteq \mathcal{F}$. Then there exists a random variable Y such that*

- *Y is $\mathcal{G}$-measurable*
- *$Y \in L^1$, and $\mathbb{E}X\mathbf{1}_A = \mathbb{E}Y\mathbf{1}_A$ for all $A \in \mathcal{G}$.*

Moreover, if Y' is another random variable satisfying these conditions, then $Y' = Y$ almost surely.

We call Y a (version of) the conditional expectation given $\mathcal{G}$.

Proof. (Existence)

Case 1: $X \in L^2$.

Recall that L^2 is a Hilbert space, and that the set of $\mathcal{G}$-measurable random variables is a closed subspace of L^2 (it is closed because the space $L^2(\Omega, \mathcal{G}, \mathbb{P})$ is complete). The projection theorem then gives us the existence and uniqueness of $Y \in L^2 \subseteq L^1$.

Case 2: $X \geq 0 \in L^1$.

Let $X_n = X \wedge n \in L^2$. Then by case 1, we can define $Y_n = \mathbb{E}(X_n \mid \mathcal{G}) \in L^2$. We make the following observation

Lemma 1.0.1. Suppose (X, Y) and (X', Y') are two pairs of random variables satisfying the conditions of the theorem, then $X \geq X'$ implies $Y \geq Y'$ almost surely.

Proof. Let $A = \{Y < Y'\}$. Then $\mathbb{E}Y\mathbf{1}_A = \mathbb{E}X\mathbf{1}_A \geq \mathbb{E}X'\mathbf{1}_A = \mathbb{E}Y'\mathbf{1}_A$, so $\mathbb{E}(Y - Y')\mathbf{1}_A \geq 0$ and $\mathbb{P}(A) = 0$. □

It follows that there is some random variable Y such that $Y_n \uparrow Y$. Clearly Y is $\mathcal{G}$-measurable. For any $A \in \mathcal{G}$, we have

$$\begin{aligned} \mathbb{E}Y\mathbf{1}_A &= \lim_{n\to\infty} \mathbb{E}Y_n\mathbf{1}_A && \text{(MCV)} \\ &= \lim_{n\to\infty} \mathbb{E}X_n\mathbf{1}_A && \text{(MCV)} \\ &= \mathbb{E}X\mathbf{1}_A \end{aligned}$$

Case 3: $X \in L^1$.

Write $X = X^+ - X^-$, and apply case 2 to X^+ and X^-.

(Uniqueness) Suppose Y and Y' are two random variables satisfying the conditions of the theorem. The $\{Y > Y'\}$ is in $\mathcal{G}$ so $\mathbb{E}Y\mathbf{1}_{\{Y>Y'\}} = \mathbb{E}Y'\mathbf{1}_{\{Y>Y'\}} \implies \mathbb{E}(Y - Y')\mathbf{1}_{\{Y>Y'\}} = 0 \implies \mathbb{P}(Y > Y') = 0$. Similarly, $\mathbb{P}(Y' > Y) = 0$.

Remark. The above can also be proved using the Radon-Nikodym theorem.

(Proof via Radon-Nikodym) First recall the Radon-Nikodym theorem

Proposition 1.1 (Radon-Nikodym theorem)**.** Let μ, ν be two σ-finite measures on $(\Omega, \mathcal{F})$ such that $\nu \ll \mu$. Then there exists a unique (up to a.e. equivalence) $f \in L^1(\Omega, \mathcal{F}, \mu)$ such that $\nu(A) = \int_A f \, d\mu$ for all $A \in \mathcal{F}$.

Consider the measure on $(\Omega, \mathcal{G})$ given by

$$\mu(A) = \mathbb{E}X\mathbf{1}_A, \quad A \in \mathcal{G}$$

so $\mu \ll \mathbb{P}$. By the Radon-Nikodym theorem, there exists a unique $Y \in L^1(\Omega, \mathcal{G}, \mathbb{P})$ such that $\mu(A) = \int_A Y \, d\mathbb{P}$ for all $A \in \mathcal{G}$.

For general $X \in L^1$, we can write $X = X^+ - X^-$ and apply the above to X^+ and X^-. □

Proposition 1.2 (Equivalent definition for conditional expectation)**.** Let $X, \mathcal{G}$ be as above. Then there exists a random variable Y such that

- Y is $\mathcal{G}$-measurable
- $Y \in L^1$ and $\mathbb{E}XZ = \mathbb{E}YZ$ for all $Z \in L^\infty(\mathcal{G})$

Moreover, $Y = \mathbb{E}(X \mid \mathcal{G})$ almost surely.

Proof. (Existence) Set $Y = \mathbb{E}(X \mid \mathcal{G})$. It is straightforward to see that Y satisfies the conditions of the proposition for simple functions Z. Note that simple functions that are in L^p are dense in L^p for $1 \leq p \leq \infty$. Let $Z_n \in L^\infty(\mathcal{G})$ be a sequence of simple functions such that $Z_n \to Z$ in L^∞ (in particular, we have almost sure pointwise convergence). Then

$$\begin{aligned}
\mathbb{E}XZ &= \lim_{n\to\infty} \mathbb{E}XZ_n && \text{(DCT)} \\
&= \lim_{n\to\infty} \mathbb{E}YZ_n \\
&= \mathbb{E}YZ && \text{(DCT)}
\end{aligned}$$

□

(Uniqueness) Note that any two random variables satisfying the conditions of the proposition are versions of the conditional expectation given $\mathcal{G}$, which was shown to be unique.

Lemma 1.2.1 (Conditional expectation as a function)**.** Let $X, Y : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Then Y is measurable with respect to $\sigma(X)$ if and only if there exists a Borel-measurable function $f : \mathbb{R} \to \mathbb{R}$ such that $Y(\omega) = f(X(\omega))$ for all $\omega \in \Omega$.

Proposition 1.3 (Properties of conditional expectation)**.** All (in)equality relations below hold almost surely.

1. If $X \geq 0$ a.s., then $\mathbb{E}(X \mid \mathcal{G}) \geq 0$
2. If X and $\mathcal{G}$ are independent, then $\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}[X]$
3. If $\alpha, \beta \in \mathbb{R}$ and $X_1, X_2 \in L^1$, then
$$\mathbb{E}(\alpha X_1 + \beta X_2 \mid \mathcal{G}) = \alpha\mathbb{E}(X_1 \mid \mathcal{G}) + \beta\mathbb{E}(X_2 \mid \mathcal{G}).$$
4. *Tower property*: If $\mathcal{H} \subseteq \mathcal{G}$, then
$$\mathbb{E}(\mathbb{E}(X \mid \mathcal{G}) \mid \mathcal{H}) = \mathbb{E}(X \mid \mathcal{H}).$$
5. If Z is bounded and $\mathcal{G}$-measurable, then
$$\mathbb{E}(ZX \mid \mathcal{G}) = Z\mathbb{E}(X \mid \mathcal{G}).$$
6. Let $X \in L^1$ and $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$. Assume that $\sigma(X, \mathcal{G})$ is independent of $\mathcal{H}$. Then
$$\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(X \mid \sigma(\mathcal{G}, \mathcal{H})).$$

Proof. 1. Follows from the proof of existence and uniqueness of conditional expectation, or just use monotonicity.

2. Let $A \in \mathcal{G}$. Then $\mathbb{E}(\mathbb{E}(X)\mathbf{1}_A) = \mathbb{E}X\mathbb{E}1_A = \mathbb{E}(X1_A)$
3. Use linearity of conditional expectation.
4. Let $A \in \mathcal{H}$. Then $\mathbb{E}\left[\mathbb{E}(\mathbb{E}(X \mid \mathcal{G}) \mid \mathcal{H})\mathbf{1}_A\right] = \mathbb{E}[\mathbb{E}(X \mid \mathcal{G})\mathbf{1}_A] = \mathbb{E}(X\mathbf{1}_A)$
5. Easy if Z is an indicator function. Then use linearity and covergence theorems.
6. Note $\mathbb{E}(X \mid \mathcal{G})$ is $\sigma(\mathcal{G}, \mathcal{H})$-measurable and $\sigma(\mathcal{G}, \mathcal{H})$ is generated by the π-system $\{A \cap B : A \in \mathcal{G}, B \in \mathcal{H}\}$. We show that $\mathbb{E}(X \mid \mathcal{G})$ satisfies the defining property of $\mathbb{E}(X \mid \sigma(\mathcal{G}, \mathcal{H}))$. Let $A \in \mathcal{G}$ and $B \in \mathcal{H}$. Then for any element of the π-system, we have
$$\mathbb{E}(\mathbb{E}(X \mid \mathcal{G})\mathbf{1}_{A\cap B}) = \mathbb{E}[\mathbb{E}(X \mid \mathcal{G})\mathbf{1}_A\mathbf{1}_B] = \mathbb{E}[\mathbb{E}(X\mathbf{1}_A \mid \mathcal{G})\mathbf{1}_B] = \mathbb{E}(\underbrace{X\mathbf{1}_A}_{\in\sigma(\mathcal{G},X)})\mathbb{E}(\mathbf{1}_B) = \mathbb{E}(X\mathbf{1}_{A\cap B})$$
Since finite measures extend uniquely from π-systems, the above holds if $A \cap B$ is replaced by any element of $\sigma(\mathcal{G}, \mathcal{H})$

□

Proposition 1.4 (Properties of conditional expectation)**.** All (in)equality relations below hold almost surely.

1. *Jensen's inequality*: If $c : \mathbb{R} \to \mathbb{R}$ is convex, then
$$\mathbb{E}(c(X) \mid \mathcal{G}) \geq c(\mathbb{E}(X \mid \mathcal{G})).$$
2. *Conditional expectation is a contraction* For $p \geq 1$,
$$\|\mathbb{E}(X \mid \mathcal{G})\|_p \leq \|X\|_p.$$

3. *Monotone convergence theorem* Suppose $X_n \uparrow X$ is a sequence of non-negative random variables. Then
$$\mathbb{E}(X_n \mid \mathcal{G}) \uparrow \mathbb{E}(X \mid \mathcal{G}).$$
4. *Fatou's lemma*: If X_n are non-negative measurable, then
$$\mathbb{E}\left(\liminf_{n\to\infty} X_n \mid \mathcal{G}\right) \leq \liminf_{n\to\infty} \mathbb{E}(X_n \mid \mathcal{G}).$$
5. *Dominated convergence theorem*: If $X_n \to X$ and $Y \in L^1$ such that $Y \geq |X_n|$ for all n, then
$$\mathbb{E}(X_n \mid \mathcal{G}) \to \mathbb{E}(X \mid \mathcal{G}).$$

Proof. 1. Note that a convex function is the supremum of countably many affine functions $c(x) = \sup_{i\in I} a_i x + b_i$. Then
$$\begin{aligned}\mathbb{E}(c(X) \mid \mathcal{G}) &= \mathbb{E}\left(\sup_{i\in I} (a_i X + b_i) \mid \mathcal{G}\right)\\ &\geq \mathbb{E}(a_i X + b_i \mid \mathcal{G}) \quad \forall i \in I && \text{(monotonicity)}\end{aligned}$$

So $\mathbb{E}(c(X) \mid \mathcal{G}) \geq \sup_{i\in I} \mathbb{E}(a_i X + b_i \mid \mathcal{G}) = c(\mathbb{E}(X \mid \mathcal{G}))$.

2. Jensen

3. By monotonicity, $\mathbb{E}(X_n \mid \mathcal{G}) \uparrow Y$ for some Y. By the usual monotone convergence theorem, $\mathbb{E}\mathbb{E}(X_n \mid \mathcal{G}) = \mathbb{E}X_n \to \mathbb{E}Y \leq \mathbb{E}X$ so $Y \in L^1$. Since each of the $\mathbb{E}(X_n \mid \mathcal{G})$ are $\mathcal{G}$-measurable, so is Y. Finally, for any $A \in \mathcal{G}$,
$$\begin{aligned}\mathbb{E}Y\mathbf{1}_A &= \lim_{n\to\infty} \mathbb{E}\mathbb{E}(X_n \mid \mathcal{G})\mathbf{1}_A && \text{(MCV)}\\ &= \lim_{n\to\infty} \mathbb{E}X_n\mathbf{1}_A\\ &= \mathbb{E}X\mathbf{1}_A && \text{(MCV)}\end{aligned}$$

4.
$$\begin{aligned}\mathbb{E}\left(\liminf_{n\to\infty} X_n \mid \mathcal{G}\right) &= \mathbb{E}\left(\lim_{n\to\infty} \underbrace{\inf_{m\geq n} X_m}_{increasing} \mid \mathcal{G}\right)\\ &= \lim_{n\to\infty} \mathbb{E}\left(\inf_{m\geq n} X_m \mid \mathcal{G}\right) && \text{(MCV)}\\ &= \liminf_{n\to\infty} \mathbb{E}\left(\underbrace{\inf_{m\geq n} X_m}_{\leq X_n} \mid \mathcal{G}\right)\\ &\leq \liminf_{n\to\infty} \mathbb{E}(X_n \mid \mathcal{G}) && \text{(monotonicity)}\end{aligned}$$

5. Use Fatou's lemma on $Y + X_n$ and $Y - X_n$.

□

2 Martingales

Definition ((Discrete) stochastic process)**.** A *stochastic process* (in discrete time) is a collection of random variables $(X_n)_{n\in\mathbb{N}}$. A stochastic process is is *integrable* if $X_n \in L^1$ for all n.

Definition (Filtration)**.** A *filtration* is a sequence of σ-algebras $\mathcal{F}_n \subseteq \mathcal{F}$ such that $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ for all n. We define $F_\infty = \sigma(\bigcup_{n=1}^{\infty} \mathcal{F}_n)$. The *natural filtration* of a stochastic process X is the filtration $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. A stochastic process is *adapted* to a filtration $\mathcal{F}_n$ if X_n is $\mathcal{F}_n$-measurable for all n.

Definition (Martingale)**.** An integrable adapted process $(X_n)_{n\geq 0}$ is a *martingale* if for all $n \geq m$, we have

$$\mathbb{E}(X_n \mid \mathcal{F}_m) = X_m.$$

We say it is a *super-martingale* if

$$\mathbb{E}(X_n \mid \mathcal{F}_m) \leq X_m,$$

and a *sub-martingale* if

$$\mathbb{E}(X_n \mid \mathcal{F}_m) \geq X_m,$$

By the tower property, it is sufficient to check the martingale property for $n = m + 1$.

Theorem 2.1 (Doob decomposition, non-examinable)**.** *Let X_n be an integrable adapted process. Then there exists a martingale M_n and an integrable predictable process A_n such that $X_n = M_n + A_n$ and $A_0 = 0$, where predictable means that A_n is $\mathcal{F}_{n-1}$-measurable for all $n \geq 1$. Moreover, M_n and A_n are unique up to a.s. equivalence.*

Proof. (Existence) Add up the 'known' bits to get A and the 'surprises' to get M. Formally,

$$\begin{aligned} A_n &= A_{n-1} + \mathbb{E}(X_n \mid \mathcal{F}_{n-1}) - X_{n-1} \\ M_n &= M_{n-1} + \underbrace{X_n - \mathbb{E}(X_n \mid \mathcal{F}_{n-1})}_{\text{surprise}} \end{aligned}$$

(Uniqueness) Let $X_n = M_n + A_n = M'_n + A'_n$. Then $M_n - M'_n = A'_n - A_n$ is $\mathcal{F}_{n-1}$-measurable. But $M_n - M'_n$ is a martingale, so $\mathbb{E}(M_n - M'_n \mid \mathcal{F}_{n-1}) = 0$ so $M_n = M'_n$ almost surely. Similarly, $A_n = A'_n$ almost surely. $\square$

Definition (Stopping time)**.** A random variable $T : \Omega \to \mathbb{N} \cup \{\infty\}$ is a *stopping time* if $\{T \leq n\} \in \mathcal{F}_n$ for all n.

In the discrete case, we can equivalently require that $\{T = n\} \in \mathcal{F}_n$ for all n.

Definition (X_T)**.** Let X be a stochastic process and T a stopping time. Then $X_T : \Omega \to \mathbb{R}$ is defined by cases

$$X_T(\omega) = \begin{cases} X_n(\omega) & T(\omega) = n \\ 0 & T(\omega) = \infty \end{cases}$$

Definition (Stopped σ-algebra)**.** Let T be a stopping time. Then the *stopped σ-algebra* is

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

Example. Let $N = \#$ of times a random walk hits -5 before it first hits 10 and T be the first time the random walk hits 10. N is $\mathcal{F}_T$-measurable

Definition (Stopped process)**.** Let X be a stochastic process and T a stopping time. Then the *stopped process* is $X_n^T = X_{T \wedge n}$

Proposition 2.1.

1. If $T, S, (T_n)_{n \geq 0}$ are all stopping times, then
$$T \vee S, T \wedge S, \sup_n T_n, \inf T_n, \limsup T_n, \liminf T_n$$
are all stopping times.
2. $\mathcal{F}_T$ is a σ-algebra
3. If $S \leq T$, then $\mathcal{F}_S \subseteq \mathcal{F}_T$.
4. $X_T \mathbf{1}_{T<\infty}$ is $\mathcal{F}_T$-measurable.
5. If (X_n) is an adapted process, then so is $(X_n^T)_{n \geq 0}$ for any stopping time T.
6. If (X_n) is an integrable process, then so is $(X_n^T)_{n \geq 0}$ for any stopping time T.

Proof.

1. Elementary
2. Elementary
3. Let $A \in \mathcal{F}_S$. For any n, we have $A \cap \{S \leq n\} \in \mathcal{F}_n$ and $A \cap \{T \leq n\} = A \cap \{S \leq n\} \cap \{T \leq n\} \in \mathcal{F}_n$.
4. $X_T \mathbf{1}_{T<\infty} = \sum_{n=1}^{\infty} X_n \mathbf{1}_{\{T=n\}}$ where each of the terms is $\mathcal{F}_T$-measurable.
5. $X_n^T = X_n \mathbf{1}_{\{T \geq n\}} + X_T \mathbf{1}_{\{T<n\}} \mathbf{1}_{\{T<\infty\}}$.
6. $X_n^T = X_n \mathbf{1}_{\{T \geq n\}} + \sum_{k=1}^{n-1} X_k \mathbf{1}_{\{T=k\}}$ so $E|X_n^T| \leq E|X_n| + \sum_{k=1}^{n-1} E|X_k| < \infty$.

$\square$

Theorem 2.2 (Equivalent definitions for super-martingales)**.** *Let $(X_n)_{n \geq 0}$ be an integrable and adapted process. Then the following are equivalent:*

1. $(X_n)_{n \geq 0}$ *is a super-martingale.*
2. *For any bounded stopping time T and any stopping time S,*
$$\mathbb{E}(X_T \mid \mathcal{F}_S) \leq X_{S \wedge T}.$$

3. (X_n^T) *is a super-martingale for any stopping time* T.

4. *For bounded stopping times* S, T *such that* $S \leq T$, *we have*

$$\mathbb{E}X_T \leq \mathbb{E}X_S.$$

Proof.
- (2) $\Rightarrow$ (1): Let $n \geq m$ and set $T = n, S = m$.
- (2) $\Rightarrow$ (4): Tower rule
- (2) $\Rightarrow$ (3): Let $n \geq m$
$$\mathbb{E}(X_n^T \mid \mathcal{F}_m) = \mathbb{E}(X_{T\wedge n} \mid \mathcal{F}_m) \leq X_{T\wedge m \wedge n} = X_m^T.$$
- (1) $\Rightarrow$ (2) Let $T \leq N$
$$X_T = X_{S\wedge T} + \sum_{k=0}^{N} (X_{k+1} - X_k)\mathbf{1}_{S\leq k<T} \tag{$*$}$$
Let $A \in \mathcal{F}_S$.
$$\begin{aligned}\mathbb{E}\left[(X_{k+1} - X_k)\mathbf{1}_{S\leq k<T}\mathbf{1}_A\right] &= \mathbb{E}\left[\mathbb{E}\left[(X_{k+1} - X_k)\underbrace{\mathbf{1}_{S\leq k<T}\mathbf{1}_A}_{\in\mathcal{F}_k} \mid \mathcal{F}_k\right]\right] \\ &= \mathbb{E}\left[\mathbf{1}_{S\leq k<T}\mathbf{1}_A \underbrace{\mathbb{E}\left[(X_{k+1} - X_k) \mid \mathcal{F}_k\right]}_{\leq 0}\right] \\ &\leq 0\end{aligned}$$
so $\mathbb{E}X_T\mathbf{1}_A \leq \mathbb{E}X_{S\wedge T}\mathbf{1}_A$. By Radon-Nikodym, $\mathbb{E}(X_{S\wedge T} - X_T \mid \mathcal{F}_S) \geq 0$. But $X_{S\wedge T}$ is $\mathcal{F}_S$-measurable, so $X_{S\wedge T} - X_T \geq 0$ almost surely.
- (4) $\Rightarrow$ (2) Let $n \geq m$ and $A \in \mathcal{F}_m$. One can check that $T = m\mathbf{1}_A + n\mathbf{1}_{A^c} \leq n$ is a stopping time such that
$$\mathbb{E}((X_n - X_m)\mathbf{1}_A) = \mathbb{E}(X_n - X_T) \leq 0$$
By Radon-Nikodym, $\mathbb{E}(X_m - X_n \mid \mathcal{F}_m) \geq 0$ so $\mathbb{E}(X_n \mid \mathcal{F}_m) \leq X_m$.
- (3) $\Rightarrow$ (1) Let $T = \infty$

$\square$

Proposition 2.2 (Convex transformations of martingales)**.** Let (X_n) be a martingale and $c : \mathbb{R} \to \mathbb{R}$ a convex function. Then $(c(X_n))$ is a sub-martingale.

Proof. Let $S \leq T$ be bounded stopping times. Then

$$\begin{aligned}\mathbb{E}(c(X_T) \mid \mathcal{F}_S) &\geq c(\mathbb{E}(X_T \mid \mathcal{F}_S)) && \text{(Jensen)} \\ &= c(X_S)\end{aligned}$$

□

Theorem 2.3 (Optional stopping)**.** *Let $(X_n)_{n\geq 0}$ be a martingale and T a stopping time. Then $E(X_T) = E(X_0)$ if any of the following conditions hold:*

1. *T is almost surely bounded, i.e. there is some N such that $T \leq N$ almost surely.*
2. *X has bounded increments, i.e. there is some K such that $|X_{n+1} - X_n| \leq K$ for all n almost surely and T is integrable*
3. *There exists an integrable random variable Y such that $|X_n| \leq Y$ for all n almost surely and T is finite almost surely, i.e. $\mathbb{P}(T < \infty) = 1$.*

Proof.

1. Use (4) of the previous theorem with $S = 0$, or prove directly.
2. placeholder
3. placeholder

□

3 Convergence

Definition (Upcrossing)**.** Let (x_n) be a sequence and (a, b) an interval. An *upcrossing* of (a, b) by (x_n) is a sequence $j, j+1, \ldots, k$ such that $x_j \leq a$ and $x_k \geq b$. We define

$$\begin{aligned}U_n[a, b, (x_n)] &= \text{number of disjoint upcrossings contained in } \{1, \ldots, n\} \\ U[a, b, (x_n)] &= \lim_{n\to\infty} U_n[a, b, (x_n)].\end{aligned}$$

The notion of upcrossings is related to the notion of convergence.

Proposition 3.1. A sequence (x_n) converges to a limit in the extended real numbers if and only if $U[a, b, (x_n)] < \infty$ for all rationals $a < b$.

Proposition 3.2 (Doob's upcrossing inequality)**.** Let $X = (X_k)$ be a super-martingale and $a < b$. Then

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

Proof. Assume that X is a super-martingale. We define stopping times S_k, T_k as follows:

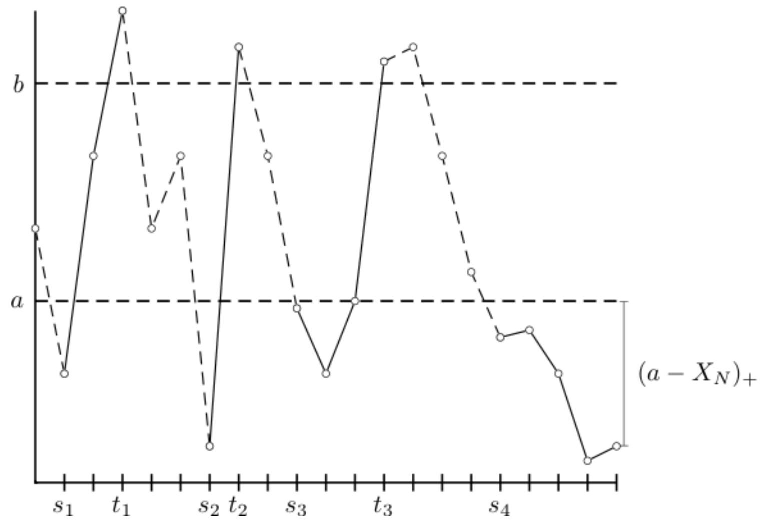


Figure 1: Three upcrossings. Source

- $T_0 = 0$
- $S_{k+1} = \inf\{n : X_n \leq a, n \geq T_n\}$
- $T_{k+1} = \inf\{n : X_n \geq b, n \geq S_{k+1}\}$.

Note that the times alternate, i.e. $S_k \leq T_k \leq S_{k+1} \leq T_{k+1}$

The idea is to sum up the increments of each upcrossing. Consier the set $\mathcal{I} := \{k : S_k < T_k < n\}$ and the sum

$$\underbrace{\sum_{k \in \mathcal{I}} (X_{T_k} - X_{S_k})}_{\geq (b-a) U_n[a,b,X]} + \begin{cases} X_n - X_{S_{\max \mathcal{I}+1}} & \text{if } n > S_{\max \mathcal{I}+1} \\ 0 & \text{otherwise} \end{cases}$$

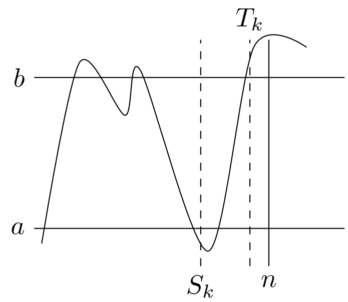
Note that each term is of the form $X_U - X_V$ for some *bounded* stopping times $U \geq V$. By the property of super-martingales, $\mathbb{E}(X_U - X_V) \leq 0$ so the sum is negative in expectation.

Hence,

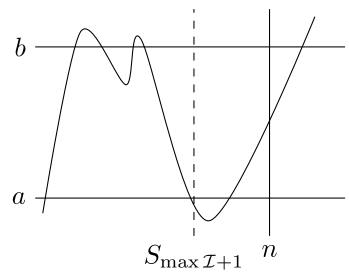
$$(b-a)\mathbb{E}U_n[a,b,X] + \mathbb{E}(\underbrace{\underbrace{(X_n - X_{S_{\max \mathcal{I}+1}})}_{\geq X_n - a \geq -(X_n - a)^-} \mathbf{1}_{n > S_{\max \mathcal{I}+1}}}_{\geq -(X_n - a)^-}) \leq 0$$

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

$\square$



(a) No extra term



(b) Extra term

Figure 2: Graphs taken from here

Proposition 3.3 (Almost sure martingale convergence theorem)**.** Suppose $X = (X_n)_{n\geq 0}$ is a super-martingale that is bounded in L^1, i.e. $\sup_n \mathbb{E}|X_n| < \infty$. Then for any $a < b$, we have $U[a, b, X] < \infty$ almost surely. In particular, there exists an $\mathcal{F}_\infty$-measurable $X_\infty \in L^1$ such that

$$X_n \to X_\infty \text{ a.s. as } n \to \infty.$$

Remark. The intuition is that you cannot make money by betting on a super-martingale (without shorting). For any $a < b$, you can devise a betting strategy where you buy at a and sell at b. If $U[a, b, X] = \infty$, then you make money almost surely.

Proof. Let $a < b$ and n be arbitrary and $\sup_m \mathbb{E}|X_m| = M$. By Doob's upcrossing inequality, we have

$$\begin{aligned}(b-a)\mathbb{E}U_n[a, b, X] &\leq M + |a| < \infty \\ (b-a)\mathbb{E}U[a, b, X] &\leq M + |a| < \infty \qquad \text{(monotone convergence, } U_n \uparrow U)\end{aligned}$$

So $U[a, b, X] < \infty$ almost surely.

Now consider the set of events where X_n converges

$$A := \cap_{a,b\in\mathbb{Q}}\{U[a, b, X] < \infty\}, \qquad \mathbb{P}(A) = 1$$

and define

$$X_\infty = \begin{cases} \lim X_n & \text{on } A \\ 0 & \text{on } A^c \end{cases}$$

which is $\mathcal{F}_\infty$-measurable. Note that $|X_\infty| = \liminf |X_n|$ almost surely and by Fatou's lemma,

$$\mathbb{E} \liminf |X_n| \leq \liminf \mathbb{E}|X_n| \leq M < \infty.$$

So $X_\infty \in L^1$. □

Remark. Sone propositions below concern non-negative submartingales. Recall convex transformations, such as $|\cdot|$, turn a martingale into a sub-martingale.

Proposition 3.4 (Doob's maximal inequality)**.** Let $X = (X_n)_{n\geq 0}$ be a non-negative sub-martingale. Then for any $\lambda > 0$ and writing $X_n^* = \max_{0\leq k\leq n} X_k$, we have

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

Proof. Let $T = \inf\{n : X_n \geq \lambda\}$

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

□

Corollary 1. Under the same hypotheses, let $X_n^* \uparrow X^*$. Then

$$\lambda \mathbb{P}(X^* \geq \lambda) \leq \sup_{n\geq 0} \mathbb{E}X_n$$

Proposition 3.5 (Doob's L^p inequality)**.** Let $X = (X_n)_{n\geq 0}$ be a martingale or a non-negative sub-martingale. Then for any $p > 1$ and $n \geq 1$, we have

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

Proof. (magic)
Let $k > 0$ and $T = \inf\{n : X_n \geq k\}$. Then

$$\begin{aligned}\int |X_n^* \wedge k|^p \,\mathrm{d}\mathbb{P} &= \int \left(\int p x^{p-1} \mathbf{1}_{\{0\leq x\leq |X_n^*\wedge k|\}} \,\mathrm{d}x \right) \mathrm{d}\mathbb{P}\\ &= \int \left(\int p x^{p-1} \mathbf{1}_{\{|X_n^*|\geq x\}} \mathbf{1}_{\{0\leq x\leq k\}} \,\mathrm{d}x \right) \mathrm{d}\mathbb{P}\\ &= \int p x^{p-1} \mathbb{P}(|X_n^*| \geq x) \mathbf{1}_{\{0\leq x\leq k\}} \,\mathrm{d}x\\ &\leq \int p x^{p-1} \frac{1}{x} \mathbb{E}|X_n| \mathbf{1}_{\{|X_n^*|\geq x\}} \mathbf{1}_{\{0\leq x\leq k\}} \,\mathrm{d}x && \text{(Doob's maximal inequality)}\\ &= \int\int p x^{p-2} |X_n| \mathbf{1}_{\{|X_n^*|\geq x\}} \mathbf{1}_{\{0\leq x\leq k\}} \,\mathrm{d}x \,\mathrm{d}\mathbb{P}\\ &= \int \frac{p}{p-1} |X_n| (|X_n^*| \wedge k)^{p-1} \,\mathrm{d}\mathbb{P}\\ &\leq \frac{p}{p-1} \|X_n\|_p \||X_n^*| \wedge k\|_p^{p-1} && \text{(Hölder's inequality)}\end{aligned}$$

By monotone convergence and taking $k \to \infty$, we get

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

and the result follows. □

Corollary 2. Under the same hypotheses, let $X_n^* \uparrow X^*$. Then

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

Proposition 3.6 (Equivalent conditions for L^p convergence, $p > 1$)**.** Let $X = (X_n)_{n\geq 0}$ be a martingale, and $p > 1$. Then the following are equivalent:

1. $(X_n)_{n\geq 0}$ is bounded in L^p, i.e. $M = \sup_n \mathbb{E}|X_i|^p < \infty$.
2. $(X_n)_{n\geq 0}$ converges as $n \to \infty$ to a random variable $X_\infty \in L^p$ almost surely and in L^p.
3. There exists a random variable $Z \in L^p$ such that

$$X_n = \mathbb{E}(Z \mid \mathcal{F}_n) \qquad \lim_{n\to\infty} X_n = \mathbb{E}(Z \mid \mathcal{F}_\infty) \text{ a.s.}$$

This gives a bijection between martingales bounded in L^p and $L^p(\mathcal{F}_\infty)$, sending $(X_n)_{n\geq 0} \mapsto X_\infty$.

Proof.

- (1) $\Rightarrow$ (2) By Jensen, $\mathbb{E}|X_n|^p \geq (\mathbb{E}|X_n|)^p$ so X is bounded in L^1. By the almost sure martingale convergence theorem, there exists an $\mathcal{F}_\infty$-measurable $X_\infty \in L^1$ such that $X_n \to X_\infty$ almost surely. Note also $|X|^* \geq |X_n|$ for all n and $X^* \in L^p$ by Corollary 2. Hence, $X_n \to X_\infty$ in L^p by L^p dominated convergence

- (2) $\Rightarrow$ (3) Let $Z = X_\infty$. Note $X_n \xrightarrow{L^p} X_\infty$ implies that X is bounded in L^p. By Doob's maximal inequality, $|X|^* \in L^p \subseteq L^1$. By conditional dominated convergence,

$$\mathbb{E}(X_\infty \mid \mathcal{F}_n) = \lim_{m\to\infty} \mathbb{E}(X_m \mid \mathcal{F}_n) = X_n$$

- (3) $\Rightarrow$ (1) Conditional expectation is a contraction (Proposition 1.4)

$\square$

Definition (Closed martingale)**.** A martingale in the form $X_n = \mathbb{E}(Z \mid \mathcal{F}_n)$ for some $Z \in L^p$ is called a *martingale closed in L^p*

Definition (Non-examinable, uniform integrability for general measure space)**.** Let $(f_n)_n$ be a family of absolutely integrable functions on some measure space. The family is said to be *uniformly integrable (UI)* if all of the following hold

1. Uniform bound on L^1 norm ($\sup_n \int |f_n| < \infty$)
2. No escape to vertical infinity ($\sup_n \int_{\{|f_n|>\lambda\}} |f_n| \to 0$ as $\lambda \to \infty$)
3. No escape to horizontal infinity (for any $\varepsilon > 0$, there exists a finite measure subset A such that $\sup_n \int_{A^c} |f_n| < \varepsilon$)

Example. A single integrable function is uniformly integrable.

Example. A family of functions which is dominated by some integrable function, i.e. there is $g \in L^1$ such that $|f_i| \leq g$ for all $i \in \mathcal{I}$, is uniformly integrable.

Definition (Uniform integrability for random variables)**.** A family of random variables $(X_i)_{i\in\mathcal{I}}$ is *uniformly integrable* if

$$\sup_{i\in\mathcal{I}} \mathbb{E}(|X_i|\mathbf{1}_{|X_i|>\alpha}) \to 0 \text{ as } \alpha \to \infty.$$

Remark. In the finite measure case, functions cannot escape to horizontal infinity and the conditions simplifies just to no escape to vertical infinity.

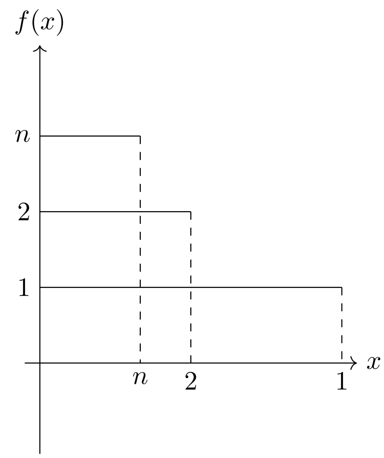


Figure 3: Non uniformly integrable sequence

Proposition 3.7 (Equivalent definition of uniform integrability)**.** A family of random variables $(X_i)_{i\in\mathcal{I}}$ is *uniformly integrable* if and only if both of the following hold

1. It is bounded in L^1 ($\sup_{i\in\mathcal{I}} \mathbb{E}(|X_i|) < \infty$)
2. It is equi-integrable (For any $\varepsilon > 0$, there exists a $\delta > 0$ such that for any $A \in \mathcal{F}$ with $\mathbb{P}(A) < \delta$, we have

$$\mathbb{E}(|X_i|\mathbf{1}_A) < \varepsilon.$$

   for all $i \in \mathcal{I}$.)

Theorem 3.1 (Non-examinable, Vitali convergence theorem)**.** *Let $f_1, f_2, \ldots$ be a sequence of integrable functions on some measure space and $1 \leq p < \infty$. Then $f_n \xrightarrow{L^p} f$ for some measurable f if and only if all of the following hold*

1. *(f_n^p) is uniformly integrable*
2. *$f_n \to f$ in measure*
3. *The sequence cannot escape to horizontal infinity, i.e. for any $\varepsilon > 0$, there exists a finite measure subset A such that $\sup_n \int_{A^c} |f_n|^p < \varepsilon$.*

Remark. In the finite measure case, the third condition is trivially true and almost sure convergence implies convergence in measure, so this implies the L^p dominated convergence theorem.

Proposition 3.8 (Conditional expectations are uniformly integrable)**.** Let $\mathcal{S}$ be a uniformly integrable family of random variables. Then the following set is uniformly integrable

$$\mathcal{S}^* = \{\mathbb{E}(X|\mathcal{G}) \mid X \in \mathcal{S}, \mathcal{G} \text{ is a sub } \sigma\text{-algebra of } \mathcal{F}\}.$$

Proof. Since $\mathcal{S}$ is bounded in L^1, $\mathcal{S}^*$ is bounded in L^1. Let $\varepsilon > 0$. By uniform integrability of $\mathcal{S}$, there exists $\delta > 0$ such that for any $A \in \mathcal{F}$ with $\mathbb{P}(A) < \delta$, we have $\mathbb{E}(|X|\mathbf{1}_A) < \varepsilon$ for all $X \in \mathcal{S}$. Note

$$\mathbb{E}(|\mathbb{E}(X \mid \mathcal{G})|\mathbf{1}_A) \underbrace{\leq}_{Jensen} \mathbb{E}(\mathbb{E}(|X| \mid \mathcal{G})\mathbf{1}_A)$$

we wish to show that when A is of the form $\{|X| > \alpha\}$, the right hand side converges to zero as $\alpha \to \infty$. The right hand side becomes $\mathbb{E}|X|\mathbf{1}_{\{|X|>\alpha\}}$

We want to choose α such that $\mathbb{P}(|X| > \alpha) < \delta$. By Markov's inequality,

$$\mathbb{P}(|X| > \alpha) \leq \frac{\mathbb{E}|X|}{\alpha}$$

so picking any $\alpha > \frac{\mathbb{E}|X|}{\delta}$ works.

We have shown that for any $\varepsilon > 0$, for any α sufficiently large, $\mathbb{E}(|\mathbb{E}(X \mid \mathcal{G})| \, \mathbf{1}_{\{|X|>\alpha\}}) \leq \varepsilon$ □

Proposition 3.9 (Equivalent conditions for L^1 convergence)**.** Let $(X_n)_{n\geq 0}$ be a martingale. Then the following are equivalent:

1. $(X_n)_{n\geq 0}$ is uniformly integrable.
2. $(X_n)_{n\geq 0}$ converges to some $X_\infty \in L^1$ almost surely and in L^1.
3. There exists $Z \in L^1$ such that $X_n = \mathbb{E}(Z \mid \mathcal{F}_n)$ almost surely.

Moreover, $X_\infty = \mathbb{E}(Z \mid \mathcal{F}_\infty)$

Proof.

- (1) ⇒ (2) X is L^1 bounded and hence converges to some X_∞ almost surely by Proposition 3.3. By Vitali, $X_n \to X_\infty$ in L^1.
- (2) ⇒ (3) Let $Z = X_\infty$. Then $X_n = \mathbb{E}(Z \mid \mathcal{F}_n)$ almost surely by Proposition 1.4. (Same as previous proposition) Let $Z = X_\infty$.

  $$\|X_n - \mathbb{E}(X_\infty|\mathcal{F}_n)\|_1 = \|\mathbb{E}(X_m - X_\infty \mid \mathcal{F}_n)\|_1 \leq \|X_m - X_\infty\|_1$$

  for any $m \geq n$ and the right hand side converges to 0 by L^1 convergence.

– (3) ⇒ (1) Conditional expectation is uniformly integrable, see previous example.

□

Lemma 3.9.1 (Stopped UI process)**.** Let X be a uniformly integrable martingale and T be any stopping time. Then the following statements about the stopped process X^T hold:

1. $X_{T\wedge n} = \mathbb{E}(X_\infty|\mathcal{F}_{T\wedge n})$
2. X^T is uniformly integrable
3. $X_n^T \to X_T$ in L^1 and almost surely

Proof. Since X is UI, we use the fact that the martingale can be represented as a conditional expectation of some $X_\infty \in L^1$. Then

$$\begin{aligned} X_{T\wedge n} &= \mathbb{E}(X_n|\mathcal{F}_{T\wedge n}) \\ &= \mathbb{E}(\mathbb{E}(X_\infty \mid \mathcal{F}_n) \mid \mathcal{F}_{T\wedge n}) \\ &= \mathbb{E}(X_\infty|\mathcal{F}_{T\wedge n}) \end{aligned}$$

By Proposition 3.8, conditional expectations are uniformly integrable. By Proposition 3.9, $X_n^T \to X_\infty^T$ in L^1 and almost surely for some $X_\infty^T \in L^1$. By considering different values of T, one can see that $X_\infty^T = X_T$ almost surely. □

Proposition 3.10 (Optional stopping for arbitrary stopping times)**.** If $(X_n)_{n\geq 0}$ is a uniformly integrable martingale, and S, T are arbitrary stopping times, then $\mathbb{E}(X_T \mid \mathcal{F}_S) = X_{S\wedge T}$. In particular $\mathbb{E}X_T = X_0$.

Note that we are now allowing arbitrary stopping times, so T may be infinite with non-zero probability. Hence we define

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

Proof. We have proven the result for bounded stopping times. For the stopped process $X^T = (X_{T\wedge n})_{n\geq 0}$, we have $\mathbb{E}(X_n^T \mid \mathcal{F}_S) = X_{S\wedge T\wedge n}$. What we would like to do is it take the limit as $n \to \infty$.

From Lemma 3.9.1,

$$\|\mathbb{E}(X_{T\wedge n} - X_T|\mathcal{F}_S)\|_1 \leq \|X_{T\wedge n} - X_T\|_1 \to 0$$

so $X_{S\wedge T\wedge n} \xrightarrow{L^1} \mathbb{E}(X_T \mid \mathcal{F}_S)$.

Lemma 3.9.1 also says $X_{S\wedge T\wedge n} \xrightarrow{L^1} X_{S\wedge T}$ so $X_{S\wedge T} = \mathbb{E}(X_T \mid \mathcal{F}_S)$ almost surely □