Topics:

- Martingales.
- Filtration. Stopping times.
- Probability harmonic functions.
- Optional Stopping Theorem.
- Martingale Convergence Theorem.
Conditional expectation.

Consider a probability space \((\Omega, \mathcal{F}, P)\) and a random variable \(X \in \mathcal{F}\).

Let \(\mathcal{G} \subseteq \mathcal{F}\) be a smaller \(\sigma\)-algebra.

**Definition.** Conditional expectation \(E[X|\mathcal{G}]\) is a unique function from \(\Omega\) to \(\mathbb{R}\) satisfying:

1. \(E[X|\mathcal{G}]\) is \(\mathcal{G}\)-measurable

\[
\int_A E[X|\mathcal{G}] \, dP(\omega) = \int_A X \, dP(\omega) \quad \text{for all} \quad A \in \mathcal{G}
\]

The existence and uniqueness of \(E[X|\mathcal{G}]\) comes from the Radon-Nikodym theorem.

**Lemma.** If \(X \in \mathcal{G}\), \(Y(\omega) \in L^1(\Omega, P)\), and \(X(\omega) \cdot Y(\omega) \in L^1(\Omega, P)\), then

\[
E[X \cdot Y|\mathcal{G}] = X \cdot E[Y|\mathcal{G}]
\]
Conditional expectation.

Consider a probability space \((\Omega, \mathcal{F}, P)\) and a random variable \(X \in \mathcal{F}\).

**Lemma.** If \(\mathcal{G} \subseteq \mathcal{F}\), then \(E[E[X|\mathcal{G}]] = E[X]\)

Let \(\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}\) be smaller sub-\(\sigma\)-algebras.

**Lemma.**

\[
E\left[ E[X|\mathcal{G}_2] \mid \mathcal{G}_1 \right] = E[X|\mathcal{G}_1]
\]

**Proof.** For any \(A \in \mathcal{G}_1 \subseteq \mathcal{G}_2\),

\[
\int_A E\left[ E[X|\mathcal{G}_2] \mid \mathcal{G}_1 \right](\omega) \, dP(\omega) = \int_A E[X|\mathcal{G}_2](\omega) \, dP(\omega)
\]

\[
= \int_A X(\omega) \, dP(\omega) = \int_A E[X|\mathcal{G}_1](\omega) \, dP(\omega)
\]
Filtration.

Definition. Consider an arbitrary linear ordered set $T$: A sequence of sub-$\sigma$-algebras $\{F_t\}_{t \in T}$ of $\mathcal{F}$ is said to be a filtration if

$$F_s \subseteq F_t \quad a.s. \quad \forall s < t \in T$$

Example. Consider a sequence of random variables $X_1, X_2, \ldots$ on $(\Omega, \mathcal{F}, P)$, and let $F_n = \sigma(X_1, X_2, \ldots, X_n)$ is the smallest $\sigma$-algebra such that $X_1, X_2, \ldots, X_n$ are $F_n$-measurable. Then $F_n$ is the smallest filtration that $X_n$ is adapted to, i.e. $X_n \in F_n$.

Important: When filtration $F_n$ is not mentioned in defining the martingale, submartingale, or supermartingale,

$$F_n = \sigma(X_1, X_2, \ldots, X_n)$$

Definition. Consider a filtration $\{F_n\}$. A sequence of random variables $X_1, X_2, \ldots \in L^1(\Omega, P)$ adapted to $F_n$ (i.e. $X_n \in F_n$) is said to be a martingale with respect to $\{F_n\}$ if

$$E[X_{n+1} \mid F_n] = X_n \quad a.s. \quad \forall n > 1$$
Martingales.

Definition. Consider a filtration \( \{ \mathcal{F}_n \} \). A sequence of random variables \( X_1, X_2, \ldots \in L^1(\Omega, P) \) adapted to \( \mathcal{F}_n \) (i.e. \( X_n \in \mathcal{F}_n \)) is said to be a martingale with respect to \( \{ \mathcal{F}_n \} \) if
\[
E[X_{n+1} | \mathcal{F}_n] = X_n \quad \text{a.s.} \quad \forall n > 1
\]

Example. Let \( \xi_1, \xi_2, \ldots \) be independent \( L^1(\Omega, P) \) random variables such that
\[
E[\xi_j] = 0 \quad \forall j \in \mathbb{N}
\]
Now, let \( X_n = \xi_1 + \ldots + \xi_n \). Then
\[
E[X_{n+1} | \mathcal{F}_n] = E[X_n + \xi_{n+1} | \mathcal{F}_n] = X_n + E[\xi_{n+1} | \mathcal{F}_n] = X_n + E[\xi_{n+1}] = X_n
\]
as \( \xi_{n+1} \) is independent of \( \mathcal{F}_n \). Specifically, \( \forall m \in \mathbb{N} \) s.t. \( 1 \leq m \leq n \), and any Borel \( A \in \mathcal{B} \),
\[
\int_{X_m^{-1}(A)} E[\xi_{n+1} | \mathcal{F}_n](\omega) \, dP(\omega) = \int_{X_m^{-1}(A)} \xi_{n+1}(\omega) \, dP(\omega) = E[\xi_{n+1} \cdot 1_{X_m \in A}]
\]
\[
= E[\xi_{n+1}] \cdot E[1_{X_m \in A}] = \int_{X_m^{-1}(A)} E[\xi_{n+1}] \, dP(\omega)
\]
Martingales.

**Definition.** Consider a filtration \( \{ \mathcal{F}_n \} \). A sequence of random variables \( X_1, X_2, \ldots \in L^1(\Omega, P) \) adapted to \( \mathcal{F}_n \) (i.e. \( X_n \in \mathcal{F}_n \)) is said to be a **martingale** with respect to \( \{ \mathcal{F}_n \} \) if

\[
E[X_{n+1} \mid \mathcal{F}_n] = X_n \quad a.s. \quad \forall n > 1
\]

**Definition.** Consider a filtration \( \{ \mathcal{F}_n \} \). A sequence of random variables \( X_1, X_2, \ldots \in L^1(\Omega, P) \) adapted to \( \mathcal{F}_n \) is said to be a **supermartingale** with respect to \( \{ \mathcal{F}_n \} \) if

\[
E[X_{n+1} \mid \mathcal{F}_n] \leq X_n \quad a.s. \quad \forall n > 1
\]

**Definition.** Consider a filtration \( \{ \mathcal{F}_n \} \). A sequence of random variables \( X_1, X_2, \ldots \in L^1(\Omega, P) \) adapted to \( \mathcal{F}_n \) is said to be a **submartingale** with respect to \( \{ \mathcal{F}_n \} \) if

\[
E[X_{n+1} \mid \mathcal{F}_n] \geq X_n \quad a.s. \quad \forall n > 1
\]

All these definitions can be extended to an arbitrary linear ordered set \( T \): Consider a filtration \( \{ \mathcal{F}_t \}_{t \in T} \). A sequence of random variables \( \{ X_t \}_{t \in T} \) adapted to \( \mathcal{F}_t \) is said to be a **martingale** if

\[
E[X_t \mid \mathcal{F}_s] = X_s \quad a.s. \quad \forall s < t \in T
\]
Probability harmonic functions.

Consider a sequence of random variables $X_1, X_2, \ldots$ with associated $\sigma$-algebras $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$.

**Definition.** A function $h(x)$ is said to be a probability harmonic function if $M_t = h(X_t)$ is a martingale sequence.

**Example.** Random walk on $\mathbb{Z}$. Take $p \in (0, 1)$, and let $\xi_1, \xi_2, \ldots$ be i.i.d. Bernoulli random variables such that

$$
\xi_j = \begin{cases} 
+1 & \text{with probability } p \\
-1 & \text{with probability } q = 1 - p 
\end{cases}
$$

If $p = \frac{1}{2}$, the random walk $X_n = X_0 + \xi_1 + \ldots + \xi_n$ is a martingale.

Suppose $p \neq \frac{1}{2}$, then $X_n = X_0 + \xi_1 + \ldots + \xi_n$ is not a martingale. We need a probability harmonic function $h(x)$ such that $M_n = h(X_n)$ is a martingale. For this, we solve

$$
p \cdot h(X_n + 1) + q \cdot h(X_n - 1) = E[h(X_{n+1}) \mid \mathcal{F}_n] = h(X_n)
$$

arriving at $h(x) = A \cdot \left(\frac{q}{p}\right)^x + B$ for any choice of constants $A$ and $B$. 
Filtration. Stopping time.

Definition. Consider an arbitrary linear ordered set $T$: A sequence of sub-$\sigma$-algebras $\{\mathcal{F}_t\}_{t \in T}$ of $\mathcal{F}$ is said to be a filtration if

$$\mathcal{F}_s \subseteq \mathcal{F}_t \quad a.s. \quad \forall s < t \in T$$

Example. Consider a sequence of random variables $X_1, X_2, \ldots$ on $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$. Then $\mathcal{F}_n$ is a filtration.

Definition. Consider an arbitrary linear ordered set $T$, and a filtration $\{\mathcal{F}_t\}_{t \in T}$. A random variable $\tau$ is a stopping time if for any $t \geq 0$,

$$\{\tau \leq t\} \in \mathcal{F}_t$$

In other words knowing the trajectory of the process up to time $m$ is sufficient to determine whether $\{\tau \leq t\}$ occurred.
Filtration. Stopping time.

**Definition.** Consider an arbitrary linear ordered set $T$, and a filtration $\{F_t\}_{t \in T}$. A random variable $\tau$ is a stopping time if for any $t \geq 0$,

$$\{\tau \leq t\} \in F_t$$

In other words knowing the trajectory of the process up to time $t$ is sufficient to determine whether $\{\tau \leq t\}$ occurred.

For every stopping time $\tau$ we associate a stopped $\sigma$-algebra $F_\tau \subseteq \mathcal{F}$ defined as

$$F_\tau = \left\{ A \in \mathcal{F} : A \cap \{\tau \leq t\} \in F_t \quad \forall t \right\}$$

**Observe** that if $\{F_n\}$ is a filtration, and if $X_1, X_2, \ldots$ is a sequence of random variables adapted to $F_n$, and $\tau$ is a stopping time w.r.t. $\{F_n\}$, then

$$X_\tau = \sum_{j=1}^{\infty} X_j \cdot 1_{\tau=j} \in F_\tau$$
Filtration. Stopping time.

For every stopping time $\tau$ we associate a stopped $\sigma$-algebra $F_\tau \subseteq F$ defined as

$$F_\tau = \{ A \in F : A \cap \{ \tau \leq t \} \in F_t \quad \forall t \}$$

Lemma. Suppose $\tau_1$ and $\tau_2$ are two stopping times w.r.t. $F_n$ such that $P(\tau_1 \leq \tau_2) = 1$, then

$$F_{\tau_1} \subseteq F_{\tau_2}$$

Proof. Take $A \in F_{\tau_1}$, then $\forall t$,

$$A \cap \{ \tau_2 \leq t \} = A \cap \{ \tau_1 \leq t \} \cap \{ \tau_2 \leq t \} \quad P - a.s.$$ and therefore $A \cap \{ \tau_2 \leq t \} \in F_t$ as both $A \cap \{ \tau_1 \leq t \}$ and $\{ \tau_2 \leq t \}$ are in $F_t$. 

$\square$
Filtration. Stopping time.

For every stopping time $\tau$ we associate a stopped $\sigma$-algebra $\mathcal{F}_\tau \subseteq \mathcal{F}$ defined as

$$\mathcal{F}_\tau = \left\{ A \in \mathcal{F} : A \cap \{ \tau \leq t \} \in \mathcal{F}_t \quad \forall t \right\}$$

Lemma. Suppose $\tau$ is a stopping time w.r.t. $\mathcal{F}_n$ such that $\tau \leq K$ a.s. for some integer $K > 0$. Then, if the sequence $\{X_t\}$ is a martingale,

$$E[X_K | \mathcal{F}_\tau] = X_\tau$$

Proof. Take $A \in \mathcal{F}_\tau$, then

$$\int_A X_K(\omega) \, dP(\omega) = \sum_{j=0}^{K} \int_{A \cap \{ \tau = j \}} X_K(\omega) \, dP(\omega) = \sum_{j=0}^{K} \int_{A \cap \{ \tau = j \}} E[X_K | \mathcal{F}_j](\omega) \, dP(\omega)$$

$$= \sum_{j=0}^{K} \int_{A \cap \{ \tau = j \}} X_j(\omega) \, dP(\omega) = \sum_{j=0}^{K} \int_{A \cap \{ \tau = j \}} X_\tau(\omega) \, dP(\omega) = \int_A X_\tau(\omega) \, dP(\omega)$$
**Optional Stopping Theorem.**

**Doob’s Optional Stopping Theorem.** Consider a sequence of random variables $X_1, X_2, \ldots$ on $(\Omega, \mathcal{F}, P)$, and let $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$. Suppose $\tau_1$ and $\tau_2$ are two stopping times w.r.t. $\mathcal{F}_n$ such that either of the following conditions is satisfied:

(a) $P(\tau_1 \leq \tau_2 \leq K) = 1$ for some $K > 0$

(b) $P(\tau_1 \leq \tau_2 < \infty) = 1$ and $S = \sup_{0 \leq k \leq \tau_2} |X_k| \in L^1(\Omega, P)$

Then, if the sequence $\{X_t\}$ is a martingale,

$$E[X_{\tau_2} | \mathcal{F}_{\tau_1}] = X_{\tau_1}$$

Similarly, if the sequence $\{X_t\}$ is a supermartingale, $E[X_{\tau_2} | \mathcal{F}_{\tau_1}] \leq X_{\tau_1}$, and if the sequence $\{X_t\}$ is a submartingale, $E[X_{\tau_2} | \mathcal{F}_{\tau_1}] \geq X_{\tau_1}$.

**Proof.** (part (a)) Suppose $P(\tau_1 \leq \tau_2 \leq K) = 1$ for some integer $K > 0$. Then

$$\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2} \subseteq \mathcal{F}_K$$
Optional Stopping Theorem.

Proof. (part (a)) Suppose \( P(\tau_1 \leq \tau_2 \leq K) = 1 \) for some integer \( K > 0 \). Then

\[
\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2} \subseteq \mathcal{F}_K
\]

and

\[
E[X_{\tau_2}|\mathcal{F}_{\tau_1}] = E[E[X_K|\mathcal{F}_{\tau_2}] \mid \mathcal{F}_{\tau_1}] = E[X_K|\mathcal{F}_{\tau_1}] = X_{\tau_1}
\]

\[
\square
\]

Proof. (part (b)) Suppose \( \{X_t\} \) is a martingale. For \( K > 0 \), consider a stopped process \( Y_t = X_{t\wedge K} \). Then, \( \tau_1^* = \tau_1 \wedge K \) and \( \tau_2^* = \tau_2 \wedge K \) are both bounded stopping times, as in part (a), and

\[
E[Y_{\tau_2^*}|\mathcal{F}_{\tau_1^*}] = Y_{\tau_1^*} \iff E[Y_{\tau_2}|\mathcal{F}_{\tau_1}] = Y_{\tau_1}
\]

Therefore

\[
E[X_{\tau_2}|\mathcal{F}_{\tau_1}] + E[(X_K - X_{\tau_2}) \cdot 1_{\tau_2 > K}|\mathcal{F}_{\tau_1}] = X_{\tau_1} \cdot 1_{\tau_1 \leq K} + X_K \cdot 1_{\tau_1 > K}
\]
Optional Stopping Theorem.

Proof. (part (b)) Suppose \( \{X_t\} \) is a martingale. For \( K > 0 \), consider a stopped process \( Y_t = X_{t \wedge K} \). Then, \( \tau_1^* = \tau_1 \wedge K \) and \( \tau_2^* = \tau_2 \wedge K \) are both bounded stopping times, as in part (a), and

\[
E[Y_{\tau_2^*} | F_{\tau_1^*}] = Y_{\tau_1^*} \quad \Leftrightarrow \quad E[Y_{\tau_2} | F_{\tau_1}] = Y_{\tau_1}
\]

Therefore

\[
E[X_{\tau_2} | F_{\tau_1}] + E[(X_K - X_{\tau_2}) \cdot 1_{\tau_2 > K} | F_{\tau_1}] = X_{\tau_1} \cdot 1_{\tau_1 \leq K} + X_K \cdot 1_{\tau_1 > K},
\]

where \( \forall A \in F_{\tau_1} \),

\[
\int_A E[(X_K - X_{\tau_2}) \cdot 1_{\tau_2 > K} | F_{\tau_1}](\omega) \, dP(\omega) = \int_A (X_K(\omega) - X_{\tau_2}(\omega)) \cdot 1_{\tau_2 > K}(\omega) \, dP(\omega) \to 0
\]

uniformly (in \( A \)) as \( K \to \infty \) by the DCT as

\[
\frac{1}{2} \cdot |X_K(\omega) - X_{\tau_2}(\omega)| \leq S(\omega) = \sup_{0 \leq k \leq \tau_2} |X_k(\omega)| \in L^1(\Omega, P)
\]

Finally,

\[
X_{\tau_1} \cdot 1_{\tau_1 \leq K} + X_K \cdot 1_{\tau_1 > K} \to X_{\tau_1} \quad \text{in } L^1(\Omega, P)
\]

as

\[
E[|X_K| \cdot 1_{\tau_1 > K}] \leq E[S \cdot 1_{\tau_1 > K}] \to 0
\]
Optional Stopping Theorem.

**Example.** Random walk on $\mathbb{Z}$. Take $p \in (0, 1)$, and let $\xi_1, \xi_2, \ldots$ be i.i.d. Bernoulli random variables such that

$$
\xi_j = \begin{cases} 
+1 & \text{with probability } p \\
-1 & \text{with probability } q = 1 - p
\end{cases}
$$

Consider integers $0 < x_0 < M$. Let $X_0 = x_0$ and $X_n = X_0 + \xi_1 + \ldots + \xi_n$. Then, the first hitting time

$$
\tau = \min\{t > 0 : X_t = 0 \text{ or } X_t = M\}
$$

is a stopping time w.r.t. filtration $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$.

We want to find $P(X_\tau = M)$.

If $p = \frac{1}{2}$, the random walk $X_n = X_0 + \xi_1 + \ldots + \xi_n$ is a martingale, and by part (b) of the Optional Stopping Theorem,

$$
P(X_\tau = M) = \frac{x_0}{M}
$$
Optional Stopping Theorem.

Example. Random walk on $\mathbb{Z}$. Take $p \in (0, 1)$, and let $\xi_1, \xi_2, \ldots$ be i.i.d. Bernoulli random variables such that

$$\xi_j = \begin{cases} +1 & \text{with probability } p \\ -1 & \text{with probability } q = 1 - p \end{cases}$$

Consider integers $0 < x_0 < M$. Let $X_0 = x_0$ and $X_n = X_0 + \xi_1 + \ldots + \xi_n$. Then, the first hitting time

$$\tau = \min\{t > 0 : X_t = 0 \text{ or } X_t = M\}$$

is a stopping time w.r.t. filtration $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$.

We want to find $P(X_\tau = M)$.

If $p \neq \frac{1}{2}$, then $X_n = X_0 + \xi_1 + \ldots + \xi_n$ is not a martingale, but $M_n = h(X_n)$ is a martingale when $h(x) = A \cdot \left(\frac{q}{p}\right)^x + B$ for any choice of constants $A$ and $B$. Then, taking $A \neq 0$, by part (b) of the Optional Stopping Theorem,

$$P(X_\tau = M) = \frac{h(x_0) - h(0)}{h(M) - h(0)} = \frac{1 - \left(\frac{q}{p}\right)^{x_0}}{1 - \left(\frac{q}{p}\right)^M}$$
Martingale Convergence Theorem.

Jensen's inequality: If $\varphi$ is a convex function, then
\[
E[\varphi(X)|\mathcal{G}] \geq \varphi(E[X|\mathcal{G}]) \quad a.s.
\]

Proposition. If $X_n$ is a submartingale w.r.t. $\mathcal{F}_n$ and $\varphi$ is an non-decreasing convex function with $E[|\varphi(X_n)|] < \infty$ for all $n$, then $\varphi(X_n)$ is a submartingale w.r.t. $\mathcal{F}_n$.

Proof. By Jensen's inequality,
\[
E[\varphi(X_{n+1})|\mathcal{F}_n] \geq \varphi(E[X_{n+1}|\mathcal{F}_n]) \geq \varphi(X_n) \quad a.s.
\]
Martingale Convergence Theorem.

Suppose $X_n$ is a submartingale:

$$E[X_{n+1} \mid \mathcal{F}_n] \geq X_n \quad \text{a.s.} \quad \forall n > 1$$

Let $a < b$ and let $N_0 = -1$,

$$N_{2k+1} = \inf\{n > N_{2k} : X_n \leq a\} \quad k = 0, 1, \ldots,$$

$$N_{2k} = \inf\{n > N_{2k-1} : X_n \geq b\} \quad k = 1, 2, \ldots$$

Then $N_j$ are stopping times,

$$\{N_{2k-1} < n \leq N_{2k}\} = \{N_{2k-1} \leq n - 1\} \cap \{N_{2k} \leq n - 1\}^c \in \mathcal{F}_{n-1}$$

and

$$H_n = \begin{cases} 1 & \text{if } N_{2k-1} < n \leq N_{2k} \text{ for some } k \geq 1 \\ 0 & \text{otherwise} \end{cases} \in \mathcal{F}_{n-1}$$

Such time intervals $[N_{2k-1}, N_{2k}]$ are called upcrossings.

Let $U_n = \sup\{k : N_{2k} \leq n\}$ denote the number of upcrossings by time $n$. 
Martingale Convergence Theorem.

\[ H_n = \begin{cases} 1 & \text{if } N_{2k-1} < n \leq N_{2k} \text{ for some } k \geq 1 \\ 0 & \text{otherwise} \end{cases} \in \mathcal{F}_{n-1} \]

Such time intervals \([N_{2k-1}, N_{2k}]\) are called upcrossings.

Let \( U_n = \sup\{k : N_{2k} \leq n\} \) denote the number of upcrossings by time \( n \).

The Upcrossing Inequality. If \( \{X_n\}_{n=0,1,...} \) is a submartingale, then

\[(b - a) \cdot E[U_n] \leq E[(X_n - a)^+] - E[(X_0 - a)^+]\]

Proof. Observe that \( Y_n = a + (X_n - a)^+ \) is also a submartingale, and it upcrosses \([a, b]\) the same number of times as \( X_n \) does, and therefore

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

as \((H \cdot Y)_n\) adds up the upcrossings \( Y(N_{2k}) - Y(N_{2k-1}) \geq b - a \) of \( Y \).

Finally, \((b - a) \cdot E[U_n] \leq E[(H \cdot Y)_n] \leq E[Y_n - Y_0] = E[(X_n - a)^+] - E[(X_0 - a)^+]\)
**Martingale Convergence Theorem.**

*Proof.* Observe that \( Y_n = a + (X_n - a)^+ \) is also a submartingale, and it upcrosses \([a, b]\) the same number of times as \( X_n \) does, and therefore

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

as \((H \cdot Y)_n\) adds up the upcrossings \( Y(N_{2k}) - Y(N_{2k-1}) \geq b - a\) of \( Y \).

Finally, \((b - a) \cdot E[U_n] \leq E[(H \cdot Y)_n] \leq E[Y_n - Y_0] = E[(X_n - a)^+] - E[(X_0 - a)^+]\)

as \( H_n \in \mathcal{F}_{n-1} \) and

\[
E[Y_n - Y_0] - E[(H \cdot Y)_n] = E \left[ \sum_{m=1}^{n} (1 - H_m) \cdot (Y_m - Y_{m-1}) \right]
\]

\[
= E \left[ \sum_{m=1}^{n} (1 - H_m) \cdot E[Y_m - Y_{m-1} \mid \mathcal{F}_{m-1}] \right] \geq 0
\]
Martingale Convergence Theorem.

The Martingale Convergence Theorem. Suppose $X_n$ is a submartingale such that

$$ \sup_n E[X_n^+] < \infty $$

Then, as $n \to \infty$, $X_n \to X \quad \text{a.s.}$

where $X \in L^1(\Omega, P)$.

Proof. From the Upcrossing Inequality, $\forall a < b,$

$$(b - a) \cdot E[U_n] \leq E[(X_n - a)^+] - E[(X_0 - a)^+]$$

and, as $(x - a)^+ \leq x^+ + |a|,$

$$E[U_n] \leq \frac{E[X_n^+] + |a|}{b - a}$$

Thus, since $\sup_n E[X_n^+] < \infty$, and since $U_n$ is an increasing sequence,

$U_n \uparrow U,$ \quad \text{where } E[U] < \infty \quad \text{and } U < \infty \quad \text{a.s.}$
Martingale Convergence Theorem.

Proof. From the Upcrossing Inequality, \( \forall a < b, \)
\[
(b - a) \cdot E[U_n] \leq E[(X_n - a)^+] - E[(X_0 - a)^+] 
\]
and, as \( (x - a)^+ \leq x^+ + |a|, \)
\[
E[U_n] \leq \frac{E[X_n^+] + |a|}{b - a} 
\]
Thus, since \( \sup_n E[X_n^+] < \infty, \) and since \( U_n \) is an increasing sequence,
\[
U_n \uparrow U, \quad \text{where } E[U] < \infty \text{ and } U < \infty \text{ a.s.} 
\]
Thus
\[
P\left( \bigcup_{a,b \in \mathbb{Q}} \{ \liminf_{n \to \infty} X_n < a < b < \limsup_{n \to \infty} X_n \} \right) = 0 
\]
and therefore
\[
\liminf_{n \to \infty} X_n = \limsup_{n \to \infty} X_n \quad \text{a.s.} 
\]
Martingale Convergence Theorem.

Proof. (continued)

\[ \liminf_{n \to \infty} X_n = \limsup_{n \to \infty} X_n \quad \text{a.s.} \]

Finally, we need to show that \( X = \lim_{n \to \infty} X_n \) is in \( L^1(\Omega, P) \).

By Fatou’s Lemma,

\[ E[X^+] \leq \liminf_{n \to \infty} E[X^n_+] < \infty \]

Now, since \( X_n \) is a submartingale,

\[ E[X^-] = E[X^+] - E[X_n] \leq E[X^+] - E[X_0], \]

and by Fatou’s Lemma,

\[ E[X^-] \leq \liminf_{n \to \infty} E[X^-] \leq \sup_{n} E[X^+] - E[X_0] < \infty \]

\( \square \)
Polya’s Urn.

We begin with $R_0$ red marbles and $G_0$ green marbles in the urn, at time $t = 0$. At each iteration, a marble is selected from the urn, uniformly at random. Then the marble is returned to the urn, and $D$ marbles of the same color as the selected marble are added into the urn.

Let $R_n$ and $G_n$ denote respectively the number of red and green marbles after $n$ iterations. Then the fraction of the red marbles at time $n$,

$$
\rho_n = \frac{R_n}{R_n + G_n}
$$

is a martingale:

$$
E[\rho_{n+1} \mid \mathcal{F}_n] = \frac{R_n + D}{R_n + G_n + D} \cdot \frac{R_n}{R_n + G_n} + \frac{R_n}{R_n + G_n + D} \cdot \frac{G_n}{R_n + G_n} = \frac{R_n}{R_n + G_n} = \rho_n
$$
Polya’s Urn.

The fraction of the red marbles at time \( n \),

\[
\rho_n = \frac{R_n}{R_n + G_n}
\]

is a **martingale**:

\[
E[\rho_{n+1} \mid \mathcal{F}_n] = \frac{R_n + D}{R_n + G_n + D} \cdot \frac{R_n}{R_n + G_n} + \frac{R_n}{R_n + G_n + D} \cdot \frac{G_n}{R_n + G_n} = \frac{R_n}{R_n + G_n} = \rho_n
\]

Thus, by the Martingale Convergence Theorem,

\[
\rho_n \to \rho_\infty \quad \text{a.s.}
\]

Here one can show that \( \rho_\infty \) is a beta random variable with parameters \((R_0 + D, G_0 + D)\) and density function

\[
f(x) = \frac{1}{B(R_0 + D, G_0 + D)} \cdot x^{R_0+D-1}(1 - x)^{G_0+D-1} \quad 0 \leq x \leq 1
\]