

MTH 664
Lectures 26 - 29

Yevgeniy Kovchegov
Oregon State University

Topics:

- Martingales.
- Filtration. Stopping times.
- Probability harmonic functions.
- Optional Stopping Theorem.
- Martingale Convergence Theorem.
- Maximal inequality.
- Doob-Kolmogorov inequality.

Conditional expectation.

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

Let $\mathcal{G} \subseteq \mathcal{F}$ be a smaller σ -algebra.

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

1. $E[X|\mathcal{G}]$ is \mathcal{G} -measurable
2. $\int_A E[X|\mathcal{G}] dP(\omega) = \int_A X dP(\omega)$ for all $A \in \mathcal{G}$

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 (Ω, \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- σ -algebras.

Lemma.

$$E[E[X|\mathcal{G}_2] | \mathcal{G}_1] = E[X|\mathcal{G}_1]$$

Proof. For any $A \in \mathcal{G}_1 \subseteq \mathcal{G}_2$,

$$\begin{aligned} \int_A E[E[X|\mathcal{G}_2] | \mathcal{G}_1](\omega) dP(\omega) &= \int_A E[X|\mathcal{G}_2](\omega) dP(\omega) \\ &= \int_A X(\omega) dP(\omega) = \int_A E[X|\mathcal{G}_1](\omega) dP(\omega) \end{aligned}$$

□

Filtration.

Definition. Consider an arbitrary linear ordered set T : A sequence of sub- σ -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_0, X_1, \dots on (Ω, \mathcal{F}, P) , and let $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$ is the smallest σ -algebra such that X_0, X_1, \dots, X_n are \mathcal{F}_n -measurable. Then \mathcal{F}_n is the smallest filtration that X_n is adapted to, i.e. $X_n \in \mathcal{F}_n$.

Important: When filtration \mathcal{F}_n is not mentioned in defining the martingale, submartingale, or supermartingale,

$$\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$$

Definition. Consider a filtration $\{\mathcal{F}_n\}$. A sequence of random variables $X_0, X_1, \dots \in L^1(\Omega, P)$ 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

$$E[X_{n+1} \mid \mathcal{F}_n] = X_n \quad a.s. \quad \forall n \geq 0$$

Martingales.

Definition. Consider a filtration $\{\mathcal{F}_n\}$. A sequence of random variables $X_0, X_1, \dots \in L^1(\Omega, P)$ 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

$$E[X_{n+1} \mid \mathcal{F}_n] = X_n \quad a.s. \quad \forall n \geq 0$$

Example. Let ξ_1, ξ_2, \dots be independent $L^1(\Omega, P)$ random variables such that

$$E[\xi_j] = 0 \quad \forall j \in \mathbb{N}$$

Now, let $X_0 = 0$ and $X_n = \xi_1 + \dots + \xi_n$ ($n \geq 1$). Then

$$E[X_{n+1} \mid \mathcal{F}_n] = E[X_n + \xi_{n+1} \mid \mathcal{F}_n] = X_n + E[\xi_{n+1} \mid \mathcal{F}_n] = X_n + E[\xi_{n+1}] = X_n$$

as ξ_{n+1} is independent of \mathcal{F}_n . Specifically, $\forall m \in \mathbb{N}$ s.t. $1 \leq m \leq n$, and any Borel $A \in \mathcal{B}$,

$$\begin{aligned} \int_{X_m^{-1}(A)} E[\xi_{n+1} \mid \mathcal{F}_n](\omega) dP(\omega) &= \int_{X_m^{-1}(A)} \xi_{n+1}(\omega) dP(\omega) = E[\xi_{n+1} \cdot \mathbf{1}_{X_m \in A}] \\ &= E[\xi_{n+1}] \cdot E[\mathbf{1}_{X_m \in A}] = \int_{X_m^{-1}(A)} E[\xi_{n+1}] dP(\omega) \end{aligned}$$

Martingales.

Definition. Consider a filtration $\{\mathcal{F}_n\}$. A sequence of random variables $X_0, X_1, \dots \in L^1(\Omega, P)$ 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

$$E[X_{n+1} \mid \mathcal{F}_n] = X_n \quad a.s. \quad \forall n \geq 0$$

Definition. Consider a filtration $\{\mathcal{F}_n\}$. A sequence of random variables $X_0, X_1, \dots \in L^1(\Omega, P)$ adapted to \mathcal{F}_n (i.e., $X_n \in \mathcal{F}_n$) is said to be a **supermartingale** with respect to $\{\mathcal{F}_n\}$ if

$$E[X_{n+1} \mid \mathcal{F}_n] \leq X_n \quad a.s. \quad \forall n \geq 0$$

Definition. Consider a filtration $\{\mathcal{F}_n\}$. A sequence of random variables $X_0, X_1, \dots \in L^1(\Omega, P)$ adapted to \mathcal{F}_n (i.e., $X_n \in \mathcal{F}_n$) is said to be a **submartingale** with respect to $\{\mathcal{F}_n\}$ if

$$E[X_{n+1} \mid \mathcal{F}_n] \geq X_n \quad a.s. \quad \forall n \geq 0$$

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_0, X_1, X_2, \dots with associated σ -algebras $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$.

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 ξ_1, ξ_2, \dots 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 $X_n = X_0 + \xi_1 + \dots + \xi_n$ is a martingale.

Suppose $p \neq \frac{1}{2}$, then $X_n = X_0 + \xi_1 + \dots + \xi_n$ 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- σ -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_0, X_1, \dots on (Ω, \mathcal{F}, P) , and let $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$. 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 τ 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 t is sufficient to determine whether $\{\tau \leq t\}$ occurred.

Filtration. Stopping time.

Definition. Consider an arbitrary linear ordered set T , and a filtration $\{\mathcal{F}_t\}_{t \in T}$. A random variable τ 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 t is sufficient to determine whether $\{\tau \leq t\}$ occurred.

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

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

Observe that if $\{\mathcal{F}_n\}$ is a filtration, and if X_0, X_1, \dots is a sequence of random variables adapted to \mathcal{F}_n , and τ is a stopping time w.r.t. $\{\mathcal{F}_n\}$, then

$$X_\tau = \sum_{j=1}^{\infty} X_j \cdot \mathbf{1}_{\tau=j} \in \mathcal{F}_\tau$$

Filtration. Stopping time.

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

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

Lemma. Suppose τ_1 and τ_2 are two stopping times w.r.t. \mathcal{F}_n such that $P(\tau_1 \leq \tau_2) = 1$, then

$$\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$$

Proof. Take $A \in \mathcal{F}_{\tau_1}$, 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 $A \cap \{\tau_2 \leq t\} \in \mathcal{F}_t$ as both $A \cap \{\tau_1 \leq t\}$ and $\{\tau_2 \leq t\}$ are in \mathcal{F}_t .

□

Filtration. Stopping time.

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

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

Lemma. Suppose τ 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

$$\begin{aligned} \int_A X_K(\omega) dP(\omega) &= \sum_{j=0}^K \int_{A \cap \{\tau=j\}} X_K(\omega) dP(\omega) = \sum_{j=0}^K \int_{A \cap \{\tau=j\}} E[X_K | \mathcal{F}_j](\omega) dP(\omega) \\ &= \sum_{j=0}^K \int_{A \cap \{\tau=j\}} X_j(\omega) dP(\omega) = \sum_{j=0}^K \int_{A \cap \{\tau=j\}} X_\tau(\omega) dP(\omega) = \int_A X_\tau(\omega) dP(\omega) \end{aligned}$$

□

Optional Stopping Theorem.

Doob's Optional Stopping Theorem. Consider a sequence of random variables X_1, X_2, \dots on (Ω, \mathcal{F}, P) , and let $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$. Suppose τ_1 and τ_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,

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

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

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

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

□

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

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

Therefore

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

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

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

Therefore

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

where $\forall A \in \mathcal{F}_{\tau_1}$,

$$\int_A E[(X_K - X_{\tau_2}) \cdot \mathbf{1}_{\tau_2 > K} | \mathcal{F}_{\tau_1}](\omega) dP(\omega) = \int_A (X_K(\omega) - X_{\tau_2}(\omega)) \cdot \mathbf{1}_{\tau_2 > K}(\omega) dP(\omega) \rightarrow 0$$

uniformly (in A) as $K \rightarrow \infty$ 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,

$$X_{\tau_1} \cdot \mathbf{1}_{\tau_1 \leq K} + X_K \cdot \mathbf{1}_{\tau_1 > K} \rightarrow X_{\tau_1} \quad \text{in } L^1(\Omega, P)$$

as

$$E[|X_K| \cdot \mathbf{1}_{\tau_1 > K}] \leq E[S \cdot \mathbf{1}_{\tau_1 > K}] \rightarrow 0$$

□

Optional Stopping Theorem.

Example. Random walk on \mathbb{Z} . Take $p \in (0, 1)$, and let ξ_1, ξ_2, \dots 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 $X_n = X_0 + \xi_1 + \dots + \xi_n$. 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 $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$.

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

If $p = \frac{1}{2}$, the random walk $X_n = X_0 + \xi_1 + \dots + \xi_n$ 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 ξ_1, ξ_2, \dots 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 $X_n = X_0 + \xi_1 + \dots + \xi_n$. 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 $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$.

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

If $p \neq \frac{1}{2}$, then $X_n = X_0 + \xi_1 + \dots + \xi_n$ 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 φ 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 φ 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 $N_0 = -1$,

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

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

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 $\{X_n\}_{n=0,1,\dots}$ 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

$$\begin{aligned} E[Y_n - Y_0] - E[(H \cdot Y)_n] &= E \left[\sum_{m=1}^n (1 - H_m) \cdot (Y_m - Y_{m-1}) \right] \\ &= E \left[\sum_{m=1}^n (1 - H_m) \cdot E[Y_m - Y_{m-1} | \mathcal{F}_{m-1}] \right] \geq 0 \end{aligned}$$

□

Martingale Convergence Theorem.

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

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

Then, as $n \rightarrow \infty$,

$$X_n \rightarrow X \quad a.s.$$

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,

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

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

$$P\left(\bigcup_{a,b \in \mathbb{Q}} \left\{ \liminf_{n \rightarrow \infty} X_n < a < b < \limsup_{n \rightarrow \infty} X_n \right\}\right) = 0$$

and therefore

$$\liminf_{n \rightarrow \infty} X_n = \limsup_{n \rightarrow \infty} X_n \quad \text{a.s.}$$

Martingale Convergence Theorem.

Proof. (continued)

$$\liminf_{n \rightarrow \infty} X_n = \limsup_{n \rightarrow \infty} X_n \quad a.s.$$

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

By Fatou's Lemma,

$$E[X^+] \leq \liminf_{n \rightarrow \infty} E[X_n^+] < \infty$$

Now, since X_n is a submartingale,

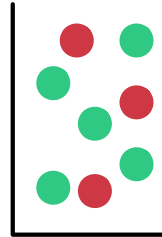
$$E[X_n^-] = E[X_n^+] - E[X_n] \leq E[X_n^+] - E[X_0],$$

and by Fatou's Lemma,

$$E[X^-] \leq \liminf_{n \rightarrow \infty} E[X_n^-] \leq \sup_n E[X_n^+] - E[X_0] < \infty$$

□

Polya's Urn.



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.

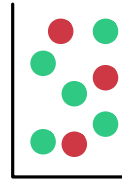
Let R_n and G_n denote respectively the number of red and green marbles after n iterations. Then the fraction of the red marbles at time n ,

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

is a **martingale**:

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

Polya's Urn.



Polya's Urn

The fraction of the red marbles at time n ,

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

is a **martingale**:

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

Thus, by the Martingale Convergence Theorem,

$$\rho_n \rightarrow \rho_\infty \quad a.s.$$

Here one can show that ρ_∞ is a beta random variable with parameters $(R_0/D, G_0/D)$ and density function

$$f(x) = \frac{1}{\mathcal{B}(R_0/D, G_0/D)} \cdot x^{R_0/D-1} (1-x)^{G_0/D-1} \quad 0 \leq x \leq 1$$

Submartingales. Doob's decomposition.

Definition. Consider a filtration $\{\mathcal{F}_n\}$. A sequence of random variables $X_0, X_1, \dots \in L^1(\Omega, P)$ adapted to \mathcal{F}_n (i.e., $X_n \in \mathcal{F}_n$) is said to be a **submartingale** with respect to $\{\mathcal{F}_n\}$ if

$$E[X_{n+1} \mid \mathcal{F}_n] \geq X_n \quad a.s. \quad \forall n \geq 0$$

Definition. Consider a filtration $\{\mathcal{F}_n\}$. A sequence of random variables S_n adapted to \mathcal{F}_n is said to be a **predictable** if S_n is \mathcal{F}_{n-1} measurable for all $n \geq 1$. A predictable sequence of random variables (process) S_n adapted to \mathcal{F}_n is said to be **increasing** if $S_0 = 0$ and $P(S_{n-1} \leq S_n) = 1$ for all $n \geq 1$.

Doob's decomposition. A submartingale $X_0, X_1, \dots \in L^1(\Omega, P)$ adapted to filtration \mathcal{F}_n can be expressed uniquely as

$$X_n = M_n + S_n,$$

where M_n is a **martingale** adapted to \mathcal{F}_n , and S_n is an **increasing predictable** sequence (process).

Submartingales. Doob's decomposition.

Definition. Consider a filtration $\{\mathcal{F}_n\}$. A sequence of random variables S_n adapted to \mathcal{F}_n is said to be a **predictable** if S_n is \mathcal{F}_{n-1} measurable for all $n \geq 1$. A predictable sequence of random variables (process) S_n adapted to \mathcal{F}_n is said to be **increasing** if $S_0 = 0$ and $P(S_{n-1} \leq S_n) = 1$ for all $n \geq 1$.

Doob's decomposition. A submartingale $X_0, X_1, \dots \in L^1(\Omega, P)$ adapted to filtration \mathcal{F}_n can be expressed uniquely as

$$X_n = M_n + S_n,$$

where M_n is a **martingale** adapted to \mathcal{F}_n , and S_n is an **increasing predictable** sequence (process).

Proof. Let $M_0 = X_0$ and $S_0 = 0$. For all integer $n \geq 1$, let

$$M_n - M_{n-1} = X_n - E[X_n | \mathcal{F}_{n-1}] \quad \text{and} \quad S_n - S_{n-1} = E[X_n | \mathcal{F}_{n-1}] - X_{n-1}.$$

Then, $E[M_n - M_{n-1} | \mathcal{F}_{n-1}] = 0$, $S_n - S_{n-1} \in \mathcal{F}_{n-1}$, and $S_n - S_{n-1} \geq 0$ *a.s.* □

Maximal inequality.

Maximal inequality. If X_n is a submartingale adapted to filtration \mathcal{F}_n , then

$$P\left(\max_{j: j \leq n} X_j \geq a\right) \leq \frac{E[X_n^+]}{a} \quad \forall a > 0.$$

Proof. Recall that X_n^+ is a nonnegative submartingale adapted to \mathcal{F}_n . For a given $a > 0$, let

$$T_a = \min\{k : X_k \geq a\} = \min\{k : X_k^+ \geq a\}.$$

Next, for a fixed n , submartingale X_k^+ with $\tau_1 = T_a \wedge n$ and $\tau_2 = n$ satisfy the conditions in the **Optional Stopping Theorem** (for submartingales). Hence, $X_n^+ = E[X_{\tau_2}^+ | \mathcal{F}_{\tau_1}] \geq X_{\tau_1}^+$, and therefore,

$$E[X_n^+] = E[X_{\tau_2}^+] \geq E[X_{\tau_1}^+] = E[X_{T_a \wedge n}^+].$$

Thus,

$$E[X_n^+] \geq E[X_{T_a \wedge n}^+] = E[X_{T_a}^+ \mathbf{1}_{T_a \leq n}] + E[X_n^+ \mathbf{1}_{T_a > n}] \geq a P(T_a \leq n) + E[X_n^+ \mathbf{1}_{T_a > n}].$$

$$\text{Finally, } P\left(\max_{j: j \leq n} X_j \geq a\right) = P(T_a \leq n) \leq \frac{E[X_n^+] - E[X_n^+ \mathbf{1}_{T_a > n}]}{a}.$$

□

Maximal and Doob-Kolmogorov inequalities.

Maximal inequality. If X_n is a submartingale adapted to filtration \mathcal{F}_n , then

$$P\left(\max_{j: j \leq n} X_j \geq a\right) \leq \frac{E[X_n^+]}{a} \quad \forall a > 0.$$

Similarly, if X_n is a supermartingale adapted to \mathcal{F}_n , then

$$P\left(\max_{j: j \leq n} X_j \geq a\right) \leq \frac{E[X_0] + E[X_n^-]}{a} \quad \forall a > 0.$$

Doob-Kolmogorov inequality. If X_n is a martingale adapted to filtration \mathcal{F}_n such that $E[X_n^2] < \infty$, then

$$P\left(\max_{j: 0 \leq j \leq n} |X_j| \geq a\right) \leq \frac{E[X_n^2]}{a^2} \quad \forall a > 0.$$

Proof. First, notice that by Jensen's inequality, X_n^+ is a submartingale, and so is $X_n^- = (-X_n)^+$. Hence, by Jensen's inequality, $X_n^2 = (X_n^+)^2 + (X_n^-)^2$ is also a submartingale.

Therefore,
$$P\left(\max_{j: 0 \leq j \leq n} |X_j| \geq a\right) = P\left(\max_{j: 0 \leq j \leq n} X_j^2 \geq a^2\right) \leq \frac{E[X_n^2]}{a^2}.$$

□

Kolmogorov inequality via martingales.

Maximal inequality. If X_n is a submartingale adapted to filtration \mathcal{F}_n , then

$$P\left(\max_{j: j \leq n} X_j \geq a\right) \leq \frac{E[X_n^+]}{a} \quad \forall a > 0.$$

Doob-Kolmogorov inequality. If X_n is a martingale adapted to filtration \mathcal{F}_n such that $E[X_n^2] < \infty$, then

$$P\left(\max_{j: 0 \leq j \leq n} |X_j| \geq a\right) \leq \frac{E[X_n^2]}{a^2} \quad \forall a > 0.$$

An alternative proof of the **Kolmogorov's inequality** follows.

Kolmogorov's Maximal Inequality. Let $S_n = X_1 + \dots + X_n$. If X_1, X_2, \dots are independent random variables in $L^2(\Omega, P)$, then $\forall \lambda > 0$ and any $n \in \mathbb{N}$,

$$P\left(\max_{1 \leq k \leq n} |S_k - E[S_k]| \geq \lambda\right) \leq \frac{\text{Var}(S_n)}{\lambda^2}$$