# Martingales

## Defn (Martingale | Submartingale | Supermartingale)

A real-valued stochastic process $\{X_n\}_{n \geq 0}$ s.t. $\forall n \geq 0$

i) $X_n$ is adapted to an underlying filtration $\mathcal{F}_n$

ii) $\mathbb{E}[|X_n|] < \infty$

is said to be a

- martingale if $\mathbb{E}[X_{n+1} | \mathcal{F}_n] = X_n \ \forall n$
- Submartingale if $\mathbb{E}[X_{n+1} | \mathcal{F}_n] \geq X_n \ \forall n$
- Supermartingale if $\mathbb{E}[X_{n+1} | \mathcal{F}_n] \leq X_n \ \forall n$

- $X_n$ is said to be a $\mathcal{F}_n$-martingale

---

Eg - $Y_n$ indep, $\mathbb{E}[Y_n] = 0 \ \forall n$, $S_n = Y_0 + Y_1 + \dots + Y_n$

Let $\mathcal{F}_n = \sigma(Y_0, \dots, Y_n)$, then $S_n$ is a $\mathcal{F}_n$-martingale

To see this, note $\mathbb{E}[S_{n+1} | \mathcal{F}_n] = \mathbb{E}[Y_{n+1} + S_n | \mathcal{F}_n] = S_n$

- Now let $\text{Var}(Y_i) = \sigma^2$. Then $M_n = S_n^2 - n\sigma^2$ is a $\mathcal{F}_n$-martingale, as $\mathbb{E}[M_{n+1} | \mathcal{F}_n] = \mathbb{E}[S_{n+1}^2 - (n+1)\sigma^2 | \mathcal{F}_n]$

$$= S_n^2 - n\sigma^2 + \mathbb{E}[Y_{n+1}^2 + 2\underbrace{Y_{n+1} S_n}_{Y_{n+1} \perp\!\!\!\perp S_n} - \sigma^2 | \mathcal{F}_n]$$

$$= S_n^2 - n\sigma^2 = M_n$$

- Finally suppose $\mathbb{E}[e^{\theta Y_i}] < \infty$ for some $\theta$. Then we have that $G_n = e^{\theta S_n} / \mathbb{E}[e^{\theta Y}]^n$ is a $\mathcal{F}_n$-martingale, as

$$\mathbb{E}[G_{n+1} | \mathcal{F}_n] = \mathbb{E}\left[ e^{\theta S_n} e^{\theta Y_{n+1}} / \mathbb{E}[e^{\theta Y_i}]^{n+1} \middle| \mathcal{F}_n \right] = G_n \frac{\cancel{\mathbb{E}[e^{\theta Y_{n+1}}]}}{\cancel{\mathbb{E}[e^{\theta Y_i}]}} = G_n$$

---

Eg - $X_n$ ind, $Z = f(X_1, X_2, \dots, X_n)$, $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$

Let $Y_i = \mathbb{E}[Z | \mathcal{F}_i]$, $i \in \{0, 1, \dots n\}$

(So $Y_0 = \mathbb{E}[Z]$, $Y_n = \mathbb{E}[Z | \mathcal{F}_n] = f(X_1, \dots, X_n)$)

$Y_i$ is a $\mathcal{F}_n$-martingale (Doob Martingale)

$$- \mathbb{E}[Y_{k+1} | \mathcal{F}_k] = \mathbb{E}\big[\mathbb{E}[Z | \mathcal{F}_{k+1}] \big| \mathcal{F}_k\big]$$
$$= \mathbb{E}[Z | \mathcal{F}_k] = Y_k \quad \text{(tower rule)}$$

- For example, consider $N_0 \equiv \#$ of isolated vertices in a $G_{np}$ random graph. Let $(1, 2, \dots, \binom{n}{2})$ be an ordering of potential edges (ie, of $V^2$) and $X_i = \mathbb{1}\{\text{edge } i \text{ is present}\} \sim \text{Ber}(p)$.
  Then $Y_0 = \mathbb{E}[N_0]$, $Y_k = \mathbb{E}[N_0 | X_1, X_2, \dots, X_k]$, $Y_{\binom{n}{2}} = N_0$

- Similarly, suppose we drop $m$ balls in $n$ bins, and let $N_0 = \#$ of empty bins. Let $X_i = \mathbb{1}\{i^{\text{th}} \text{ ball falls in empty bin}\} \Rightarrow$
  $X_i \sim \text{Ber}(Y_i/n)$, where $Y_i = \#$ of empty bins when $i^{\text{th}}$ ball is dropped.

- Let $Z_k = E[N_0 | X_1, \dots, X_k] \Rightarrow Z_0 = \mathbb{E}[N_0] = n(1-1/n)^m$

$Z_m = N_0$, and $Z_k = E[N_0 | Y_k] = Y_k (1-1/n)^{m-k}$

By the earlier example, we have $Z_m$ is a martingale

---

- Defn (Martingale Difference Sequence) A rv sequence $\{D_n\}_{n \geq 1}$ relative to filtration $\{\mathcal{F}_n\}_{n \geq 0}$ s.t.

  i) $\{D_n\}$ is adapted to $\{\mathcal{F}_n\}$

  ii) $\mathbb{E}[|D_n|] < \infty \ \forall n \geq 1$

  iii) $E[D_n | \mathcal{F}_{n-1}] = 0$ a.s $\forall n \geq 1$

- A random process $\{H_n\}_{n \geq 1}$ is said to be predictable wrt filtration $\{\mathcal{F}_n\}_{n \geq 0}$ if $H_n$ is $\mathcal{F}_{n-1}$ measurable

- (Doob Decomposition) For any random sequence $\{Y_n\}_{n \geq 1}$ with $E[|Y_i|] < \infty$, and adapted to some filtration $\mathcal{F}_n$, we can write $Y_n = H_n + D_n$, where $H_n = E[Y_n | \mathcal{F}_{n-1}]$ is a predictable process, and $D_n = Y_n - H_n$ is a martingale difference seqn wrt $\{\mathcal{F}_n\}$.

  $-E[D_n | \mathcal{F}_{n-1}] = E[Y_n - E[Y_n | \mathcal{F}_{n-1}] | \mathcal{F}_{n-1}] = 0$

- (Martingale transform) If $\{D_n\}_{n \geq 1}$ is a martingale difference seqn, and $\{H_n\}_{n \geq 1}$ is a predictable process wrt $\{\mathcal{F}_n\}$, then $(H \circ D)_n = H_n D_n$ is a martingale difference seqn wrt $\mathcal{F}_n$.

- Interpretation - $H_n$ is the 'dollar amount' bet in a gamble, and $D_n$ is the net gain per dollar of the gamble. $H_n$ being predictable $\Rightarrow$ it can not 'know' the outcome of the gamble. Thus, if a gambler started with $M_0$, then after $k$ rounds she has $M_k = M_0 + \sum_{i=1}^{k} H_i D_i$, where $M_k$ is a martingale wrt $\{\mathcal{F}_n\}$.

- Also
$$\mathbb{E}[(H_k D_k)^2] = \mathbb{E}\left[\mathbb{E}[(H_k D_k)^2 \mid \mathcal{F}_{k-1}]\right]$$
$$= \mathbb{E}[H_k^2\, \mathbb{E}[D_k^2 \mid \mathcal{F}_{k-1}]]$$
$$\Rightarrow \mathbb{E}[(M_k - M_0)^2] = \sum_{i=1}^{k} \mathbb{E}[H_i^2\, \mathbb{E}[D_i^2 \mid \mathcal{F}_i]]$$

---

- (Stopped Martingale) Let $\{Y_n\}_{n\geq 0}$ be a $\mathcal{F}_n$-martingale (submartingale) and $\tau$ a stopping time wrt $\{\mathcal{F}_n\}$. Then $Z_n = Y_{n\wedge\tau}$, $n \geq 0$ is a $\mathcal{F}_n$-martingale (submartingale), and
$$\mathbb{E}[Y_{n\wedge\tau}] \geq \mathbb{E}[Y_0] \quad , n \geq 0$$

Pf - Let $H_n = \mathbb{1}\{n \leq \tau\}$, which is $\mathcal{F}_n$-predictable. Thus
$$Y_{n\wedge\tau} = Y_0 + \sum_{k=1}^{n\wedge\tau} (Y_k - Y_{k-1})$$
$$= Y_0 + \sum_{k=1}^{n} H_k (Y_k - Y_{k-1}) \equiv \text{martingale transform}$$

# Azuma-Hoeffding Inequality

The first important use of martingales is in generalizing the tail bounds (Chernoff-Hoeffding) we saw for iid r.v.s

We first need a lemma -

**Lemma** - If $D$ is a rv with $\mathbb{E}[D]=0$ and $\mathbb{P}[|D|\le d]=1$ for constant $d\ge 0$ (ie, $D\in[-d,d]$ a.s). Then for any $a\in\mathbb{R}$

$$\mathbb{E}[e^{aD}] \le e^{(ad)^2/2}$$

Pf - Since $e^{ax}$ is convex for any $a\in\mathbb{R}$, we have that

$$e^{ax} \le \frac{d-x}{2d}e^{ad} + \frac{x+d}{2d}e^{-ad} \quad \forall x\in[-d,d]$$

$$\Rightarrow \mathbb{E}[e^{aD}] \le \mathbb{E}\left[\frac{d-D}{2d}e^{ad} + \frac{d+D}{2d}e^{-ad}\right] = \frac{e^{-ad}+e^{ad}}{2}$$

Now consider $f(x) = \ln\left(\frac{e^{x}+e^{-x}}{2}\right)$, $f(0)=0$, $f'(x)=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$, $f'(0)=0$

$$f''(x) = \frac{(e^{x}+e^{-x})^2-(e^{x}-e^{-x})^2}{(e^{x}+e^{-x})^2} = \frac{4}{(e^{x}+e^{-x})^2} \le 1 \ \forall x\in[0,1]$$

for some $\xi\in[0,1]$

$$\Rightarrow f(x) \le f(0)+f'(0)x+f''(\xi)x^2/2 \le x^2/2 \ \forall x\in[0,1]$$

- Finally $\mathbb{E}[e^{aD}] \le \exp(f(ad)) \le e^{a^2d^2/2}$ ∎

We are now ready to prove the main tail bound

**Thm** (Azuma-Haeffding Inequality) Suppose $\{M_n\}_{n\geq 0}$ is a martingale wrt $\{\mathcal{F}_n\}$, s.t. $\mathbb{P}[|M_{n+1} - M_n| \leq d_n] = 1 \ \forall n \geq 0$

Then.
$$\mathbb{P}[|M_n - M_0| \geq \lambda] \leq 2\exp\left(\frac{-\lambda^2}{2\sum_{i=1}^n d_i^2}\right)$$

---

Pf - $\forall a \in \mathbb{R}$, $\mathbb{E}[e^{a(M_n - M_0)}] = \mathbb{E}[\mathbb{E}[e^{a(M_n - M_{n-1} + M_{n-1} - M_0)} | \mathcal{F}_{n-1}]]$

$$= \mathbb{E}[\mathbb{E}[e^{a(M_n - M_{n-1})} | \mathcal{F}_{n-1}] e^{a(M_{n-1} - M_0)}]$$

Note - $\mathbb{E}[M_n - M_{n-1} | \mathcal{F}_{n-1}] = 0$, $|M_n - M_{n-1}| \leq d_n$ a.s, and so $M_n - M_{n-1} | \mathcal{F}_{n-1} \leq ds$

$$\leq e^{(a d_n)^2/2} \mathbb{E}[e^{a(M_{n-1} - M_0)}]$$

Iterating, we get $\mathbb{E}[e^{a(M_n - M_0)}] \leq e^{a^2 \sum_{i=1}^n d_i^2 / 2}$

. Now via the usual Chernoff bound, we have

$$\mathbb{P}[M_n - M_0 \geq \lambda] \leq \mathbb{E}[e^{a(M_n - M_0)}] e^{-a\lambda} \leq e^{a^2 (\sum d_i^2)/2 - a\lambda} \quad \forall a \in \mathbb{R}_+$$

Setting $a = \lambda / \sum d_i^2$, we get $\mathbb{P}[M_n - M_0 \geq \lambda] \leq e^{-\lambda^2 / 2\sum_{i=1}^n d_i^2}$

Repeating for $\mathbb{P}[M_n - M_0 \leq -\lambda]$ (with $M_n$ replaced with $-M_n$) we get the result ∎

---

Note - If $B_k \leq M_k - M_{k-1} \leq B_k + d_k$ for some predictable process $B_k$, then we can tighten this to $\mathbb{P}[|M_n - M_0| > \lambda] \leq 2e^{-2\lambda^2/(\sum d_k^2)}$.

(Try to see where this modifies the proof)

Eg - Let $X_1, X_2, \ldots$ be indep r.v, $\mathbb{E}[X_i] = p_i$, $|X_i| \leq d_i$ a.s.
Now suppose $S_n = \sum_{i=1}^{n} X_i$ and $M_n = S_n - \sum_{i=1}^{n} p_i$. Then
$M_n$ is a martingale, and thus via Azuma-Hoeffding

(Hoeffding's Inequality) $$\mathbb{P}\left[\left|S_n - \sum_{i=1}^{n} p_i\right| > t\right] \leq 2\exp\left(\frac{-t^2}{2\sum_{i=1}^{n} d_i^2}\right)$$

---

- Def (Bounded Differences) - A fn $f: \mathcal{X}^n \to \mathbb{R}$ is said to satisfy bounded differences if $\forall i \in [n]$, $x_1, x_2, \ldots, x_n, y \in \mathcal{X}$

$$|f(x_1, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, x_i, y, x_{i+1}, \ldots, x_n)| \leq c$$

- Thm (McDiarmid's Inequality) - Suppose $f$ satisfies bounded differences condition with parameter $c$, and $X_1, \ldots, X_n$ are independent r.v. Then

$$\mathbb{P}\left[\left|f(X_1, \ldots, X_n) - \mathbb{E}[f(X_1, \ldots, X_n)]\right| > \lambda\right] \leq 2e^{-2\lambda^2/nc^2}$$

Pf - Let $M_i = \mathbb{E}[f(X_1, \ldots, X_n) \mid X_1, \ldots, X_i]$ $\forall i \in \{0, \ldots, n\}$
$M_0 = \mathbb{E}[f(X_1, \ldots, X_n)]$, $M_n = f(X_1, \ldots, X_n)$ and $M_i$ is a Doob martingale (wrt $\mathcal{F}_i = \sigma(X_1, \ldots, X_i)$).
Now we want to use Azuma-Hoeffding. Note that $M_k - M_{k-1}$
$= \mathbb{E}[f(X_1^n) \mid X_1^k] - \mathbb{E}[f(X_1^n) \mid X_1^{k-1}] \leq \sup_x \left(\mathbb{E}[f(X_1^n) \mid X_1^{k-1}, X_k = x] - \mathbb{E}[f(X_1^n) \mid X_1^{k-1}]\right)$

Similarly $M_k - M_{k-1} \geqslant \inf_x \left( \mathbb{E}[f(x^n) | X_1^{k-1}, X_k = x] - \mathbb{E}[f(X_1^n) | X_1^{k-1}] \right)$

Claim - $\sup_x \left( \mathbb{E}[f(x^n) | X_1^{k-1}, X_k = x] - \mathbb{E}[f(X_1^n) | X_1^{k-1}] \right) - \inf_y \left( \mathbb{E}[f(x^n) | X_1^{k-1}, X_k = x] - \mathbb{E}[f(X_1^n) | X_1^{k-1}] \right) < c$

This follows by applying the bounded difference property for all possible values of $(X_{k+1}, X_{k+2}, \dots, X_n)$, and then taking expectations.

The theorem now follows from using the stronger version of A-H ineq. ∎

---

Eg - Let $I = \{1, 2, \dots, \binom{n}{2}\}$ be an enumeration of all edges in a $G_{n,p}$ random graph $G$, and Let $\alpha(G)$ denote the size of its largest independent set. Now consider the Doob martingale $F_k = \mathbb{E}[\alpha(G) | X_1, X_2, \dots, X_k]$, where $X_i \equiv \mathbb{1}\{\text{edge } i \text{ is present}\} \sim \text{Ber}(p)$, ind, for $k \in [\binom{n}{2}]$.

Now note that $\alpha(G)$ Changes by at most 1 given information about a single edge $\Rightarrow$ it Satisfies the bounded differences property with $c = 1$. Thus

$$\mathbb{P}\left[ |\alpha(G) - \mathbb{E}[\alpha(G)]| \geqslant \lambda \right] \leq 2 \exp\left( \frac{-4\lambda^2}{n(n-1)} \right)$$

- Can this be tightened? Consider $V = \{1, \dots, n\}$ be an ordering of vertices of $G$, and define $Y_i \equiv \{ j \in [i-1], (i,j) \in E \}$ be the set of edges of $i$ to preceeding nodes in $V$.

Now let $\hat{F}_k^n = \mathbb{E}[\alpha(G) \mid Y_1^k]$. This is again a martingale, and again, $\alpha(G)$ changes at most by 1 given $Y_k$! Thus by AH-

$$\mathbb{P}\left[ \left| \alpha(G) - \mathbb{E}[\alpha(G)] \right| \geq \lambda \right] \leq 2\exp\left( -\frac{2\lambda^2}{n} \right) \quad \blacksquare$$

# Optional Stopping

- $\tau$ is a stopping time w.r.t filtration $(\mathcal{F}_n)_{n\geq 0}$ if (recall)
$$\{T \leq n\} \in \mathcal{F}_n \quad \forall\, n \geq 0$$
- Earlier we saw that for any $\mathcal{F}_t$-martingale $M_t$, for fixed $n \in \mathbb{N}$, we have -
$$\mathbb{E}[M_n] = \mathbb{E}[\mathbb{E}[M_n | \mathcal{F}_{n-1}]] = \mathbb{E}[M_{n-1}] = \dots = \mathbb{E}[M_0]$$

Q: Is $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$ for a stopping time $\tau$?

Eg - Let $X_i \sim Ber(1/2)$, iid, $S_n = \sum_{i=1}^{n} X_i$

Define $\tau = \inf\{t \geq 0 \mid S_t = b\}$

Then $\mathbb{E}[M_\tau] = b \neq \mathbb{E}[M_0] = 0$

---

Lemma (Stopped Martingale) - If $M_n$ is a martingale and $\tau$ a stopping time wrt filtration $(\mathcal{F}_n)_{n\geq 0}$. Then $\forall\, n \in \mathbb{N}$, we have
$$\mathbb{E}[M_{n\wedge\tau}] = \mathbb{E}[M_0]$$

Pf - $M_{(n+1)\wedge\tau} - M_{n\wedge\tau} = (M_{n+1} - M_n)\, \mathbb{1}\{\tau > n\}$

$$\Rightarrow \mathbb{E}[M_{(n+1)\wedge\tau} - M_{n\wedge\tau}] = \mathbb{E}[\mathbb{E}[(M_{n+1} - M_n)\mathbb{1}\{\tau > n\} | \mathcal{F}_n]]$$

$$\xrightarrow{\mathbb{1}\{\tau>n\} \text{ is adapted to } \mathcal{F}_n} = \mathbb{E}[(\mathbb{E}[M_{n+1}|\mathcal{F}_n] - M_n)\mathbb{1}\{\tau > n\}]$$

$$\xrightarrow{M_n \text{ is a martingale}} = 0$$

Thus $\mathbb{E}[M_{(n+1)\wedge\tau}] = \mathbb{E}[M_{n\wedge\tau}]\ \forall n \geq 0 \Rightarrow \mathbb{E}[M_{n\wedge\tau}] = \mathbb{E}[M_0]$

• Thus, for $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$, we need conditions to ensure $\lim_{n\to\infty} \mathbb{E}[M_{\tau\wedge n}] = \mathbb{E}[M_\tau]$

**Thm** (*Optional Stopping*) If $M_n$ is a martingale and $\tau$ a stopping time wrt filtration $(\mathcal{F}_n)_{n\geq 0}$. Then $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$ if one of the following is true:

i) $\mathbb{P}[\tau \leq n_0] = 1$

ii) $|M_n| \leq c$ a.s $\forall n$

iii) $\mathbb{E}[\tau] < \infty$ & $\mathbb{E}[|M_{n+1} - M_n| \mid \mathcal{F}_n] < c$

Pf - From before, we have $M_\tau = M_{\tau\wedge n} + (M_\tau - M_n)\mathbb{1}_{\{\tau > n\}}$ and also $\mathbb{E}[M_{\tau\wedge n}] = \mathbb{E}[M_0]$ $\forall n$.

Thus we have $\mathbb{E}[M_\tau] = \mathbb{E}[M_0] + \mathbb{E}[(M_\tau - M_n)\mathbb{1}_{\{\tau > n\}}]$

i) If $\mathbb{P}[\tau \leq n_0] = 1$ for some $n_0$, then $\forall n \geq n_0$, $\tau \wedge n = \tau$ a.s

ii) $|M_{n\wedge\tau}| \leq c$ a.s $\forall n \Rightarrow$ by the dominated convergence theorem, $\lim_{n\to\infty} \mathbb{E}[M_{\tau\wedge n}] = \mathbb{E}[M_\tau]$

iii) Let $Y_\tau = |M_0| + |M_1 - M_0| + \dots + |M_\tau - M_{\tau-1}| \Rightarrow M_{\tau\wedge n} \leq Y_\tau \; \forall n \geq 0$

Also
$$\mathbb{E}[Y_\tau] = \mathbb{E}[|M_0|] + \sum_{i=1}^{\infty} \mathbb{E}[|M_i - M_{i-1}|\mathbb{1}_{\{i \leq \tau\}}]$$
$$= \mathbb{E}[|M_0|] + \sum_{i=1}^{\infty} \mathbb{E}[\mathbb{E}[|M_i - M_{i-1}|\mathbb{1}_{\{i\leq\tau\}} \mid \mathcal{F}_{i-1}]]$$
$$\leq \mathbb{E}[|M_0|] + c\sum_{i=1}^{\infty} \mathbb{P}[i \leq \tau] < \infty$$

Thus we can now use dominated convergence to get $\lim_{n\to\infty} \mathbb{E}[M_{\tau\wedge n}] = \mathbb{E}[M_\tau]$

$\blacksquare$

Eg - $X_i = \mp 1$ wp $1/2$, $S_n = \sum_{i=1}^{n} X_i$ (Gambler's Ruin)

- For any $a, b \in \mathbb{N}$, let $\tau_{ab} = \min\{n \mid S_n = -a \text{ or } S_n = b\}$
- $S_n$ is a martingale $\Rightarrow S_{n \wedge \tau}$ is a martingale
- $\mathbb{E}[\tau] < \infty$ and $\mathbb{E}[|S_{n+1} - S_n| \mid \mathcal{F}_n] < 1$

$\Rightarrow$ We can use optional stopping. Let $p_a = \mathbb{P}[S_{\tau_{ab}} = -a]$

$p_b = 1 - p_a = \mathbb{P}[S_{\tau_{ab}} = b]$

$\Rightarrow \mathbb{E}[S_\tau] = -p_a a + p_b b = \mathbb{E}[S_0] = 0 \Rightarrow p_a = b/a+b$

- $M_n = S_n^2 - n$ is a martingale, $\mathbb{E}[|M_{n+1} - M_n| \mid \mathcal{F}_n] < \infty$

$\Rightarrow \mathbb{E}[M_\tau] = p_a\left(a^2 - \mathbb{E}[\tau \mid S_\tau = -a]\right) + p_b\left(b^2 - \mathbb{E}[\tau \mid S_\tau = b]\right) = \mathbb{E}[M_0] = 0 \Rightarrow \mathbb{E}[\tau] = ab$

---

Eg - (Harmonic fns of MCs) Let $X_n$ be an ergodic MC with transition matrix $P$ on countable space $\mathcal{X}$, and $\psi : S \to \mathbb{R}$ be a bounded harmonic fn, ie.

$$\sum_y P(x,y)\psi(y) = \psi(x) \quad \forall x \in \mathcal{X}$$

Let $\tau = \min\{n \mid X_n = i\}$ (ie, first-passage time to state $i$)

$\Rightarrow \mathbb{E}[\tau] < \infty$ and $\mathbb{E}[|\psi(X_{n+1}) - \psi(X_n)| \mid \mathcal{F}_n] < \infty$

Thus $\mathbb{E}[\psi(X_\tau)] = \mathbb{E}[\psi(X_0)] = \psi(x) \quad \forall x \in \mathcal{X}$

(Since we can choose $X_0 = x$ for any $x \in \mathcal{X}$)

$\Rightarrow \psi(x)$ must be a constant function!

· Eg (Wald's Lemma) · Let $X_1, X_2, \ldots$ be iid r.v. with finite mean $\mu$, and let $S_n = \sum_{i=1}^n X_i$. Then $Y_n = S_n - n\mu$ is a martingale wrt the natural $(\mathcal{F}_n)$, and

$$\mathbb{E}[|Y_{n+1} - Y_n| \mid \mathcal{F}_n] = \mathbb{E}[|X_n - \mu|] < \infty$$

Thus, for any stopping time $\tau$ with $\mathbb{E}[\tau] < \infty$, $\mathbb{E}[S_\tau] = \mu \mathbb{E}[\tau]$

- This can be generalized as follows. Suppose $X_i$ have moment generating fn $M(t) = \mathbb{E}[e^{tX_i}]$, and $\exists\, t_0 > 0$ s.t $1 \leq M(t_0) < \infty$. Now consider $Y_n = e^{t_0 S_n} / (M(t_0))^n$

Then $\mathbb{E}[|Y_{n+1} - Y_n| \mid \mathcal{F}_n] = Y_n \mathbb{E}\left[\left|\frac{e^{t_0 X_{n+1}}}{M(t_0)} - 1\right|\right] \leq \frac{Y_n}{M(t_0)} \mathbb{E}[e^{t_0 X_{n+1}} + M(t_0)] = 2Y_n$

Now if $|S_n| < C \;\forall\, n \leq \tau$ for some $\tau$, then $Y_n \leq e^{t_0 C} / M(t_0)^n < \infty$

$\Rightarrow$ by optional stopping $\mathbb{E}[e^{tS_T} M(t)^{-T}] = 1$ when $1 \leq M(t) < \infty$

- To see an application of this, Suppose $X_i$ is iid, $\mathrm{Var}(X_i) > 0$ and $\tau = \min\{n \mid S_n \leq -a \text{ or } S_n \geq b\}$ (Generalizing the SRW)
- $|S_n| \leq \max(a,b)$ if $n < \tau$ and $\mathbb{E}[\tau] < \infty$

  (to see the latter, note $\exists M$ s.t $\mathbb{P}[|S_M| > a+b]$ as $X_i$ has positive variance. Now we consider $S_M, S_{2M}, \ldots$)
- Now choose $\theta$ s.t $M(\theta) = 1$. Let $\eta_a = \mathbb{E}[e^{\theta S_\tau} \mid S_\tau \leq -a]$ and $\eta_b = \mathbb{E}[e^{\theta S_\tau} \mid S_\tau \geq b]$. Then using Wald's inequality

$$\mathbb{P}[S_\tau \leq -a] = \frac{\eta_b - 1}{\eta_b - \eta_a}, \quad \mathbb{P}[S_\tau \geq b] = \frac{1 - \eta_a}{\eta_b - \eta_a}$$

▶ signifies something important, ▶▶ something very important, and ▶▶▶ the Martingale Convergence Theorem. (Williams - Probability with Martingales)

Thm (Martingale Convergence Theorem) Let $Y_n$ be a submartingale wrt filtration $\mathcal{F}_n$, and suppose $E[|Y_n|] \le M$ $\forall n$. Then $\exists$ r.v. $Y_\infty$ with $E[|Y_\infty|] < \infty$ s.t.

$$Y_n \xrightarrow{a.s} Y_\infty \quad \text{as } n \to \infty$$

Corollary - Any non-positive supermartingale / non-negative submartingale converges a.s. to an integrable r.v.

---

The proof uses a key inequality about 'upcrossings'.

Defn (Upcrossing) Suppose $(y_n)_{n \ge 0}$ is a real sequence and $[a,b] \subset \mathbb{R}$. Let $T_1 = \min\{n \mid y_n \le a\}$, $T_2 = \min\{n > T_1 \mid y_n \ge b\}$, and subsequently $T_{2k-1} = \min\{n \ge T_{2k-2} \mid y_n \le a\}$, $T_{2k} = \min\{n \ge T_{2k-1} \mid y_n \ge b\}$ for $k \ge 2$. Then $[T_{2k-1}, T_{2k}]$ are called the upcrossings of $[a,b]$.

Lemma - Let $U_n(a,b;y) \equiv$ # of upcrossings of $[a,b]$ by $\{y_0, y_1, \ldots, y_n\}$, and define $U(a,b;y) = \lim_{n\to\infty} U_n(a,b;y)$. If $U(a,b;y) < \infty$ for all rational $a,b$ s.t. $a<b$. Then $\lim_{n\to\infty} y_n$ exists (may be infinite).

Pf - Let $l = \liminf_{n\to\infty} y_n$ and $u = \limsup_{n\to\infty} y_n$ s.t. $l \le u$. Then $\exists\, a,b \in \mathbb{Q}$ s.t. $l < a < b < u$. Since $y_n \le a$ and $y_n \ge b$ for infinitely many $n \Rightarrow U(a,b;y) = \infty$ which is a contradiction. Thus $l = u$ ▨

- **Note** - The above limit was true for any sequence $y_n$! Now for any random process $Y_n$, if $U[a,b;Y] < \infty$ a.s., then it means each sample path converges wp 1!

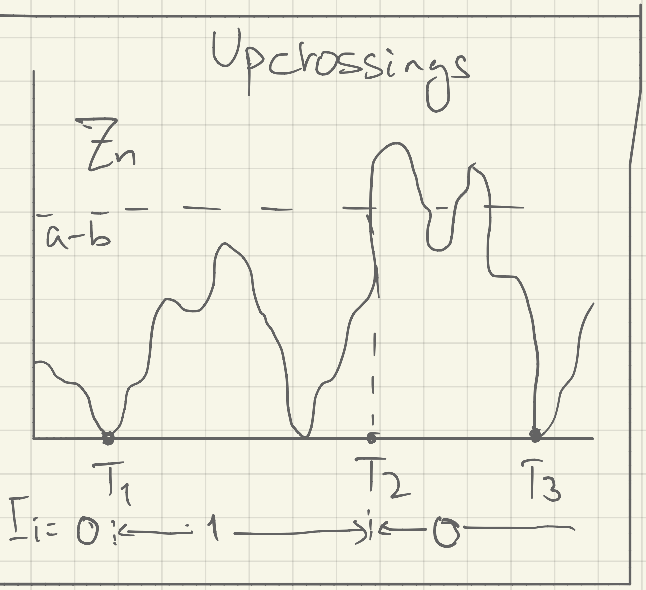
**Lemma** (Snell's Upcrossing Inequality) Let $Y_n$ be an $\mathcal{F}_n$-submartingale, and $a<b$. Then
$$\mathbb{E}\left[U_n(a,b;Y)\right] \le \frac{\mathbb{E}\left[(Y_n-a)^+\right]}{b-a}$$

Pf - Since $Y_n$ is a submartingale $\Rightarrow Z_n = (Y_n-a)^+$ is a submartingale (check). Upcrossings of $[a,b]$ by $Y_n \Leftrightarrow$ upcrossings of $[0,b-a]$ by $Z_n \Rightarrow U_n(a,b;Y) = U_n(0,b-a;Z)$.

- Let $[T_{2k-1}, T_{2k}]$, $k \ge 0$ be the upcrossings of $[0,b-a]$ by $Z_n$, and define $I_i = \mathbb{1}\{i \in (T_{2k-1}, T_{2k}] \text{ for some } k \ge 0\}$, $i \in [n]$. Note that $I_i$ is adapted to $\mathcal{F}_{i-1}$. Now we have

$$(b-a)\,\mathbb{E}\left[U_n(0,b-a;Z)\right] \le \mathbb{E}\left[\sum_{i=1}^{n}(Z_i - Z_{i-1})I_i\right]$$

$\leftarrow$ $\because$ each upcrossing contributes $\ge b-a$




$$= \sum_{i=1}^{n}\mathbb{E}\left[\mathbb{E}\left[(Z_i - Z_{i-1})I_i \mid \mathcal{F}_{i-1}\right]\right]$$
$$\le \sum_{i=1}^{n}\mathbb{E}\left[I_i\left(\mathbb{E}\left[Z_i \mid \mathcal{F}_{i-1}\right] - Z_{i-1}\right)\right]$$
$$\le \sum_{i=1}^{n}\left(\underbrace{\mathbb{E}\left[\mathbb{E}\left[Z_i \mid \mathcal{F}_{i-1}\right]\right]}_{=\mathbb{E}[Z_i]} - \mathbb{E}\left[Z_{i-1}\right]\right)$$
$$= \mathbb{E}\left[Z_n\right] - \mathbb{E}\left[Z_0\right]$$

$$\Rightarrow (b-a)\,\mathbb{E}\left[U_n(0,b-a;Z)\right] \le \mathbb{E}\left[Z_n - Z_0\right] \le \mathbb{E}\left[Z_n\right]$$ ∎

Pf of Thm - By the lemma, $\mathbb{E}[U_n(a,b;Y)] \leq \frac{\mathbb{E}[Y_n^+] + |a|}{b-a}$, and taking $n \to \infty$, $\mathbb{E}[U(a,b;Y)] \leq (M+|a|)/_{b-a}$ $\forall a<b$.

- Thus, $U(a,b;Y) < \infty$ a.s. for any given $a,b$, $a<b$.

Moreover, since $\mathbb{Q}$ is countable $\Rightarrow$ $\mathbb{E}[U(a,b;Y)] < \infty$ a.s. for all rationals $a,b$ s.t $a<b$.

$\Rightarrow$ $Y_n$ converges a.s. to some limit $Y_\infty$

Moreover by Fatou's lemma

$$\mathbb{E}\left[\lim_{n \uparrow \infty} Y_n\right] = \mathbb{E}\left[\liminf_{n\uparrow\infty} |Y_n|\right] \leq \liminf_{n\uparrow\infty} \mathbb{E}[|Y_n|] \leq M$$

$\boxtimes$