Warmup: Toss fair coin

Exp \# steps to see \( H \)?

\( HH \)?

\( HHH \)

\( HHH \)

\( HT \)?

\[
X = 2 + 1 + \frac{1}{2} X + \frac{1}{2} \left[ \frac{1}{2} HHH \right] \\
= 3 \frac{1}{2} + \frac{3}{8} X \\
\frac{1}{4} X = \frac{7}{2} \quad \Rightarrow \quad X = 14
\]

Organized approach (integers of fair gambling):

each step a new gambler arrives

bets $1 on first \( H \)

\[
\begin{align*}
\text{Lose} & \quad \text{Wms gets $2} \\
\text{bets it on T} & \quad \text{bets it on T} \\
\text{Lose} & \quad \text{Wms gets $4} \\
\text{Lose} & \quad \text{bets it on H} \\
\text{Lose} & \quad \text{Wms gets $8}
\end{align*}
\]
Every gambler bets fair bets. At end, say after n coin tosses get HTTH

\[ \text{Net profit of all gamblers} = 8 + 2 - n \]

\[ E(\text{net profit}) = E(8 + 2 - N) = 0 \]

\[ \uparrow \]

all bets are fair

\[ \Rightarrow E(N) = 10 \]

HTH 10
HHH 14
A stochastic process \( \{X_t\} \) is a **martingale** if \( \forall i \):

\[
E(X_{i+1} | X_0, \ldots, X_i) = X_i
\]

\( E(\gamma|Z) \) is a r.v. that takes value \( E(\gamma|Z=z) \) with prob \( \Pr(Z=z) \).

**Example:** gambler walks into casino with \( X_0 \)
- \( X_i \): amt of money gambler has after \( i \) games
- every game is "fair" i.e. expected winnings = 0

\( \{X_t\} \) is **martingale** with respect to \( \{Y_t\} \) where \( X_t = f(Y_0, \ldots, Y_t) \)

\[
E(X_{i+1} | Y_0, \ldots, Y_i) = X_i
\]
Some common martingales:

1. Sums of independent random variables

\[ Y_0 = 0, \ Y_1, ..., Y_n \text{ iid with } E(Y_k) = 0 \quad \forall k \]

Define \( X_0 = 0 \), \( X_n = Y_1 + \ldots + Y_n \)

\( \{ X_t \} \) is a martingale w.r.t. \( \{ Y_t \} \)

\[ E(X_{n+1} | Y_0, ..., Y_n) = E(X_n + Y_{n+1} | Y_0, ..., Y_n) \]

\[ = E(Y_{n+1} | Y_0, ..., Y_n) + E(X_n | Y_0, ..., Y_n) \]

\[ = X_n + E(Y_{n+1}) \]

\[ = Y_n \]

2. Variance of a sum

\[ Y_0 = 0, \ Y_1, ..., Y_n \text{ iid with } E(Y_k) = 0 \quad \forall k \quad E(Y_k^2) = \sigma^2 \]

Define \( X_0 = 0 \), \( X_n = (\frac{1}{n} \sum_{k=1}^{n} Y_k)^2 - n \sigma^2 \)

\( \{ X_t \} \) is a martingale w.r.t. \( \{ Y_t \} \)
\[ E(Y_{n+1} \mid Y_0, \ldots, Y_n) = E[(Y_{n+1} + \sum_{k=1}^{n} Y_k) - (n+1)\sigma^2 \mid Y_0, \ldots, Y_n] \]
\[ = E(Y_{n+1}^2 + 2Y_{n+1} \sum_{k=1}^{n} Y_k + (\sum_{k=1}^{n} Y_k)^2 - (n+1)\sigma^2 \mid Y_0, \ldots, Y_n] \]
\[ = X_n + E(\sum_{k=1}^{n} Y_k) + \sigma^2 E(Y_{n+1}) - \sigma^2 \]
\[ = X_n \]

(3) "Dubbs" martingale process

\( Y_1, Y_2, \ldots \) arbitrary seq of random vars

\( X \) r.v. w/ finite expectation

\( X_n = E(X \mid Y_1, \ldots, Y_n) \) forms martingale wrt. \( \{Y_n\} \)

\( X_0 = E(X) \)

\[
\begin{align*}
X_0 &= E(X) \\
X_1 &= E(X \mid Y_1) \\
X_2 &= E(X \mid Y_1, Y_2) \\
&\vdots \\
X_n &= E(X \mid Y_1, \ldots, Y_n) \\
X_{n+1} &= E(X \mid Y_1, \ldots, Y_{n+1}) \\
\end{align*}
\]

\( Y_{y_1 \ldots y_n} \quad E(X_{n+1} \mid Y_1 = y_1, \ldots, Y_n = y_n) = X_n \)
\[ E(X_n \mid Y_1, \ldots, Y_n) \]
\[ = E(E(X \mid Y_1, \ldots, Y_n) \mid Y_1, \ldots, Y_n) \]
\[ = E(X \mid Y_1, \ldots, Y_n) = X_n \]

Ex: Edge exposure martingale

\( G(n, p) \) random graph

label \( m = \binom{n}{2} \) potential edges \( e_1, \ldots, e_m \)

Let \( f \) be graph-theoretic fn e.g. chromatic num

\[ \text{max independent set size} \]

\[ Y_j \text{ ind. r.v. } = \begin{cases} 1 & \text{if edge } e_j \text{ present} \\ 0 & \text{otherwise} \end{cases} \]

\[ X_k = E(f(G) \mid Y_1, \ldots, Y_k) \]

\[ X_0 = E(f(G)) \quad X_m = f(G) \]
Some useful facts about martingales

1. \( E(X_n) = E(X_0) \)

by induction

\[
E(X_{n+1} | Y_0, \ldots, Y_n) = X_n
\]

\[
E\left( E(X_{n+1} | Y_0, \ldots, Y_n) \right) = E(X_n)
\]

\[
E\left( \frac{E(X_1 | Y)}{E(X_1 | Y, y)} \right) = E(X)
\]

with prob \( \Pr(Y = y) \)

\[
E(E(X | Y)) = \sum_y E(X | Y = y) \Pr(Y = y)
\]

\[
= \sum_y \sum_x x \Pr(X = x | Y = y) \Pr(Y = y)
\]

\[
= \sum_x \sum_y \Pr(X = x, Y = y)
\]

\[
= \sum_x \Pr(X = x)
\]
\[ \{Z_t\} \text{ martingale wrt. } \{X_t\} \]

2. For \( T \) a stopping time \( "\text{know it when you see it}" \)

\[ E(Z_T) = E(Z_0) \]

optional sampling thm

whenever any of following hold

- \( Z_t \) is bounded \((\exists C \text{ s.t. } \forall t \quad |Z_t| \leq C)\)
- \( T \) is bounded
- \( E(T) < \infty \) and \( \exists C \text{ s.t. } E(\|Z_{T+n} - Z_T\| \mid X_{-n}, X_T) \leq C \)

A r.v. \( T \) is called a "stopping time" wrt. \( \{X_t\} \) if

\( T \) takes values in \( \{0, 1, 2, \ldots \} \)

and \( \forall n \geq 0 \), the event \( \{T = n\} \) is determined by \( X_{0, \ldots, n} \)

i.e. can determine \( \{T = n\} \) or \( T \geq n \) from knowledge of values \( X_{0, \ldots, n} \)
1. unbiased r.w. on line starting at 0

\[ \frac{1}{a} \cdot \frac{1}{b} \]

\[ \Pr(\text{reach } -a \text{ before reaching } b) \]

\[ Y_i = \begin{cases} 1 & \text{w.prob } \frac{1}{a} \\ -1 & \text{w.prob } \frac{1}{b} \end{cases} \]

\[ X_n = \sum_{i=1}^{n} Y_i \text{ martingale} \]

\[ T = \min\{n \mid X_n = -a \text{ or } X_n = b\} \]

\[ T \text{ is a stopping time} \]

Let \( v_a = \Pr(X_n \text{ reaches } -a \text{ before reaching } b) \)

By o.s.t. \( E(X_T) = E(X_0) = 0 \)

\[ E(X_T) = v_a(-a) + (1-v_a)b = 0 \]

\[ \Rightarrow \quad v_a = \frac{b}{a+b} \]
2) Some unbiased r.w. on line, same $T$

What is $E(T)$?

$Z_n = X_n^2 - n$ is a martingale

$\text{var of a sum } E(Y_i^2) = 1$

By O.S.T.

$E(Z_T) = E(Z_0) = 0$

$E(Z_T) = [v_0 \cdot \frac{a^2}{a+b} + (1-v_0)b^2] - E(T) = 0 \text{ s.t.}$

$\Rightarrow E(T) = \frac{b}{a+b} \cdot \frac{a^2}{a+b} + \frac{a}{b+a} \cdot b^2 = ab$
Same questions: biased r.w.

\[ Y_i = \begin{cases} +1 & p \\ -1 & q \end{cases} \quad p > q \quad (=1-p) \]

\[ X_n = \sum_{i=1}^{\infty} Y_i - n(p-q) \]

\[ X_n' = \left( \frac{q}{p} \right)^{\sum_{i=1}^{\infty} Y_i} \quad (X_0'=1) \]

\[ \{X_n'\} \text{ martingale w.r.t. } \{Y_i\} \]

\[ T = \min\{n | \sum_{i=1}^{n} Y_i = -a \text{ or } = b \} \]

\[ v_a = \Pr(\sum_{i=1}^{\infty} Y_i \text{ reaches } -a \text{ before } b) \]

\[ E(X_T') = E(X_0') = 1 \]

\[ E(X_T') = v_a \left( \frac{q}{p} \right)^a + (1-v_a) \left( \frac{q}{p} \right)^b \]

\[ \Rightarrow \quad v_a = \frac{1 - \left( \frac{q}{p} \right)^b}{\left( \frac{q}{p} \right)^a - \left( \frac{q}{p} \right)^b} \]
Tail Inequalities (Large deviations)

\[ E(X_m) = E(X_0) \] how far can it be from its expectation?

\textbf{Azuma-Hoeffding Inequality}

\[ X_0, \ldots, X_m \text{ martingale s.t. } \forall k \quad |X_k - X_{k-1}| \leq c_k \]

\( (c_k \text{ may depend on } k) \)

Then \( \forall \; t > 0, \text{ any } R > 0 \)

\[ \Pr(|X_t - X_0| > R) \leq 2e^{-\left[\frac{R^2}{2\sum_{k=1}^m c_k^2}\right]} \]

\textbf{Proof:}

By convexity of \( f(x) = e^{\gamma x} \) we have \( e^{\gamma x} \leq \frac{(1-\frac{x}{2})e^{\gamma c} + (1+\frac{x}{2})e^{-\gamma c}}{2} \)

for \( x \in [-c, c] \)

Thus \( Y \) has \( E(Y) = 0 \) and \( |Y| \leq c \), then

\[ E(e^{\gamma Y}) \leq E(e^{\gamma X}) = \frac{e^{\gamma X} + e^{-\gamma X}}{2} = \frac{e^{\gamma c} + e^{-\gamma c}}{2} = \frac{e^{\gamma c} + e^{-\gamma c}}{2} \leq \frac{e^{\gamma c} + e^{-\gamma c}}{2} \]

Therefore \( E(e^{\gamma X_t} | H_t) \leq e^{\gamma^2 c^2/2} \)
\[ E(e^{X_{t+1}} | H_t) = e^{\lambda X_t} E(e^{\lambda X_{t+1}} | H_t) \leq e^{\lambda X_t} e^{\frac{\lambda^2 \sigma^2}{2}} \]

Taking expectations,
\[ E(e^{\lambda X_t}) \leq e^{\lambda^2 \sigma^2 / 2} \]

by induction on \( t \)

Finally, \( P(X_t \geq R) = P(e^{\lambda X_t} \geq e^{\lambda R}) \leq e^{-\lambda R} e^{\frac{\lambda^2 \sigma^2}{2}} \)

Optimizing, we choose \( \lambda = \frac{R}{\frac{\lambda \sigma}{\sqrt{2}} \cdot \frac{1}{2}} \quad P(X_t \geq R) \leq e^{-\frac{R^2}{2 \frac{\lambda^2 \sigma^2}{2}}} \)

\[-\left( \lambda R - \frac{\lambda^2 \sigma^2}{2} \right) = -\left( \frac{R^2}{2 \frac{\lambda^2 \sigma^2}{2}} - \frac{R^2}{2 \frac{\lambda^2 \sigma^2}{2}} \right) \]

Factor of 2 comes from \( P(X_t < -2) \) obtained by replacing \( X_t - X_{t-1} \) by \( X_{t-1} - X_t \)
Tail bounds r.w. on line

\[
y_0 = 0 \quad y_i = \begin{cases} 1 & \frac{i}{n} \\ -1 & \frac{i}{n} \\ \frac{1}{2} & \end{cases}
\]

\[
x_n = \text{position of particle at time } n = \sum_{i=1}^{n} y_i \quad x_0 = 0
\]

\[
|X_k - X_{k-1}| \leq 1 \quad \implies \sum_{k=1}^{t} c_k = t
\]

\[
\Pr(|X_t - X_0| > 2) \leq 2 e^{-\frac{\lambda^2}{4t}} \quad \text{by Azuma-Hoeffding}
\]

\[
t = n \quad \lambda = \sqrt{n + \varepsilon}
\]

\[
\Pr(\text{particle \geq \sqrt{n} steps from origin after } n \text{ steps}) \leq 2 e^{-\frac{n^{1+\varepsilon}}{2n}} = O(e^{-n^{\varepsilon}})
\]

Chromatic \# in random graph \( G(n, \frac{1}{2}) \)

Azuma \implies \text{sharp concentration of chromatic } \# \text{ around its expectation}
Finding "interesting" patterns (e.g. in DNA sequences)

Let \( X = (X_1, \ldots, X_n) \) be a sequence of characters chosen independently and u.a.r. from \( \Sigma \) (e.g. \( \Sigma = \{A, T, C, G\} \))

\( B = (b_1, \ldots, b_k) \) fixed string of characters.

\( F : \#	ext{ occurrences of } F \)

\( E(F) = ? \)

\( E(F) = (n-k+1) \left( \frac{1}{|S|} \right)^k \)

Let \( Z_i = E(F | X_1, \ldots, X_i) \)  

Dobbs martingale

\( Z_n = F, \quad Z_0 = E(F) \)
\[ |Z_{i+1} - Z_i| \leq k \quad \text{since each character can participate in } \leq k \text{ matches} \]

\[ \Rightarrow \text{By Azuma-Hoeffding} \]

\[ \Pr (|F - E(F)| \geq 2) \leq 2e^{-\frac{\lambda^2}{2nk^2}} \]

or for \( n = c k \sqrt{n} \)

\[ \Pr (|F - E(F)| > c k \sqrt{n}) \leq 2e^{-\frac{\lambda^2}{8nk^2}} \]