Martingales

Definition

A sequence of random variables $Z_0, Z_1, \ldots$ is a *martingale* with respect to the sequence $X_0, X_1, \ldots$ if for all $n \geq 0$ the following hold:

1. Z_n is a function of $X_0, X_1, \ldots, X_n$;
2. $\mathbf{E}[|Z_n|] < \infty$;
3. $\mathbf{E}[Z_{n+1} | X_0, X_1, \ldots, X_n] = Z_n$;

Definition

A sequence of random variables $Z_0, Z_1, \ldots$ is a *martingale* when it is a martingale with respect to itself, that is

1. $\mathbf{E}[|Z_n|] < \infty$;
2. $\mathbf{E}[Z_{n+1} | Z_0, Z_1, \ldots, Z_n] = Z_n$;

Example

I play series of fair games (win with probability $1/2$).

Game 1: bet $\$1$.

Game $i > 1$: bet 2^i if won in round $i - 1$; bet i otherwise.

$X_i =$ amount won in ith game. ($X_i < 0$ if ith game lost).

$Z_i =$ total winnings at end of ith game.

Example

$X_i =$ amount won in ith game. ($X_i < 0$ if ith game lost).

$Z_i =$ total winnings at end of ith game.

$Z_1, Z_2, \ldots$ is martingale with respect to $X_1, X_2, \ldots$

$\mathbf{E}[X_i] = 0.$

$\mathbf{E}[Z_i] = \sum \mathbf{E}[X_i] = 0 < \infty.$

$\mathbf{E}[Z_{i+1} | X_1, X_2, \ldots, X_i] = Z_i + \mathbf{E}[X_{i+1}] = Z_i.$

Doob Martingale

Let $X_0, X_1, \ldots, X_n$ be sequence of random variables. Let Y be a random variable with $\mathbf{E}[|Y|] < \infty$. In general Y is a function of $X_1, X_2, \ldots, X_n$.

Let $Z_i = \mathbf{E}[Y | X_0, X_1, \ldots, X_i]$, $i = 0, 1, \ldots, n$.

$Z_0, Z_1, \ldots, Z_n$ is martingale with respect to $X_0, X_1, \ldots, X_n$.

(Often $Z_0 = \mathbf{E}[Y]$.)

Proof

Fact

$\mathbf{E}[\mathbf{E}[V|U, W]|W] = \mathbf{E}[V|W]$.

$Z_i = \mathbf{E}[Y|X_0, X_1, \ldots, X_i]$, $i = 0, 1, \ldots, n$

$$
\begin{aligned}
\mathbf{E}[Z_{i+1}|X_0, X_1, \ldots, X_i] &= \mathbf{E}[\mathbf{E}[Y|X_0, X_1, \ldots, X_{i+1}]|X_0, X_1, \ldots, X_i] \\
&= \mathbf{E}[Y|X_0, X_1, \ldots, X_i] \\
&= Z_i.
\end{aligned}
$$

Example: Edge Exposure Martingale

Let G random graph from $G_{n,p}$. Consider $m = \binom{n}{2}$ possible edges in arbitrary order.

$$X_i = \begin{cases} 1 & \text{if } i\text{th edge is present} \\ 0 & \text{otherwise} \end{cases}$$

$F(G) =$ size maximum clique in G.

$Z_0 = \mathbf{E}[F(G)]$

$Z_i = \mathbf{E}[F(G) | X_1, X_2, \ldots, X_i]$, for $i = 1, \ldots, m$.

$Z_0, Z_1, \ldots, Z_m$ is a Doob martingale.

($F(G)$ could be any finite-valued function on graphs.)

Back to Gambling

I play series of fair games (win with probability $1/2$).

Game 1: bet $\$1$.

Game $i > 1$: bet 2^i if I won in round $i-1$; bet i otherwise.

$X_i =$ amount won in ith game. ($X_i < 0$ if ith game lost).

$Z_i =$ total winnings at end of ith game.

Assume that (before starting to play) I decide to quit after k games: what are my expected winnings?

Lemma

If $Z_0, Z_1, \ldots, Z_n$ *is a martingale with respect to* $X_0, X_1, \ldots, X_n$*, then*

$$\mathbf{E}[Z_n] = \mathbf{E}[Z_0].$$

Proof.

Since Z_i defines a martingale

$$Z_i = \mathbf{E}[Z_{i+1} | X_0, X_1, \ldots, X_i].$$

Then

$$\mathbf{E}[Z_i] = \mathbf{E}[\mathbf{E}[Z_{i+1} | X_0, X_1, \ldots, X_i]] = \mathbf{E}[Z_{i+1}].$$

□

Back to Gambling

I play series of fair games (win with probability $1/2$).

Game 1: bet $\$1$.

Game $i > 1$: bet 2^i if I won in round $i - 1$; bet i otherwise.

$X_i =$ amount won in ith game. ($X_i < 0$ if ith game lost).

$Z_i =$ total winnings at end of ith game.

Assume that (before starting to gamble) we decide to quit after k games: what are my expected winnings?

$\mathbf{E}[Z_k] = \mathbf{E}[Z_1] = 0$.

A Different Strategy

Same gambling game. What happens if I:

- play a random number of games?
- decide to stop only when I have won (or lost) \$1000?

Stopping Time

Definition

A non-negative, integer *random variable* T is a *stopping time* for the sequence $Z_0, Z_1, \ldots$ if the event "$T = n$" depends only on the value of random variables $Z_0, Z_1, \ldots, Z_n$.

Intuition: corresponds to a strategy for determining when to stop a sequence based only on values seen so far.

In the gambling game:

- first time I win 10 games in a row: is a stopping time;
- the last time when I win: is not a stopping time.

Consider again the gambling game: let T be a stopping time.

$Z_i =$ total winnings at end of ith game.

What are my winnings at the stopping time, i.e. $\mathbf{E}[Z_T]$?

Consider again the gambling game: let T be a stopping time.

$Z_i =$ total winnings at end of ith game.

What are my winnings at the stopping time, i.e. $\mathbf{E}[Z_T]$?

Fair game: $\mathbf{E}[Z_T] = \mathbf{E}[Z_0] = 0$?

Consider again the gambling game: let T be a stopping time.

$Z_i =$ total winnings at end of ith game.

What are my winnings at the stopping time, i.e. $\mathbf{E}[Z_T]$?

Fair game: $\mathbf{E}[Z_T] = \mathbf{E}[Z_0] = 0$?

"T=first time my total winnings are at least \$1000" is a stopping time, and $\mathbf{E}[Z_T] \geq 1000$...

Consider again the gambling game: let T be a stopping time.

$Z_i =$ total winnings at end of ith game.

What are my winnings at the stopping time, i.e. $\mathbf{E}[Z_T]$?

Fair game: $\mathbf{E}[Z_T] = \mathbf{E}[Z_0] = 0$?

"T=first time my total winnings are at least \$1000" is a stopping time, and $\mathbf{E}[Z_T] > 1000$...

This is a particular stopping time: it may not be finite!

Martingale Stopping Theorem

Theorem

If $Z_0, Z_1, \ldots$ is a martingale with respect to $X_1, X_2, \ldots$ and if T is a stopping time for $X_1, X_2, \ldots$ then

$$\mathbf{E}[Z_T] = \mathbf{E}[Z_0]$$

whenever one of the following holds:

- *there is a constant c such that, for all i, $|Z_i| \leq c$;*
- *T is bounded;*
- *$\mathbf{E}[T] < \infty$, and there is a constant c such that $\mathbf{E}[|Z_{i+1} - Z_i||X_1, \ldots, X_i] < c$.*

Example: The Gambler's Ruin

- Consider a sequence of independent, two players, fair gambling games.
- In each round a player wins a dollar with probability $1/2$ or loses a dollar with probability $1/2$.
- $X_i =$ amount won by player 1 on ith round.
- If player 1 has lost in round i: $X_i < 0$.
- $Z_i =$ total amount won by player 1 after ith rounds.
- $Z_0 = 0$.
- Player 1 must end the game if she loses ℓ_1 dollars ($Z_t = -\ell_1$); player 2 must terminate when she loses ℓ_2 dollars ($Z_t = \ell_2$).
- $q =$ probability that the game ends with player 1 wining ℓ_2 dollars.

Example: The Gambler's Ruin

- $T =$ first time player 1 wins ℓ_2 dollars or loses ℓ_1 dollars.
- T is a stopping time for $X_1, X_2, \ldots$.
- $Z_0, Z_1, \ldots$ is a martingale.
- Z_i's are bounded.
- Martingale Stopping Theorem: $\mathbf{E}[Z_T] = \mathbf{E}[Z_0] = 0$.

$$\mathbf{E}[Z_T] = q\ell_2 - (1-q)\ell_1 = 0$$

$$q = \frac{\ell_1}{\ell_1 + \ell_2}$$

Example: A Ballot Theorem

- Candidate A and candidate B run for an election.
- Candidate A gets a votes.
- Candidate B gets b votes.
- $a > b$.
- Votes are counted in random order: chosen from all permutations on $n = a + b$ votes.
- What is the probability that A is always ahead in the count?

Example: A Ballot Theorem

- $S_k =$ number of votes A is leading by after k votes counted (if A is trailing: $S_k < 0$).

- $S_n = a - b$.

- For $0 \leq k \leq n-1$: $X_k = \frac{S_{n-k}}{n-k}$.

- Consider $X_0, X_1, \ldots, X_n$. It relates to the counting process in backwards order.

$$\mathbf{E}[X_k | X_0, X_1, \ldots, X_{k-1}] = ?$$

Example: A Ballot Theorem

$$\mathbf{E}[X_k | X_0, X_1, \dots, X_{k-1}] = ?$$

- Conditioning on $X_0, X_1, \dots, X_{k-1}$: equivalent to conditioning on $S_n, S_{n-1}, \dots, S_{n-k+1}$, equivalent on conditioning on values of count when counting $k-1$ last votes.
- $a_k =$ number of votes for A after first k votes are counted.
- $b_k =$ number of votes for B after first k votes are counted.

Conditioning on S_{n-k+1}:

$$a_{n-k+1} = \frac{a_{n-k+1} + b_{n-k+1} + a_{n-k+1} - b_{n-k+1}}{2} = \frac{n-k+1+S_{n-k+1}}{2}$$

$$b_{n-k+1} = \frac{a_{n-k+1} + b_{n-k+1} - (a_{n-k+1} - b_{n-k+1})}{2} = \frac{n-k+1-S_{n-k+1}}{2}$$

Example: A Ballot Theorem

- $n-k+1$th vote: random vote among these first $n-k+1$ votes.

$$S_{n-k} = \begin{cases} S_{n-k+1}+1 & \text{if } n-k+1\text{th vote is for B} \\ S_{n-k+1}-1 & \text{if } n-k+1\text{th vote is for A} \end{cases}$$

$$\begin{aligned} \mathbf{E}[S_{n-k}|S_{n-k+1}] &= (S_{n-k+1}+1)\frac{n-k+1-S_{n-k+1}}{2(n-k+1)} \\ &+ (S_{n-k+1}-1)\frac{n-k+1+S_{n-k+1}}{2(n-k+1)} \\ &= S_{n-k+1}\frac{n-k}{n-k+1} \end{aligned}$$

Example: A Ballot Theorem

$$\mathbf{E}[S_{n-k}|S_{n-k+1}] = S_{n-k+1}\frac{n-k}{n-k+1}$$

$$\begin{aligned}\mathbf{E}[X_k|X_0, X_1, \ldots, X_{k-1}] &= \mathbf{E}\left[\left.\frac{S_{n-k}}{n-k}\right| S_n, \ldots, S_{n-k+1}\right] \\ &= \frac{S_{n-k+1}}{n-k+1} \\ &= X_{k-1}\end{aligned}$$

$X_0, X_1, \ldots, X_n$ is a martingale.

Example: A Ballot Theorem

$$T = \begin{cases} \min\{k : X_k = 0\} & \text{if such } k \text{ exists} \\ n-1 & \text{otherwise} \end{cases}$$

- T is a stopping time.
- T is bounded.
- Martingale Stopping Theorem:

$$\mathbf{E}[X_T] = \mathbf{E}[X_0] = \frac{\mathbf{E}[S_n]}{n} = \frac{a-b}{a+b}.$$

Two cases:

1. A leads throughout the count.
2. A does not lead throughout the count.

Example: A Ballot Theorem

1. A leads throughout the count.

For $0 \leq k \leq n-1$: $S_{n-k} > 0$, then $X_k > 0$.

$T = n - 1$.

$X_T = X_{n-1} = S_1$.

A gets the first vote in the count: $S_1 = 1$, then $X_T = 1$.

Example: A Ballot Theorem

2. A does not lead throughout the count.

A leads at the end. If at a certain point B leads, at a certain moment k: $S_k = 0$. Then $X_k = 0$.

$T = k < n - 1$.

$X_T = 0$.

Example: A Ballot Theorem

Putting it all together:

1. A leads throughout the count: $X_T = 1$.
2. A does not lead throughout the count: $X_T = 0$

$$\mathbf{E}[X_T] = \frac{a-b}{a+b} = 1\dot{\Pr}(\text{Case } 1) + 0\dot{\Pr}(\text{Case } 2).$$

That is

$$\Pr(\text{A leads throughout the count}) = \frac{a-b}{a+b}$$

A Different Gambling Game

Two stages:

1. roll one die; let X be the outcome;
2. roll X standard dice; your gain Z is the sum of the outcomes of the X dice.

What is your expected gain?

Wald's Equation

Theorem

Let $X_1, X_2, \ldots$ be nonnegative, independent, identically distributed random variables with distribution X. Let T be a stopping time for this sequence. If T and X have bounded expectation, then

$$\mathbf{E}\left[\sum_{i}^{T} X_i\right] = \mathbf{E}[T]\mathbf{E}[X].$$

Corollary of the martingale stopping theorem.

Stopping Time: Sequence of Independent r.v.

Definition

Let $Z_0, Z_1, \dots$ be a sequence of independent random variables. A nonnegative, integer-valued random variable T is a stopping time for the sequence if the event "$T = n$" is independent of $Z_{n+1}, Z_{n+2}, \dots$.

A Different Gambling Game

Two stages:

1. roll one die; let X be the outcome;
2. roll X standard dice; your gain Z is the sum of the outcomes of the X dice.

What is your expected gain?

$Y_i =$ outcome of ith die in second stage.

$$\mathbf{E}[Z] = \mathbf{E}\left[\sum_{i=1}^{X} Y_i\right].$$

X is a stopping time for $Y_1, Y_2, \ldots$.

By Wald's equation:

$$\mathbf{E}[Z] = \mathbf{E}[X]\mathbf{E}[Y_i] = \left(\frac{7}{2}\right)^2.$$

Example

n servers: each has queue with packets to send.

Time divided in discrete slots; servers send packets to communicate.

Communicate through *shared* channel:

- if exactly 1 packet sent in time slot, transmission is successful;
- if > 1 packet sent in time slot, *none* is successful.

At each time slot:

- if queue is not empty, the first packet in the queue with probability $\frac{1}{n}$.

Assume: Queues are never empty.

Expected number of time slots until each server successfully sends at least one packet?

$T =$ number of time slots until each server successfully sends at least one packet.

$N =$ number of packets successfully sent until each server has successfully sent at least one packet.

$t_i =$ time slot ith successfully transmitted packet is sent. $t_0 = 0$.

$r_i = t_i - t_{i+1}$.

$$T = \sum_{i=1}^{N} r_i.$$

$N =$ number of packets successfully sent until each server has successfully sent at least one packet.

$t_i =$ time slot ith successfully transmitted packet is sent. $t_0 = 0$.

$r_i = t_i - t_{i+1}$.

Easy to check that:

- N is independent of $r_0, r_1, \ldots$;
- $\mathbf{E}[N] < \infty$.

Then N is a stopping time for $r_0, r_1, \ldots$.

$$\mathbf{E}[T] = \mathbf{E}\left[\sum_{i=1}^{N} r_i\right] = \mathbf{E}[N]\mathbf{E}[r_i].$$

$p =$ probability a packet successfully sent in a time slot

$$p = \binom{n}{1} \left(\frac{1}{n}\right) \left(1 - \frac{1}{n}\right)^{n-1} \approx e^{-1}.$$

r_i: geometric random variable $G(p)$.

$$\mathbf{E}[r_i] = 1/p \approx e.$$

$N =$ number of packets successfully sent until each server has successfully sent at least one packet.

Coupon collector: $\mathbf{E}[N] = nH(n) = n \ln n + O(n)$.

$$\mathbf{E}[T] = \mathbf{E}\left[\sum_{i=1}^{N} r_i\right] = \mathbf{E}[N]\mathbf{E}[r_i] = \frac{nH(n)}{p} \approx en \ln n.$$

Tail Inequalities

Theorem (Azuma-Hoeffding Inequality)

Let $Z_0, Z_1, \ldots, Z_n$ *be a martingale such that*

$$|Z_k - Z_{k-1}| \leq c_k.$$

Then, for all $t \geq 0$ *and any* $\lambda > 0$

$$\Pr(|Z_t - Z_0| \geq \lambda) \leq 2e^{-\lambda^2/(2\sum_{k=1}^{t} c_k^2)}.$$

Tail Inequalities: A More General Form

Theorem (Azuma-Hoeffding Inequality)

Let $Z_0, Z_1, \dots, Z_n$ *be a martingale such that*

$$B_k \leq Z_k - Z_{k-1}| \leq B_k + c_k$$

for some constants c_k *and for some random variables* B_k *that may be functions of* $X_0, X_1, \dots, X_{k-1}$*. Then, for all* $t \geq 0$ *and any* $\lambda > 0$

$$\Pr(|Z_t - Z_0| \geq \lambda) \leq 2e^{-2\lambda^2/(\sum_{k=1}^{t} c_k^2)}.$$

Tail Inequalities: Doob Martingales

Let $X_1, \dots, X_n$ be sequence of random variables.

Random variable Y:

- Y is a function of $X_1, X_2, \dots, X_n$;
- $\mathbf{E}[|Y|] < \infty$.

Let $Z_i = \mathbf{E}[Y | X_1, \dots, X_i]$, $i = 0, 1, \dots, n$.

$Z_0, Z_1, \dots, Z_n$ is martingale with respect to $X_1, \dots, X_n$.

If we can use Azuma-Hoeffding inequality:

$$\Pr(|Z_n - Z_0| \geq \lambda) \leq \varepsilon(\lambda, \dots)$$

that is

$$\Pr(|Y - \mathbf{E}[Y]| \geq \lambda) \leq \varepsilon(\lambda, \dots).$$

A General Formalization

$f(X_1, X_2, \ldots, X_n)$ satisfies *Lipschitz condition* with bound c if for any i and any set of values $x_1, \ldots, x_n$ and y:

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

$$Z_0 = \mathbf{E}[f(X_1, X_2, \ldots, X_n)].$$

$$Z_k = \mathbf{E}[f(X_1, X_2, \ldots, X_n) | X_1, \ldots, X_k].$$

$Z_0, Z_1, \ldots, Z_n$ is a Doob martingale.

If $X_1, X_2, \ldots, X_k$ are independent random variables: there exists B_k depending only on $Z_0, Z_1, \ldots, Z_{k-1}$ with

$$B_k \leq Z_k - Z_{k-1} \leq B_k + c.$$

A General Formalization

$$Z_0 = \mathbf{E}[f(X_1, X_2, \ldots, X_n)].$$

$$Z_k = \mathbf{E}[f(X_1, X_2, \ldots, X_n) | X_1, \ldots, X_k].$$

$Z_0, Z_1, \ldots, Z_n$ is a Doob martingale.

If $X_1, X_2, \ldots, X_k$ are independent random variables: there exists B_k depending only on $Z_0, Z_1, \ldots, Z_{k-1}$ with

$$B_k \leq Z_k - Z_{k-1} \leq B_k + c.$$

By Azuma-Hoeffding:

$$\begin{aligned} \Pr(|Z_n - Z_0| \geq \lambda) &= \Pr(|f(\ldots) - \mathbf{E}[f(\ldots)]| \geq \lambda) \\ &\leq 2e^{-2\lambda^2/(\sum_{k=1}^{n} c_k^2)}. \end{aligned}$$

Example: Pattern Matching

Given a string and a pattern: is the pattern interesting?

Does it appear more often than is expected in a random string?

Is the number of occurrences of the pattern concentrated around the expectation?

$S = (S_1, S_2, \ldots, S_n)$ string of characters, each chosen independently and uniformly at random from σ, with $s = |\sigma|$.

pattern: $B = (b_1, \ldots, b_k)$ fixed string, $b_i \in \sigma$.

$F=$ number occurrences of B in random string S.

$$\mathbf{E}[F] = ?$$

$S = (S_1, S_2, \ldots, S_n)$ string of characters, each chosen independently and uniformly at random from Σ, with $m = |\Sigma|$.

pattern: $B = (b_1, \ldots, b_k)$ fixed string, $b_i \in \Sigma$.

$F=$ number occurrences of B in random string S.

$$\mathbf{E}[F] = (n - k + 1)\left(\frac{1}{m}\right)^k .$$

Can we bound the deviation of F from its expectation?

F= number occurrences of B in random string S.

$Z_0 = \mathbf{E}[F]$

$Z_i = \mathbf{E}[F|S_1, \ldots, S_i]$, for $i = 1, \ldots, n$.

$Z_0, Z_1, \ldots, Z_n$ is a Doob martingale.

$Z_n = F$.

$F =$ number occurrences of B in random string S.

$Z_0 = \mathbf{E}[F]$

$Z_i = \mathbf{E}[F | S_1, \ldots, S_i]$, for $i = 1, \ldots, n$.

$Z_0, Z_1, \ldots, Z_n$ is a Doob martingale.

$Z_n = F$.

Each character in S can participate in no more than k occurrences of B:

$$|Z_i - Z_{i+1}| \leq k.$$

Azuma-Hoeffding inequality (version 1):

$$\Pr(|F - \mathbf{E}[F]| \geq \lambda) \leq 2e^{-\lambda^2/(2nk^2)}.$$

Slightly better bound:

$$F = f(S_1, S_2, \ldots, S_n).$$

Each character in S can participate in no more than k occurrences of B: for all i, for all $s_1, \ldots, s_n$ and y

$$|f(s_1, \ldots, s_{i-1}, s_i, s_{i+1}, \ldots, s_n) - f(s_1, \ldots, s_{i-1}, y, s_{i+1}, \ldots, s_n)| \leq k.$$

Azuma-Hoeffding inequality (general framework):

$$\Pr(|F - \mathbf{E}[F]| \geq \lambda) \leq 2e^{-2\lambda^2/(nk^2)}.$$

$$\Pr(|F - \mathbf{E}[F]| \geq ck\sqrt{n}) \leq 2e^{-2c^2}.$$