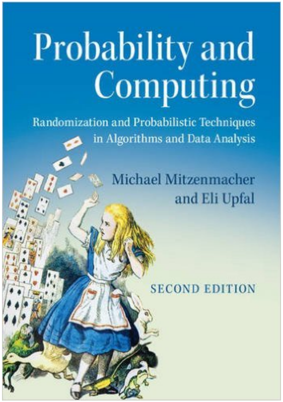
CS155/254: Probabilistic Methods in Computer Science

Chapter 13.1: Martingales



Hoeffding's Bound

Theorem

Let $X_1, \ldots, X_n$ *be* **independent** *random variables with* $\mathbf{E}[X_i] = \mu_i$ *and* $Pr(B_i \leq X_i \leq B_i + c_i) = 1$, *then*

$$Pr(|\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i| \geq \epsilon) \leq e^{-\frac{2\epsilon^2}{\sum_{i=1}^{n} c_i^2}}$$

Do we need independence?

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

Conditioning Defines a Probability Space

Let $(\Omega, Pr(\cdot))$ be a probability space, and let $B \subseteq \Omega$ be an event in Ω, with $Pr(B) > 0$.

We show that $(B, Pr(\cdot \mid B))$ is a probability space:

1. For any $E \subseteq B$,

$$0 \le Pr(E \mid B) = \frac{Pr(E \cap B)}{Pr(B)} \le Pr(B \mid B) = 1$$

2. Let E_1 and E_2 be disjoint events in B,

$$\begin{aligned} Pr(E_1 \cup E_1 \mid B) &= \frac{Pr((E_1 \cup E_2) \cap B)}{Pr(B)} \\ &= \frac{Pr(E_1 \cap B)}{Pr(B)} + \frac{Pr(E_2 \cap B)}{Pr(B)} \\ &= Pr(E_1 \mid B) + Pr(E_2 \mid B) \end{aligned}$$

Conditional Expectation

Definition

$$\mathbf{E}[Y \mid Z = z] = \sum_{y} y \Pr(Y = y \mid Z = z) \ ,$$

where the summation is over all y in the range of Y.

Note that $\mathbf{E}[Y \mid Z]$ is a random variable (a function of Z)

Lemma

For any random variables X and Y,

$$\mathbf{E}[X] = \mathbf{E}_Y[\mathbf{E}_X[X \mid Y]] = \sum_{y} \Pr(Y = y)\mathbf{E}[X \mid Y = y] \ ,$$

where the sum is over all values in the range of Y.

Lemma

For any random variables X *and* Y,

$$\mathbf{E}[X] = \mathbf{E}_Y[\mathbf{E}_X[X \mid Y]] = \sum_y \Pr(Y = y)\mathbf{E}[X \mid Y = y] \;,$$

where the sum is over all values in the range of Y.

Proof.

$$\begin{aligned}
& \sum_y \Pr(Y = y)\mathbf{E}[X \mid Y = y] \\
= \; & \sum_y \Pr(Y = y) \sum_x x \Pr(X = x \mid Y = y) \\
= \; & \sum_x \sum_y x \Pr(X = x \mid Y = y) \Pr(Y = y) \\
= \; & \sum_x \sum_y x \Pr(X = x \cap Y = y) = \sum_x x \Pr(X = x) = \mathbf{E}[X].
\end{aligned}$$

□

Martingales

Definition

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

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

Definition

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

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

How to read it

$\mathbf{E}[Z_{n+1}|X_0, X_1, \ldots, X_n] = Z_n$ is a short form for:

$$\begin{aligned} & \mathbf{E}[Z_{n+1}|X_0 = x_0, X_1 = x_1, \ldots, X_n = x_n] \\ = \quad & Z_n|[X_0 = x_0, X_1 = x_1, \ldots, X_n = x]_n = z_n \end{aligned}$$

Since conditioning on $[X_0 = x_0, X_1 = x_1, \ldots, X_n = x]$, Z_n is a constant.
In many applications we just use $E[Z_{n+1}|Z_n] = Z_n$ which stands for $E[Z_{n+1}|Z_n = z_n] = z_n$.

Example: Sequence of Fair Games

- $X_i =$ amount won/lost in i-th game. $E[|X_i|] < \infty$.
- Fair game: $\mathbf{E}[X_i] = 0$.
- $Z_i =$ total winnings at end of i-th game.

Lemma

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

Proof.

$\mathbf{E}[|Z_i|] \leq \sum_{j=1}^{i} \mathbf{E}[|X_j|] < \infty$, and
$\mathbf{E}[Z_{i+1}|X_1, X_2, \ldots, X_i] = Z_i + \mathbf{E}[X_{i+1}] = Z_i$. □

The outcomes of the games do not need to be independent.

Efficient Market Hypothesis

The efficient markets hypothesis (EMH) maintains that market prices fully reflect all available information. Samuelson (1965), Fama (1963);

For simplicity assume an asset that is paying no dividend, and assume 0 interest rate (so value is not discounted in time).

Let X_t be the price of a unit asset at time t.
If I know that at time $t+1$ the price will be $X_{t+1} = c$, I will not sale the asset now for less than c.
If I know that at time $t+1$ the price will be $X_{t+1} = c$, I will not buy the asset now for more than c.

$$X_t = E[X_{t+1} \mid X_0, \ldots X_t]$$

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

Gambling Strategies

You play series of fair games. (If you bet (pay) i, with probability $1/2$ you win $2i$, else 0.)

Game 1: you bet $\$1$.

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

$X_i =$ amount won/lost in ith game.

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

Assume that (before starting to play) you decide to quit after exactly k games.
What is $E[Z_i]$?

Lemma

Let $Z_0, Z_1, Z_{2,} \ldots$ be a martingale with respect to $X_0, X_1, \ldots$. For any fixed n,

$$\mathbf{E}_{X[0:n]}[Z_n] = \mathbf{E}_{X_0}[Z_0] \ .$$

($X[0:i] = X_0, \ldots, X_i$)

Proof.

Since Z_i is a martingale $\mathbf{E}_{X_i}[Z_i | X_0, X_1, \ldots, X_{i-1}] = Z_{i-1}$.
Then

$$\mathbf{E}_{X[0:i-1]}[Z_{i-1}] = \mathbf{E}_{X[0:i-1]}[\mathbf{E}_{X_i}[Z_i | X_0, X_1, \ldots, X_{i-1}]] = \mathbf{E}_{X[0:i]}[Z_i]$$

Thus,

$$\mathbf{E}_{X[0:n]}[Z_n] = \mathbf{E}_{X[0:n-1]}[Z_{n-1}] = \ldots, = \mathbf{E}[Z_0]$$

□

Gambling Strategies

You play series of fair games. (If you bet (pay) i, with probability $1/2$ you win $2i$, else 0.)

Assume that before you start playing:

- You decide to stop after exactly k games.
- You decide to stop after each game with probability $1/2$.
- You decide to stop after k-th loss.
- You decide to stop before the k-th loss.
- You decide to stop when you win \$1000.

What is $E[Z_i]$ in each case?

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 I win*: is not a stopping time.

Martingale Optional 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:

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

Proof of Martingale Stopping Theorem (Sketch)

Define a sequence $Y_0, Y_1, \ldots$ such that

$$Y_i = \begin{cases} Z_i & \text{if } T > i \\ Z_T & \text{if } T \leq i \end{cases}$$

Lemma

The sequence $Y_0, Y_1, \ldots$ is a martingale with respect to $Z_0, Z_1, \ldots$.

Proof.

1. Y_n is determined by $Z_0, \ldots, Z_n$.
2. $E[|Y_n|] \leq \max_{0 \leq i \leq n} E[|X_i|] \leq \sum_{i=1}^{n} E[|X_i|] < \infty$
3. $\mathbf{E}[Y_{n+1}|Z_0, Z_1, \ldots, Z_n] = Y_n + E_{Z_{n+1}}[(Y_{n+1} - Y_n)1_{(T>n)}|Z_0, Z_1, \ldots, Z_n] = Y_n + E_{Z_{n+1}}[(Z_{n+1} - Z_n)|Z_0, Z_1, \ldots, Z_n]Pr(T > n) = Y_n$;

   Since $Pr(T > n)$ is independent of Z_{n+1}, and $E[(Z_{n+1} - Z_n)] = 0$.

□

Since $Y_0, Y_1, \ldots$ is a martingale, for any $n \geq 0$, $E[Y_n] = E[Z_0]$, and

$$\lim_{n\to\infty} E[Y_n] = E[Y_0] = E[Z_0].$$

Since T is finite, $Z_T = \lim_{n\to\infty} Z_{\min(n,T)} = lim_{n\to\infty} Y_n$.

We want to show that $E[Z_T] = \lim_{n\to\infty} E[Y_n] = E[Z_0]$.

This is not always true:
Example: Let $W_1, W_2, \ldots$ be random variables with distributions:

$$W_n = \begin{cases} n & \text{with probability} \frac{1}{n} \\ 0 & \text{otherwise} \end{cases}$$

$$\lim_{n\to\infty} W_n = 0,$$

but

$$\lim_{n\to\infty} E[W_n] = 1.$$

Since $Y_0, Y_1, \ldots$ is a martingale, for any $n \geq 0$, $E[Y_n] = E[Z_0]$, and

$$\lim_{n\to\infty} E[Y_n] = E[Y_0] = E[Z_0].$$

Since T is finite, $Z_T = \lim_{n\to\infty} Z_{\min(n,T)} = lim_{n\to\infty} Y_n$.
We want to show that $E[Z_T] = \lim_{n\to\infty} E[Y_n] = E[Z_0]$.

We use a simple version of the Dominated Convergence Theorem:

Theorem

Let $W_0, W_1, \ldots$ be a sequence of random variables such that $\lim_{n\to\infty} W_n = W$ (pointwise), and $max_i|W_i| \leq M$, where M is either a constant or a random variable with $E[|M|] < \infty$, then

$$\lim_{n\to\infty} E[W_n] = E[W].$$

Proof of Martingale Stopping Theorem (Sketch)

Since T is finite, $\lim_{n\to\infty} Y_n = \lim_{n\to\infty} Z_{\min(n,T)} = Z_T$.

We need to show that $|Y_n| \leq M$.

1. there is a constant c such that, for all i, $|Z_i| \leq c$ - $|Y_n| \leq max_{0\leq i\leq n}|Z_i| \leq c$, $c = M < \infty$.
2. T is bounded - $|Y_n| \leq max_{0\leq i\leq \max T}|Z_i| \leq M < \infty$
3. $\mathbf{E}[T] < \infty$, and there is a constant c such that $\mathbf{E}\left[|Z_{i+1} - Z_i|\middle|X_1,\ldots,X_i\right] < c$

$$Y_n = Z_0 + \sum_{i=1}^{\infty}(Z_{i+1} - Z_i)\mathbf{1}_{i\leq T} \leq |Z_0| + \sum_{i=1}^{\infty}|Z_{i+1} - Z_i|\mathbf{1}_{i\leq T} = M.$$

$$\begin{aligned} E[|M|] &= E[|Z_0|] + \sum_{i=1}^{\infty}\mathbf{E}[\mathbf{E}\left[|Z_{i+1} - Z_i|\middle|X_1,\ldots,X_i\right]\mathbf{1}_{i\leq T}] \\ &\leq E[|Z_0|] + c\sum_{i=1}^{\infty}Pr(T \geq i) \\ &\leq E[|Z_0|] + cE[T] < \infty \end{aligned}$$

Martingale Stopping Theorem Applications

We play a sequence of fair games with:

1. T is bounded, $\mathbf{E}[Z_T] = \mathbf{E}[Z_0]$.
2. $E[T] < \infty$ and $E[|X_i|] < \infty$, $\mathbf{E}[Z_T] = \mathbf{E}[Z_0]$.
3. Double the bet until the first win. $\mathbf{E}[T] = 2$ but $\mathbf{E}\left[|Z_{i+1} - Z_i| \middle| X_1, \ldots, X_i\right]$ is unbounded.
4. T is the first time we made $\$1000$: $\mathbf{E}[T]$ is unbounded.
5. Stop before the first loss. Not a stopping time.

Example: The Gambler's Ruin

- Consider a sequence of independent, fair 2-player gambling games.
- In each round, each player wins or loses $\$1$ with probability $\frac{1}{2}$.
- $X_i =$ 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.
  - $Z_0 = 0$.
- Game ends when one player runs out of money
  - Player 1 must stop when she loses net ℓ_1 dollars ($Z_t = -\ell_1$)
  - Player 2 terminates when she loses net ℓ_2 dollars ($Z_t = \ell_2$).
- $q =$ probability game ends with player 1 winning 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

- $S_i =$ number of votes **A** is leading by after i votes counted
  - If **A** is trailing: $S_i < 0$.
  - $S_n = a - b$.
- For $0 \leq k \leq n-1$: $X_k = \frac{S_{n-k}}{n-k}$.
- Consider $X_0 = \frac{a-b}{a+b}, X_1, \ldots, X_{n-1}$.
  - This sequence goes backward in time!

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

Example: A Ballot Theorem

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

This sequence goes backward in time! Equivalent to starting with all the votes counted, and removing at each step a random vote from the count.

$$\begin{aligned}
\mathbf{E}[X_{k+1}|X_0, X_1, \ldots, X_{k1}] &= \mathbf{E}\left[\frac{S_{n-k-1}}{n-k-1}\middle| S_n, \ldots, S_{n-(k-1)}\right] \\
&= (\frac{S_{n-k}}{n-k})\frac{S_{n-k}-1}{n-k-1} + (1-\frac{S_{n-k}}{n-k})\frac{S_{n-k}}{n-k-1} \\
&= \frac{S_{n-k}(n-k-1)}{(n-k-1)(n-k)} \\
&= X_{k-1}
\end{aligned}$$

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

Example: A Ballot Theorem

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

- Conditioning on $X_0, X_1, \ldots, X_{k-1}$: equivalent to conditioning on $S_n, S_{n-1}, \ldots, S_{n-(k-1)}$,
- $S_{n-k} = a_{n-k} - b_{n-k} =$ the number of votes for **A** minus the number of votes for **B** after the first $n-k$ votes are counted.
- $(n-(k-1))$-th vote: random vote among the first $n-k+1$ votes.

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

$$S_{n-k} = \begin{cases} S_{n-(k-1)} + 1 & \text{with prob. } \frac{n-k+1-a_{n-(k-1)}}{n-(k-1)} \\ S_{n-(k-1)} - 1 & \text{with prob. } \frac{a_{n-(k-1)}}{n-(k-1)} \end{cases}$$

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

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

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

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

Example: A Ballot Theorem

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

- 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.

1. **A** leads throughout the count.

For $0 \le k \le n-1$: $S_{n-k} > 0$, 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 $X_T = 1$.

[$X_{n-1} = S_1$ cannot be 0. If it's -1, then there must be $k < n-1$ such that $X_k = 0$.]

2. **A** does not lead throughout the count.

For some k: $S_k = 0$. Then $X_k = 0$.

$T = k < n - 1$.

$X_T = 0$.

Example: A Ballot Theorem

Putting all together:

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

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

That is

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