

CMPSCI 711: “Really Advanced Algorithms”

Lecture 15 & 16 – Martingales

Andrew McGregor

Two Examples

Martingale Definitions

Azuma-Hoeffding Inequality

Outline

Two Examples

Martingale Definitions

Azuma-Hoeffding Inequality

Example Application: Ballot Counting

- ▶ Suppose two candidates, A and B , run for an election.
- ▶ A gets a votes and B gets $b < a$ votes.
- ▶ If the votes are counted in random order, what is probability that A is always ahead during the voting?

Example Application: Gambling

- ▶ Suppose a gambler places bets on a sequence of fair games: bets can increase/decrease based on history and he can stop at any time.
- ▶ Let X_t be amount he wins at step t (could be negative),
- ▶ Let $Z_t = \sum_{i \in [t]} X_i$ be total winning at end of t -th step.
- ▶ Can he win money if he comes up with a clever strategy about when to quit?

Outline

Two Examples

Martingale Definitions

Azuma-Hoeffding Inequality

Martingales

Definition

Given two r.v.'s X, Y , $\mathbb{P}[X = x|Y]$ is the r.v. taking value $\mathbb{P}[X = x|Y = y]$ with prob. $\mathbb{P}[Y = y]$. Consequently $\mathbb{E}[X|Y]$ is r.v. taking value $\sum_x x\mathbb{P}[X = x|Y = y]$ with prob. $\mathbb{P}[Y = y]$.

Martingales

Definition

Given two r.v.'s X, Y , $\mathbb{P}[X = x|Y]$ is the r.v. taking value $\mathbb{P}[X = x|Y = y]$ with prob. $\mathbb{P}[Y = y]$. Consequently $\mathbb{E}[X|Y]$ is r.v. taking value $\sum_x x\mathbb{P}[X = x|Y = y]$ with prob. $\mathbb{P}[Y = y]$.

Definition

A sequence of r.v.'s $Z_0, Z_1, \dots$ is a martingale with respect to sequence $X_0, X_1, \dots$ if for all $n \geq 0$:

- ▶ Z_n is a function of $X_0, X_1, \dots, X_n$;
- ▶ $\mathbb{E}[|Z_n|] < \infty$;
- ▶ $\mathbb{E}[Z_{n+1}|X_0, \dots, X_n] = Z_n$

$Z_0, Z_1, \dots$ is a martingale if it's a martingale with respect to itself.

Martingales

Definition

Given two r.v.'s X, Y , $\mathbb{P}[X = x|Y]$ is the r.v. taking value $\mathbb{P}[X = x|Y = y]$ with prob. $\mathbb{P}[Y = y]$. Consequently $\mathbb{E}[X|Y]$ is r.v. taking value $\sum_x x\mathbb{P}[X = x|Y = y]$ with prob. $\mathbb{P}[Y = y]$.

Definition

A sequence of r.v.'s $Z_0, Z_1, \dots$ is a martingale with respect to sequence $X_0, X_1, \dots$ if for all $n \geq 0$:

- ▶ Z_n is a function of $X_0, X_1, \dots, X_n$;
- ▶ $\mathbb{E}[|Z_n|] < \infty$;
- ▶ $\mathbb{E}[Z_{n+1}|X_0, \dots, X_n] = Z_n$

$Z_0, Z_1, \dots$ is a martingale if it's a martingale with respect to itself.

E.g., In the gambling problem:

$$\mathbb{E}[Z_{i+1}|X_1, \dots, X_i] = Z_i + \mathbb{E}[X_{i+1}] = Z_i$$

Simple Lemmas

Lemma

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|Z]]$$

Lemma

$$\mathbb{E}[X|Y] = \mathbb{E}[X|f(Y)] \text{ if } f \text{ is invertible.}$$

Lemma

$$\mathbb{E}[X|X, Y] = X$$

Deciding when to stop ahead of time...

Suppose, before he starts, the gambler decides he'll quit after the n -th game. What's the gambler's expected winning?

Deciding when to stop ahead of time...

Suppose, before he starts, the gambler decides he'll quit after the n -th game. What's the gambler's expected winning?

Lemma

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

$$\mathbb{E}[Z_n] = \mathbb{E}[Z_0]$$

Deciding when to stop ahead of time...

Suppose, before he starts, the gambler decides he'll quit after the n -th game. What's the gambler's expected winning?

Lemma

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

$$\mathbb{E}[Z_n] = \mathbb{E}[Z_0]$$

Proof.

- By the martingale definition: $Z_i = \mathbb{E}[Z_{i+1} | X_0, \dots, X_i]$



Deciding when to stop ahead of time...

Suppose, before he starts, the gambler decides he'll quit after the n -th game. What's the gambler's expected winning?

Lemma

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

$$\mathbb{E}[Z_n] = \mathbb{E}[Z_0]$$

Proof.

- ▶ By the martingale definition: $Z_i = \mathbb{E}[Z_{i+1} | X_0, \dots, X_i]$
- ▶ Take expectation: $\mathbb{E}[Z_i] = \mathbb{E}[\mathbb{E}[Z_{i+1} | X_0, \dots, X_i]] = \mathbb{E}[Z_{i+1}]$



Deciding when to stop ahead of time...

Suppose, before he starts, the gambler decides he'll quit after the n -th game. What's the gambler's expected winning?

Lemma

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

$$\mathbb{E}[Z_n] = \mathbb{E}[Z_0]$$

Proof.

- ▶ By the martingale definition: $Z_i = \mathbb{E}[Z_{i+1} | X_0, \dots, X_i]$
- ▶ Take expectation: $\mathbb{E}[Z_i] = \mathbb{E}[\mathbb{E}[Z_{i+1} | X_0, \dots, X_i]] = \mathbb{E}[Z_{i+1}]$
- ▶ Repeating argument gives result.



Deciding when to stop based on evolving sequence...

Definition (Stopping Time)

A non-negative, integer-valued random variable T is a stopping time for sequence $\{Z_n : n \geq 0\}$ if the event $T = n$ depends only on the value of the random variables $Z_0, Z_1, \dots, Z_n$.

Deciding when to stop based on evolving sequence...

Definition (Stopping Time)

A non-negative, integer-valued random variable T is a stopping time for sequence $\{Z_n : n \geq 0\}$ if the event $T = n$ depends only on the value of the random variables $Z_0, Z_1, \dots, Z_n$.

Theorem (Martingale Stopping Theorem)

If $Z_0, Z_1, \dots$ is a martingale with respect to $X_0, X_2, \dots$ and if T is a stopping time for $X_0, X_2, \dots$ then

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

whenever one of the following holds:

- ▶ Z_i are bounded, i.e., for constant c , $|Z_i| \leq c$ for all i ;
- ▶ T is bounded;
- ▶ $\mathbb{E}[T] < \infty$ and there exists constant c such that

$$\mathbb{E}[|Z_{i+1} - Z_i| | X_0, \dots, X_i] < c$$

Ballot Example (1/4)

- ▶ Two candidates, A and B , run for an election with n voters.
- ▶ A gets a votes and B gets $b < a$ votes.
- ▶ If the votes are counted in a random order, what is the probability that A is always ahead during the voting?

Ballot Example (2/4)

- ▶ Let $S_k = \#$ votes by which A leads after k votes counted
- ▶ Hence, A has $(k + S_k)/2$ votes and B has $(k - S_k)/2$ votes after k votes have been counted

Ballot Example (2/4)

- ▶ Let $S_k = \#$ votes by which A leads after k votes counted
- ▶ Hence, A has $(k + S_k)/2$ votes and B has $(k - S_k)/2$ votes after k votes have been counted
- ▶ Note that $S_k = S_{k+1} \pm 1$ and

$$\mathbb{P}[S_k = S_{k+1} + 1 | S_{k+1}] = \frac{(k + 1 - S_{k+1})/2}{k + 1}$$

$$\mathbb{P}[S_k = S_{k+1} - 1 | S_{k+1}] = \frac{(k + 1 + S_{k+1})/2}{k + 1}$$

Ballot Example (3/4)

- ▶ Define martingale $X_k = S_{n-k}/(n-k)$ for $0 \leq k \leq n-1$:

Ballot Example (3/4)

- ▶ Define martingale $X_k = S_{n-k}/(n-k)$ for $0 \leq k \leq n-1$:

$$\mathbb{E}[X_k | X_0, \dots, X_{k-1}]$$

Ballot Example (3/4)

- ▶ Define martingale $X_k = S_{n-k}/(n-k)$ for $0 \leq k \leq n-1$:

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

Ballot Example (3/4)

- Define martingale $X_k = S_{n-k}/(n-k)$ for $0 \leq k \leq n-1$:

$$\begin{aligned} & \mathbb{E}[X_k | X_0, \dots, X_{k-1}] \\ = & \mathbb{E}\left[\frac{S_{n-k}}{n-k} \mid S_n, \dots, S_{n-k+1}\right] \\ = & \frac{1}{n-k} (\mathbb{P}[S_{n-k} = S_{n-k+1} + 1 | S_{n-k+1}] (S_{n-k+1} + 1) \\ & + \mathbb{P}[S_{n-k} = S_{n-k+1} - 1 | S_{n-k+1}] (S_{n-k+1} - 1)) \end{aligned}$$

Ballot Example (3/4)

- Define martingale $X_k = S_{n-k}/(n-k)$ for $0 \leq k \leq n-1$:

$$\begin{aligned} & \mathbb{E}[X_k | X_0, \dots, X_{k-1}] \\ = & \mathbb{E}\left[\frac{S_{n-k}}{n-k} \mid S_n, \dots, S_{n-k+1}\right] \\ = & \frac{1}{n-k} (\mathbb{P}[S_{n-k} = S_{n-k+1} + 1 | S_{n-k+1}](S_{n-k+1} + 1) \\ & \quad + \mathbb{P}[S_{n-k} = S_{n-k+1} - 1 | S_{n-k+1}](S_{n-k+1} - 1)) \\ = & \frac{1}{n-k} \left(\frac{(n-k+1 - S_{n-k+1})/2}{n-k+1} (S_{n-k+1} + 1) \right. \\ & \quad \left. + \frac{(n-k+1 + S_{n-k+1})/2}{n-k+1} (S_{n-k+1} - 1) \right) \end{aligned}$$

Ballot Example (3/4)

- Define martingale $X_k = S_{n-k}/(n-k)$ for $0 \leq k \leq n-1$:

$$\begin{aligned} & \mathbb{E}[X_k | X_0, \dots, X_{k-1}] \\ = & \mathbb{E}\left[\frac{S_{n-k}}{n-k} \mid S_n, \dots, S_{n-k+1}\right] \\ = & \frac{1}{n-k} (\mathbb{P}[S_{n-k} = S_{n-k+1} + 1 | S_{n-k+1}](S_{n-k+1} + 1) \\ & \quad + \mathbb{P}[S_{n-k} = S_{n-k+1} - 1 | S_{n-k+1}](S_{n-k+1} - 1)) \\ = & \frac{1}{n-k} \left(\frac{(n-k+1 - S_{n-k+1})/2}{n-k+1} (S_{n-k+1} + 1) \right. \\ & \quad \left. + \frac{(n-k+1 + S_{n-k+1})/2}{n-k+1} (S_{n-k+1} - 1) \right) \\ = & \frac{S_{n-k+1}}{n-k+1} = X_{k-1} \end{aligned}$$

Ballot Example (3/4)

- Define martingale $X_k = S_{n-k}/(n-k)$ for $0 \leq k \leq n-1$:

$$\begin{aligned} & \mathbb{E}[X_k | X_0, \dots, X_{k-1}] \\ &= \mathbb{E}\left[\frac{S_{n-k}}{n-k} \mid S_n, \dots, S_{n-k+1}\right] \\ &= \frac{1}{n-k} (\mathbb{P}[S_{n-k} = S_{n-k+1} + 1 | S_{n-k+1}](S_{n-k+1} + 1) \\ &\quad + \mathbb{P}[S_{n-k} = S_{n-k+1} - 1 | S_{n-k+1}](S_{n-k+1} - 1)) \\ &= \frac{1}{n-k} \left(\frac{(n-k+1 - S_{n-k+1})/2}{n-k+1} (S_{n-k+1} + 1) \right. \\ &\quad \left. + \frac{(n-k+1 + S_{n-k+1})/2}{n-k+1} (S_{n-k+1} - 1) \right) \\ &= \frac{S_{n-k+1}}{n-k+1} = X_{k-1} \end{aligned}$$

- Clearly that $\mathbb{E}[|X_{k+1}|] \leq n < \infty$.

Ballot Example (4/4)

- ▶ Let $T = \min\{n - 1\} \cup \{k : X_k = 0\}$

Ballot Example (4/4)

- ▶ Let $T = \min\{n - 1\} \cup \{k : X_k = 0\}$
- ▶ T is bounded stopping time; Martingale Stopping Thm gives

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

Ballot Example (4/4)

- ▶ Let $T = \min\{n - 1\} \cup \{k : X_k = 0\}$
- ▶ T is bounded stopping time; Martingale Stopping Thm gives

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

- ▶ If A always leads then $T = n - 1$ and $X_T = S_1 = 1$

Ballot Example (4/4)

- ▶ Let $T = \min\{n - 1\} \cup \{k : X_k = 0\}$
- ▶ T is bounded stopping time; Martingale Stopping Thm gives

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

- ▶ If A always leads then $T = n - 1$ and $X_T = S_1 = 1$
- ▶ If A doesn't always lead then $T = k < n - 1$ and $X_T = 0$

Ballot Example (4/4)

- ▶ Let $T = \min\{n - 1\} \cup \{k : X_k = 0\}$
- ▶ T is bounded stopping time; Martingale Stopping Thm gives

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

- ▶ If A always leads then $T = n - 1$ and $X_T = S_1 = 1$
- ▶ If A doesn't always lead then $T = k < n - 1$ and $X_T = 0$
- ▶ Hence, $X_T = 0$ or $X_T = 1$ and

$$\frac{a - b}{a + b} = \mathbb{E}[X_T] = 1 \cdot \mathbb{P}[\text{A always leads}] + 0 \cdot \mathbb{P}[\text{A doesn't always lead}]$$

Wald's Equation

Theorem

Let $X_1, X_2, \dots$ be non-negative, 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

$$\mathbb{E} \left[\sum_{i=1}^T X_i \right] = \mathbb{E}[T] \mathbb{E}[X] .$$

Wald's Equation

Theorem

Let $X_1, X_2, \dots$ be non-negative, 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

$$\mathbb{E} \left[\sum_{i=1}^T X_i \right] = \mathbb{E}[T] \mathbb{E}[X] .$$

Proof.

- For $i \geq 1$, let $Z_i = \sum_{j=1}^i (X_j - \mathbb{E}[X])$

Wald's Equation

Theorem

Let $X_1, X_2, \dots$ be non-negative, 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

$$\mathbb{E} \left[\sum_{i=1}^T X_i \right] = \mathbb{E}[T] \mathbb{E}[X] .$$

Proof.

- ▶ For $i \geq 1$, let $Z_i = \sum_{j=1}^i (X_j - \mathbb{E}[X])$
- ▶ $Z_1, Z_2, \dots$ is martingale w.r.t. $X_1, X_2, \dots$ and $\mathbb{E}[Z_1] = 0$

Wald's Equation

Theorem

Let $X_1, X_2, \dots$ be non-negative, 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

$$\mathbb{E} \left[\sum_{i=1}^T X_i \right] = \mathbb{E}[T] \mathbb{E}[X] .$$

Proof.

- ▶ For $i \geq 1$, let $Z_i = \sum_{j=1}^i (X_j - \mathbb{E}[X])$
- ▶ $Z_1, Z_2, \dots$ is martingale w.r.t. $X_1, X_2, \dots$ and $\mathbb{E}[Z_1] = 0$
- ▶ $\mathbb{E}[T] < \infty$ and $\mathbb{E}[|Z_{i+1} - Z_i| | X_1, \dots, X_i] \leq 2\mathbb{E}[X]$, hence:

$$\mathbb{E}[Z_T] = \mathbb{E}[Z_1] = 0$$

Wald's Equation

Theorem

Let $X_1, X_2, \dots$ be non-negative, 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

$$\mathbb{E} \left[\sum_{i=1}^T X_i \right] = \mathbb{E}[T] \mathbb{E}[X] .$$

Proof.

- ▶ For $i \geq 1$, let $Z_i = \sum_{j=1}^i (X_j - \mathbb{E}[X])$
- ▶ $Z_1, Z_2, \dots$ is martingale w.r.t. $X_1, X_2, \dots$ and $\mathbb{E}[Z_1] = 0$
- ▶ $\mathbb{E}[T] < \infty$ and $\mathbb{E}[|Z_{i+1} - Z_i| | X_1, \dots, X_i] \leq 2\mathbb{E}[X]$, hence:

$$\mathbb{E}[Z_T] = \mathbb{E}[Z_1] = 0$$

- ▶ Results follows since $\mathbb{E}[Z_T] = \mathbb{E} \left[\sum_{j=1}^T X_j \right] - \mathbb{E}[T] \mathbb{E}[X]$

Application of Wald's Equation (1/2)

- ▶ Suppose n servers communicate using shared channel.
- ▶ At each step, each server sends a message with prob. $1/n$.
- ▶ Send is successful if only one message is sent.
- ▶ What's expected time until all servers have sent ≥ 1 message?

Application of Wald's Equation (2/2)

- ▶ Let N be # packets sent successfully until each server has sent at least one packet: $\mathbb{E}[N] = nH(n)$

Application of Wald's Equation (2/2)

- ▶ Let N be # packets sent successfully until each server has sent at least one packet: $\mathbb{E}[N] = nH(n)$
- ▶ Let t_i be time of i -th successful send

Application of Wald's Equation (2/2)

- ▶ Let N be # packets sent successfully until each server has sent at least one packet: $\mathbb{E}[N] = nH(n)$
- ▶ Let t_i be time of i -th successful send
- ▶ Let $r_i = t_i - t_{i-1}$: $\mathbb{P}[\text{one message sent}] = \binom{n}{1}(\frac{1}{n})(1 - \frac{1}{n})^{n-1}$

$$\mathbb{E}[r_i] = 1/\mathbb{P}[\text{one message sent}] =: 1/p$$

Application of Wald's Equation (2/2)

- ▶ Let N be # packets sent successfully until each server has sent at least one packet: $\mathbb{E}[N] = nH(n)$
- ▶ Let t_i be time of i -th successful send
- ▶ Let $r_i = t_i - t_{i-1}$: $\mathbb{P}[\text{one message sent}] = \binom{n}{1}(\frac{1}{n})(1 - \frac{1}{n})^{n-1}$

$$\mathbb{E}[r_i] = 1/\mathbb{P}[\text{one message sent}] =: 1/p$$

- ▶ Then $T = \sum_{i=1}^N r_i$ is time at which all servers have sent ≥ 1 message:

$$\mathbb{E}[T] = \mathbb{E}\left[\sum_{i \in [N]} r_i\right] = \mathbb{E}[N] \mathbb{E}[r_i] = nH(n)/p$$

Outline

Two Examples

Martingale Definitions

Azuma-Hoeffding Inequality

Azuma-Hoeffding Inequality

Theorem (Azuma-Hoeffding Inequality)

Let $X_0, \dots, X_n$ be a martingale with $|X_k - X_{k-1}| \leq c_k$. Then for all $t \geq 0$ and any $\lambda > 0$, $\mathbb{P}[|X_t - X_0| \geq \lambda] \leq 2e^{-\lambda^2/(2\sum_{k=1}^t c_k^2)}$

Azuma-Hoeffding Inequality

Theorem (Azuma-Hoeffding Inequality)

Let $X_0, \dots, X_n$ be a martingale with $|X_k - X_{k-1}| \leq c_k$. Then for all $t \geq 0$ and any $\lambda > 0$, $\mathbb{P}[|X_t - X_0| \geq \lambda] \leq 2e^{-\lambda^2/(2\sum_{k=1}^t c_k^2)}$

Proof.

- Let $Y_i = X_i - X_{i-1}$: $|Y_i| \leq c_i$ and

$$\mathbb{E}[Y_i | X_0, X_1, \dots, X_{i-1}] = \mathbb{E}[X_i | X_0, X_1, \dots, X_{i-1}] - X_{i-1} = 0$$

Azuma-Hoeffding Inequality

Theorem (Azuma-Hoeffding Inequality)

Let $X_0, \dots, X_n$ be a martingale with $|X_k - X_{k-1}| \leq c_k$. Then for all $t \geq 0$ and any $\lambda > 0$, $\mathbb{P}[|X_t - X_0| \geq \lambda] \leq 2e^{-\lambda^2/(2\sum_{k=1}^t c_k^2)}$

Proof.

- ▶ Let $Y_i = X_i - X_{i-1}$: $|Y_i| \leq c_i$ and

$$\mathbb{E}[Y_i | X_0, X_1, \dots, X_{i-1}] = \mathbb{E}[X_i | X_0, X_1, \dots, X_{i-1}] - X_{i-1} = 0$$

- ▶ By convexity of $e^{\alpha Y_i}$:

$$e^{\alpha Y_i} \leq \frac{1 - \frac{Y_i}{c_i}}{2} e^{-\alpha c_i} + \frac{1 + \frac{Y_i}{c_i}}{2} e^{\alpha c_i} = \frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} + \frac{Y_i}{2c_i} (e^{\alpha c_i} - e^{-\alpha c_i})$$

Azuma-Hoeffding Inequality

Theorem (Azuma-Hoeffding Inequality)

Let $X_0, \dots, X_n$ be a martingale with $|X_k - X_{k-1}| \leq c_k$. Then for all $t \geq 0$ and any $\lambda > 0$, $\mathbb{P}[|X_t - X_0| \geq \lambda] \leq 2e^{-\lambda^2/(2\sum_{k=1}^t c_k^2)}$

Proof.

- ▶ Let $Y_i = X_i - X_{i-1}$: $|Y_i| \leq c_i$ and

$$\mathbb{E}[Y_i | X_0, X_1, \dots, X_{i-1}] = \mathbb{E}[X_i | X_0, X_1, \dots, X_{i-1}] - X_{i-1} = 0$$

- ▶ By convexity of $e^{\alpha Y_i}$:

$$e^{\alpha Y_i} \leq \frac{1 - \frac{Y_i}{c_i}}{2} e^{-\alpha c_i} + \frac{1 + \frac{Y_i}{c_i}}{2} e^{\alpha c_i} = \frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} + \frac{Y_i}{2c_i} (e^{\alpha c_i} - e^{-\alpha c_i})$$

- ▶ Putting above two lines together:

$$\mathbb{E}\left[e^{\alpha Y_i} | X_0, X_1, \dots, X_{i-1}\right] \leq \frac{e^{\alpha c_i} + e^{-\alpha c_i}}{2} \leq e^{(\alpha c_i)^2/2}$$

Proof Continued

- ▶ Applying the previous bound repeatedly:

$$\mathbb{E} \left[e^{\alpha(X_t - X_0)} \right] = \mathbb{E} \left[\prod_{i \in [t]} e^{\alpha Y_i} \right] \leq e^{\sum_{i \in [t]} (\alpha c_i)^2 / 2}$$

Proof Continued

- ▶ Applying the previous bound repeatedly:

$$\mathbb{E} \left[e^{\alpha(X_t - X_0)} \right] = \mathbb{E} \left[\prod_{i \in [t]} e^{\alpha Y_i} \right] \leq e^{\sum_{i \in [t]} (\alpha c_i)^2 / 2}$$

- ▶ By Markov inequality:

$$\mathbb{P} [X_t - X_0 \geq \lambda] = \mathbb{P} \left[e^{\alpha(X_t - X_0)} \geq e^{\alpha\lambda} \right] \leq \frac{\mathbb{E} \left[e^{\alpha(X_t - X_0)} \right]}{e^{\alpha\lambda}}$$

Proof Continued

- ▶ Applying the previous bound repeatedly:

$$\mathbb{E} \left[e^{\alpha(X_t - X_0)} \right] = \mathbb{E} \left[\prod_{i \in [t]} e^{\alpha Y_i} \right] \leq e^{\sum_{i \in [t]} (\alpha c_i)^2 / 2}$$

- ▶ By Markov inequality:

$$\mathbb{P} [X_t - X_0 \geq \lambda] = \mathbb{P} \left[e^{\alpha(X_t - X_0)} \geq e^{\alpha\lambda} \right] \leq \frac{\mathbb{E} \left[e^{\alpha(X_t - X_0)} \right]}{e^{\alpha\lambda}}$$

- ▶ Combing above two lines and choosing $\alpha = \lambda / \sum_{k \in [t]} c_k^2$:

$$\mathbb{P} [X_t - X_0 \geq \lambda] \leq e^{\sum_{i \in [t]} (\alpha c_i)^2 / 2 - \alpha\lambda} \leq e^{-\lambda^2 / (2 \sum_{i \in [t]} c_i^2)}$$

Proof Continued

- ▶ Applying the previous bound repeatedly:

$$\mathbb{E} \left[e^{\alpha(X_t - X_0)} \right] = \mathbb{E} \left[\prod_{i \in [t]} e^{\alpha Y_i} \right] \leq e^{\sum_{i \in [t]} (\alpha c_i)^2 / 2}$$

- ▶ By Markov inequality:

$$\mathbb{P} [X_t - X_0 \geq \lambda] = \mathbb{P} \left[e^{\alpha(X_t - X_0)} \geq e^{\alpha\lambda} \right] \leq \frac{\mathbb{E} \left[e^{\alpha(X_t - X_0)} \right]}{e^{\alpha\lambda}}$$

- ▶ Combing above two lines and choosing $\alpha = \lambda / \sum_{k \in [t]} c_k^2$:

$$\mathbb{P} [X_t - X_0 \geq \lambda] \leq e^{\sum_{i \in [t]} (\alpha c_i)^2 / 2 - \alpha\lambda} \leq e^{-\lambda^2 / (2 \sum_{i \in [t]} c_i^2)}$$

- ▶ Bound for $\mathbb{P} [X_0 - X_t \geq \lambda]$ follows similarly.

Application to Pattern Matching

- ▶ Let $X = (X_1, \dots, X_n)$ be sequence of characters drawn independently and uniformly from a size s alphabet Σ

Application to Pattern Matching

- ▶ Let $X = (X_1, \dots, X_n)$ be sequence of characters drawn independently and uniformly from a size s alphabet Σ
- ▶ Let $B = (b_1, \dots, b_k)$ be a fixed set of characters from Σ .

Application to Pattern Matching

- ▶ Let $X = (X_1, \dots, X_n)$ be sequence of characters drawn independently and uniformly from a size s alphabet Σ
- ▶ Let $B = (b_1, \dots, b_k)$ be a fixed set of characters from Σ .
- ▶ Let F be number of occurrences of the fixed string B .

Application to Pattern Matching

- ▶ Let $X = (X_1, \dots, X_n)$ be sequence of characters drawn independently and uniformly from a size s alphabet Σ
- ▶ Let $B = (b_1, \dots, b_k)$ be a fixed set of characters from Σ .
- ▶ Let F be number of occurrences of the fixed string B .
- ▶ By linearity of expectation: $\mathbb{E}[F] = (n - k + 1)(1/s)^k$

Application to Pattern Matching

- ▶ Let $X = (X_1, \dots, X_n)$ be sequence of characters drawn independently and uniformly from a size s alphabet Σ
- ▶ Let $B = (b_1, \dots, b_k)$ be a fixed set of characters from Σ .
- ▶ Let F be number of occurrences of the fixed string B .
- ▶ By linearity of expectation: $\mathbb{E}[F] = (n - k + 1)(1/s)^k$
- ▶ Define the following martingale

$$Z_0 = \mathbb{E}[F], \quad Z_i = \mathbb{E}[F|X_1, \dots, X_i] \quad \text{for } 1 \leq i \leq n$$

Application to Pattern Matching

- ▶ Let $X = (X_1, \dots, X_n)$ be sequence of characters drawn independently and uniformly from a size s alphabet Σ
- ▶ Let $B = (b_1, \dots, b_k)$ be a fixed set of characters from Σ .
- ▶ Let F be number of occurrences of the fixed string B .
- ▶ By linearity of expectation: $\mathbb{E}[F] = (n - k + 1)(1/s)^k$
- ▶ Define the following martingale

$$Z_0 = \mathbb{E}[F], \quad Z_i = \mathbb{E}[F|X_1, \dots, X_i] \quad \text{for } 1 \leq i \leq n$$

- ▶ Because the $(i + 1)$ -th character occurs in at most k matches:

$$|Z_{i+1} - Z_i| = |\mathbb{E}[F|X_1, \dots, X_{i+1}] - \mathbb{E}[F|X_1, \dots, X_i]| \leq k$$

Application to Pattern Matching

- ▶ Let $X = (X_1, \dots, X_n)$ be sequence of characters drawn independently and uniformly from a size s alphabet Σ
- ▶ Let $B = (b_1, \dots, b_k)$ be a fixed set of characters from Σ .
- ▶ Let F be number of occurrences of the fixed string B .
- ▶ By linearity of expectation: $\mathbb{E}[F] = (n - k + 1)(1/s)^k$
- ▶ Define the following martingale

$$Z_0 = \mathbb{E}[F], \quad Z_i = \mathbb{E}[F|X_1, \dots, X_i] \quad \text{for } 1 \leq i \leq n$$

- ▶ Because the $(i + 1)$ -th character occurs in at most k matches:

$$|Z_{i+1} - Z_i| = |\mathbb{E}[F|X_1, \dots, X_{i+1}] - \mathbb{E}[F|X_1, \dots, X_i]| \leq k$$

- ▶ Note that $Z_n = F$: By applying Azuma's inequality:

$$\mathbb{P}[|Z_n - Z_0| \geq \lambda] \leq 2e^{-\lambda^2/(2nk^2)}$$

Azuma-Hoeffding Inequality Generalization

Theorem (Azuma-Hoeffding Inequality)

Let $X_0, \dots, X_n$ be a martingale with

$$B_k \leq X_k - X_{k-1} \leq B_k + d_k$$

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

$$\mathbb{P}[|X_t - X_0| \geq \lambda] \leq 2e^{-2\lambda^2 / (\sum_{k=1}^t d_k^2)}$$

Applying Azuma-Hoeffding Inequality

- For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with $\mathbb{E}[|f(X_1, X_2, \dots, X_n)|] < \infty$ for random variables $X_1, \dots, X_n$ define the martingale:

$$Z_0 = \mathbb{E}[f(X_1, X_2, \dots, X_n)]$$

$$Z_k = \mathbb{E}[f(X_1, X_2, \dots, X_n) | X_1, \dots, X_k]$$

Applying Azuma-Hoeffding Inequality

- ▶ For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with $\mathbb{E}[|f(X_1, X_2, \dots, X_n)|] < \infty$ for random variables $X_1, \dots, X_n$ define the martingale:

$$Z_0 = \mathbb{E}[f(X_1, X_2, \dots, X_n)]$$

$$Z_k = \mathbb{E}[f(X_1, X_2, \dots, X_n) | X_1, \dots, X_k]$$

- ▶ Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies the Lipschitz condition with bound c : i.e., for any $i, x_1, \dots, x_n, y_i$

$$|f(x_1, \dots, x_n) - f(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)| \leq c$$

Applying Azuma-Hoeffding Inequality

- ▶ For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with $\mathbb{E}[|f(X_1, X_2, \dots, X_n)|] < \infty$ for random variables $X_1, \dots, X_n$ define the martingale:

$$Z_0 = \mathbb{E}[f(X_1, X_2, \dots, X_n)]$$

$$Z_k = \mathbb{E}[f(X_1, X_2, \dots, X_n) | X_1, \dots, X_k]$$

- ▶ Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies the Lipschitz condition with bound c : i.e., for any $i, x_1, \dots, x_n, y_i$

$$|f(x_1, \dots, x_n) - f(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)| \leq c$$

- ▶ If X_i 's are independent, it can be shown that

$$B_k \leq X_k - X_{k-1} \leq B_k + c$$

where

$$B_k = \inf_y \mathbb{E}[f(X_1, \dots, X_n) | X_1, \dots, X_{k-1}, X_k = y] \\ - \mathbb{E}[f(X_1, \dots, X_n) | X_1, \dots, X_{k-1}]$$

Applying Azuma-Hoeffding Inequality

- ▶ For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with $\mathbb{E}[|f(X_1, X_2, \dots, X_n)|] < \infty$ for random variables $X_1, \dots, X_n$ define the martingale:

$$Z_0 = \mathbb{E}[f(X_1, X_2, \dots, X_n)]$$

$$Z_k = \mathbb{E}[f(X_1, X_2, \dots, X_n) | X_1, \dots, X_k]$$

- ▶ Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfies the Lipschitz condition with bound c : i.e., for any $i, x_1, \dots, x_n, y_i$

$$|f(x_1, \dots, x_n) - f(x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)| \leq c$$

- ▶ If X_i 's are independent, it can be shown that

$$B_k \leq X_k - X_{k-1} \leq B_k + c$$

where

$$B_k = \inf_y \mathbb{E}[f(X_1, \dots, X_n) | X_1, \dots, X_{k-1}, X_k = y] \\ - \mathbb{E}[f(X_1, \dots, X_n) | X_1, \dots, X_{k-1}]$$

- ▶ Hence, by Azuma, $\mathbb{P}[|X_t - X_0| \geq \lambda] \leq 2e^{-2\lambda^2 / (\sum_{k=1}^t c^2)}$