Advanced Algorithms
(Martingales and the Method of Bounded Differences)
(Some) Concentration Inequalities

Question: probability that $X$ deviates more than $\delta$ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$, then:

for any $\delta > 0$,
$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{1+\delta}}\right)^\mu$$

for $0 < \delta < 1$,
$$\mathbb{P}(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^\mu$$

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$, then:

for any $t > 0$,
$$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n} (b_i-a_i)^2}\right)$$
$$\mathbb{P}(X \leq \mathbb{E}(X) - t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n} (b_i-a_i)^2}\right)$$
Question: probability that $X$ deviates more than $\delta$ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$, then:

for any $\delta > 0$,
$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq \left( \frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \right)^{\mu}$$

for $0 < \delta < 1$,
$$\mathbb{P}(X \leq (1 - \delta)\mu) \leq \left( \frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}} \right)^{\mu}$$

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$, then:

for any $t > 0$,
$$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \exp \left( -\sum_{i=1}^{n} \frac{2t^2}{(b_i - a_i)^2} \right)$$

$$\mathbb{P}(X \leq \mathbb{E}(X) - t) \leq \exp \left( -\sum_{i=1}^{n} \frac{2t^2}{(b_i - a_i)^2} \right)$$
Conditional Probability

The *conditional probability* that event $\mathcal{E}_1$ occurs, given that event $\mathcal{E}_2$ occurs, is:

$$P(\mathcal{E}_1 | \mathcal{E}_2) = \frac{P(\mathcal{E}_1 \cap \mathcal{E}_2)}{P(\mathcal{E}_2)}$$
Conditional Probability

The *conditional probability* that event $\mathcal{E}_1$ occurs, given that event $\mathcal{E}_2$ occurs, is:

$$
P(\mathcal{E}_1 | \mathcal{E}_2) = \frac{P(\mathcal{E}_1 \cap \mathcal{E}_2)}{P(\mathcal{E}_2)}$$

*Example:*
roll a fair six-sided dice
$\mathcal{E}_1 =$ the outcome is six
$\mathcal{E}_2 =$ the outcome is an even number
Conditional Probability

The *conditional probability* that event $\mathcal{E}_1$ occurs, given that event $\mathcal{E}_2$ occurs, is:

$$P(\mathcal{E}_1|\mathcal{E}_2) = \frac{P(\mathcal{E}_1 \cap \mathcal{E}_2)}{P(\mathcal{E}_2)}$$

*Example:*

roll a fair six-sided dice

$\mathcal{E}_1$ = the outcome is six

$\mathcal{E}_2$ = the outcome is an even number

$$P(\mathcal{E}_1|\mathcal{E}_2) = \frac{1/6}{1/2} = \frac{1}{3}$$
Conditional Expectation

The *conditional expectation* of a random variable $Y$ with respect to an event $\mathcal{E}$ is:

$$
\mathbb{E}(Y \mid \mathcal{E}) = \sum_y y \cdot \mathbb{P}(Y = y \mid \mathcal{E})
$$

In particular, if the event $\mathcal{E}$ is $X = x$, then:

$$
\mathbb{E}(Y \mid X = x) = \sum_y y \cdot \mathbb{P}(Y = y \mid X = x)
$$
Conditional Expectation

The conditional expectation of a random variable $Y$ with respect to an event $\mathcal{E}$ is:

$$
\mathbb{E}(Y \mid \mathcal{E}) = \sum_y y \cdot \mathbb{P}(Y = y \mid \mathcal{E})
$$

In particular, if the event $\mathcal{E}$ is $X = x$, then:

$$
\mathbb{E}(Y \mid X = x) = \sum_y y \cdot \mathbb{P}(Y = y \mid X = x)
$$

Example:

sample a human being uniformly at random

$Y$: height of the chosen human being

$X$: country of origin of the chosen human being
Conditional Expectation

The *conditional expectation* of a random variable $Y$ with respect to an event $\mathcal{E}$ is:

$$\mathbb{E}(Y \mid \mathcal{E}) = \sum_{y} y \cdot \mathbb{P}(Y = y \mid \mathcal{E})$$

In particular, if the event $\mathcal{E}$ is $X = x$, then:

$$\mathbb{E}(Y \mid X = x) = \sum_{y} y \cdot \mathbb{P}(Y = y \mid X = x)$$

*Example:*

sample a human being uniformly at random

$Y$: height of the chosen human being

$X$: country of origin of the chosen human being

$\mathbb{E}(Y) =$?
Conditional Expectation

The conditional expectation of a random variable \( Y \) with respect to an event \( \mathcal{E} \) is:

\[
\mathbb{E}(Y \mid \mathcal{E}) = \sum_y y \cdot \mathbb{P}(Y = y \mid \mathcal{E})
\]

In particular, if the event \( \mathcal{E} \) is \( X = x \), then:

\[
\mathbb{E}(Y \mid X = x) = \sum_y y \cdot \mathbb{P}(Y = y \mid X = x)
\]

**Example:**

sample a human being uniformly at random

\( Y \): height of the chosen human being

\( X \): country of origin of the chosen human being

\( \mathbb{E}(Y) = ? \)

\( \mathbb{E}(Y \mid X = "China") = ? \)
Conditional Expectation

The *conditional expectation* of a random variable $Y$ with respect to an event $\mathcal{E}$ is:

$$
\mathbb{E}(Y \mid \mathcal{E}) = \sum_y y \cdot \mathbb{P}(Y = y \mid \mathcal{E})
$$

In particular, if the event $\mathcal{E}$ is $X = x$, then:

$$
\mathbb{E}(Y \mid X = x) = \sum_y y \cdot \mathbb{P}(Y = y \mid X = x)
$$

**Example:**

sample a human being uniformly at random

$Y$: height of the chosen human being

$X$: country of origin of the chosen human being

$\mathbb{E}(Y) =$?

$\mathbb{E}(Y \mid X = \text{"China"}) =$?

$\mathbb{E}(Y \mid X = \text{"U.S."}) =$?
Conditional Expectation

Example:
sample a human being uniformly at random
$Y$: height of the chosen human being
$X$: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = "China") = ?$

$\mathbb{E}(Y \mid X = "U.S." ) = ?$
Conditional Expectation

**Example:**
sample a human being uniformly at random

$Y$: height of the chosen human being

$X$: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = \ ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = \ ?$

$\mathbb{E}(Y \mid X = x)$
Conditional Expectation

**Example:**

Sample a human being uniformly at random

- \( Y \): height of the chosen human being
- \( X \): country of origin of the chosen human being

\[
\mathbb{E}(Y \mid X = "China") = ?
\]
\[
\mathbb{E}(Y \mid X = "U.S.") = ?
\]

\[
f(x) = \mathbb{E}(Y \mid X = x)
\]
Conditional Expectation

Example:

Sample a human being uniformly at random

$Y$: height of the chosen human being

$X$: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = "China") = ?$

$\mathbb{E}(Y \mid X = "U.S.") = ?$

\[
f(X) = \mathbb{E}(Y \mid X)\]
Conditional Expectation

Example:
sample a human being uniformly at random
$Y$: height of the chosen human being
$X$: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = "China") = ?$
$\mathbb{E}(Y \mid X = "U.S.") = ?$

$$f(X) = \mathbb{E}(Y \mid X)$$
a random variable
Conditional Expectation

**Example:**
sample a human being uniformly at random

$Y$: height of the chosen human being

$X$: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = "China") =$?

$\mathbb{E}(Y \mid X = "U.S." ) =$?

$$f(X) = \mathbb{E}(Y \mid X)$$

a random variable

**Example:**
throw a fair six-sided dice for $n$ times

$X_i$: # of times $i$ appears in $n$ throws
Conditional Expectation

**Example:**
sample a human being uniformly at random
$Y$: height of the chosen human being
$X$: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = "China") = ?$

$\mathbb{E}(Y \mid X = "U.S.") = ?$

$f(X) = \mathbb{E}(Y \mid X)$
a random variable

**Example:**
throw a fair six-sided dice for $n$ times
$X_i$: # of times $i$ appears in $n$ throws

$\mathbb{E}(X_1 \mid X_2 = a) =$
Conditional Expectation

**Example:**
sample a human being uniformly at random
$Y$: height of the chosen human being
$X$: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = "China") = ?$

$\mathbb{E}(Y \mid X = "U.S." ) = ?$

$f(X) = \mathbb{E}(Y \mid X)$
a random variable

**Example:**
throw a fair six-sided dice for $n$ times
$X_i$: # of times $i$ appears in $n$ throws

$\mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5$
Conditional Expectation

**Example:**
sample a human being uniformly at random

\[ \hat{Y}: \text{height of the chosen human being} \]
\[ \hat{X}: \text{country of origin of the chosen human being} \]

\[ \mathbb{E}(Y \mid X = \text{"China"}) = ? \]
\[ \mathbb{E}(Y \mid X = \text{"U.S."}) = ? \]

\[ f(X) = \mathbb{E}(Y \mid X) \]

**Example:**
throw a fair six-sided dice for \( n \) times

\[ X_i: \text{# of times } i \text{ appears in } n \text{ throws} \]

\[ \mathbb{E}(X_1 \mid X_2 = a) = \frac{n - a}{5} \]

\[ \mathbb{E}(X_1 \mid X_2) = \]
Conditional Expectation

Example:
sample a human being uniformly at random
\( Y \): height of the chosen human being
\( X \): country of origin of the chosen human being
\[ \mathbb{E}(Y \mid X = "China") = ? \]
\[ \mathbb{E}(Y \mid X = "U.S.") = ? \]

\[ f(X) = \mathbb{E}(Y \mid X) \quad \text{a random variable} \]

Example:
throw a fair six-sided dice for \( n \) times
\( X_i \): \# of times \( i \) appears in \( n \) throws
\[ \mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5 \]
\[ \mathbb{E}(X_1 \mid X_2) = (n - X_2)/5 \]
Conditional Expectation

Example:
sample a human being uniformly at random $Y$: height of the chosen human being
$X$: country of origin of the chosen human being

$$f(X) = \mathbb{E}(Y \mid X)$$

Example:
throw a fair six-sided dice for $n$ times
$X_i$: # of times $i$ appears in $n$ throws

$$\mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5$$
$$\mathbb{E}(X_1 \mid X_2) = (n - X_2)/5$$
$$\mathbb{E}(X_1 \mid X_2 = a, X_3 = b) =$$
Conditional Expectation

Example:
sample a human being uniformly at random
$Y$: height of the chosen human being
$X$: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$
$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$f(X) = \mathbb{E}(Y \mid X)$

Example:
throw a fair six-sided dice for $n$ times
$X_i$: # of times $i$ appears in $n$ throws

$\mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5$
$\mathbb{E}(X_1 \mid X_2) = (n - X_2)/5$

$\mathbb{E}(X_1 \mid X_2 = a, X_3 = b) = (n - a - b)/4$
Conditional Expectation

**Example:**
sample a human being uniformly at random 
\( Y: \) height of the chosen human being 
\( X: \) country of origin of the chosen human being 
\[ \mathbb{E}(Y \mid X = \text{"China"}) =? \]
\[ \mathbb{E}(Y \mid X = \text{"U.S."}) =? \]

\[ f(X) = \mathbb{E}(Y \mid X) \]

**Example:**
throw a fair six-sided dice for \( n \) times 
\( X_i: \) # of times \( i \) appears in \( n \) throws 
\[ \mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5 \]
\[ \mathbb{E}(X_1 \mid X_2) = (n - X_2)/5 \]
\[ \mathbb{E}(X_1 \mid X_2 = a, X_3 = b) = (n - a - b)/4 \]
\[ \mathbb{E}(X_1 \mid X_2, X_3) = \]
Conditional Expectation

**Example:**
sample a human being uniformly at random
\( Y \): height of the chosen human being
\( X \): country of origin of the chosen human being

\[ \mathbb{E}(Y \mid X = "China") = ? \]
\[ \mathbb{E}(Y \mid X = "U.S.") = ? \]

\[ f(X) = \mathbb{E}(Y \mid X) \]  
\[ \text{a random variable} \]

**Example:**
throw a fair six-sided dice for \( n \) times
\( X_i \): # of times \( i \) appears in \( n \) throws

\[ \mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5 \]
\[ \mathbb{E}(X_1 \mid X_2) = (n - X_2)/5 \]
\[ \mathbb{E}(X_1 \mid X_2 = a, X_3 = b) = (n - a - b)/4 \]
\[ \mathbb{E}(X_1 \mid X_2, X_3) = (n - X_2 - X_3)/4 \]
Fundamental Facts about Conditional Expectation

*Example:* $Y$: height of the chosen human being  
$X$: country of origin of the chosen human being  
$Z$: gender of the chosen human being
Fundamental Facts about Conditional Expectation

*Example*: \( Y \): height of the chosen human being  
\( X \): country of origin of the chosen human being  
\( Z \): gender of the chosen human being

\[
\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y | X))
\]
Fundamental Facts about Conditional Expectation

*Example:* $Y$: height of the chosen human being  
$X$: country of origin of the chosen human being  
$Z$: gender of the chosen human being

\[
E(Y) = E\left(E(Y \mid X)\right)
\]

- average height of all human beings
- $= \text{weighted average of the country-by-country average heights}$
Fundamental Facts about Conditional Expectation

Example: \( Y \): height of the chosen human being
\( X \): country of origin of the chosen human being
\( Z \): gender of the chosen human being

\[
E(Y) = E\left( E(Y \mid X) \right)
\]

average height of all human beings
= weighted average of the country-by-country average heights

\[
E(Y \mid Z) = E\left( E(Y \mid X, Z) \mid Z \right)
\]
Fundamental Facts about Conditional Expectation

Example: \( Y \): height of the chosen human being  
\( X \): country of origin of the chosen human being  
\( Z \): gender of the chosen human being

\[
E(Y) = E(E(Y | X))
\]

average height of all human beings  
= weighted average of the country-by-country average heights

\[
E(Y | Z) = E(E(Y | X, Z) | Z)
\]

average height of all male/female human beings  
= weighted average of the country-by-country average male/female heights
Fundamental Facts about Conditional Expectation

\[ E(Y) = E\left( E(Y \mid X) \right) \]

average height of all human beings
\[ = \text{weighted average of the country-by-country average heights} \]

\[ E(Y \mid Z) = E\left( E(Y \mid X, Z) \mid Z \right) \]

average height of all male/female human beings
\[ = \text{weighted average of the country-by-country average male/female heights} \]

\[ E\left( E(f(X)g(X,Y) \mid X) \right) = E\left( f(X)E(g(X,Y) \mid X) \right) \]
Fundamental Facts about Conditional Expectation

*Example:* $Y$: height of the chosen human being  
$X$: country of origin of the chosen human being  
$Z$: gender of the chosen human being

\[
\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y \mid X))
\]

average height of all human beings  
= weighted average of the country-by-country average heights

\[
\mathbb{E}(Y \mid Z) = \mathbb{E}(\mathbb{E}(Y \mid X, Z) \mid Z)
\]

average height of all male/female human beings  
= weighted average of the country-by-country average male/female heights

\[
\mathbb{E}(\mathbb{E}(f(X)g(X,Y) \mid X)) = \mathbb{E}(f(X)\mathbb{E}(g(X,Y) \mid X))
\]

once $X$ is fixed to some $x$,  
\[
\mathbb{E}(f(X)g(X,Y) \mid X = x) = f(x)\mathbb{E}(g(X,Y) \mid X = x)
\]
Martingales

originally refers to a betting strategy:

“double your bet after every loss”
Martingales

originally refers to a betting strategy:

“double your bet after every loss”

when you get a win after $n$ losses: $2^n - \sum_{i=0}^{n-1} 2^i = 1$
Martingales

originally refers to a betting strategy:

“double your bet after every loss”

when you get a win after $n$ losses: $2^n - \sum_{i=0}^{n-1} 2^i = 1$

consider a fair game, with any betting strategy
Martingales

originally refers to a betting strategy:

“double your bet after every loss”

when you get a win after \( n \) losses: \( 2^n - \sum_{i=0}^{n-1} 2^i = 1 \)

consider a fair game, with any betting strategy

let \( X_i \) be our wealth after \( i \) rounds
Martingales

originally refers to a betting strategy:

“double your bet after every loss”

when you get a win after $n$ losses: $2^n - \sum_{i=0}^{n-1} 2^i = 1$

consider a fair game, with any betting strategy

let $X_i$ be our wealth after $i$ rounds

$$\mathbb{E}(X_{i+1} \mid X_0, X_1, \ldots, X_i) =$$
Martingales

originally refers to a betting strategy:

“double your bet after every loss”

when you get a win after $n$ losses: $2^n - \sum_{i=0}^{n-1} 2^i = 1$

consider a fair game, with any betting strategy

let $X_i$ be our wealth after $i$ rounds

$\mathbb{E}(X_{i+1} | X_0, X_1, \ldots, X_i) = X_i$
Martingales

originally refers to a betting strategy:

“double your bet after every loss”

when you get a win after \( n \) losses: \( 2^n - \sum_{i=0}^{n-1} 2^i = 1 \)

consider a fair game, with any betting strategy

let \( X_i \) be our wealth after \( i \) rounds

\[
\mathbb{E}(X_{i+1} \mid X_0, X_1, \ldots, X_i) = X_i
\]

since the game is fair, conditioned on past history, we expect no change to current value after one round
Martingales

A sequence of random variables $X_0, X_1, \cdots$ is a martingale if for all $i \geq 1$,

$$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$$
Example: Random Walk

a dot starting from the origin in each step, move equiprobably to one of four neighbor
Example: Random Walk

- A dot starting from the origin.
- In each step, move equiprobably to one of four neighbors.
- After $i$ steps, use $X_i$ to denote the number of hops to the origin (Manhattan distance).
Random Walk

Start from the origin and in each step, move equiprobably to one of four neighbors.

After \( i \) steps, use \( X_i \) to denote the number of hops to the origin (Manhattan distance).
Example: Random Walk

a dot starting from the origin
in each step, move equiprobably
to one of four neighbors

after $i$ steps, use $X_i$ to denote
# of hops to origin (Manhattan distance)

$$E(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$$
Example: Random Walk

A dot starting from the origin in each step, move equiprobably to one of four neighbors after \(i\) steps, use \(X_i\) to denote # of hops to origin (Manhattan distance)

\[
\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}
\]

How far the dot is away from the origin after \(n\) steps?
Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$
**Azuma’s Inequality**

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

$X_0, X_1, \cdots$ are not necessarily independent
Azuma’s Inequality in Action

After $i$ steps, use $X_i$ to denote 
# of hops to origin (Manhattan distance)

How large is $X_n$?
Azuma’s Inequality in Action

After $i$ steps, use $X_i$ to denote 
# of hops to origin (Manhattan distance)

How large is $X_n$?

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

\[ |X_k - X_{k-1}| \leq c_k \]

Then,

\[
\Pr(|X_n - X_0| \geq t) \leq 2 \exp \left( -2 \frac{t^2}{\sum_{k=1}^{n} c_k^2} \right)
\]
Azuma’s Inequality in Action

After $i$ steps, use $X_i$ to denote
# of hops to origin (Manhattan distance)

How large is $X_n$?

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,
$$|X_k - X_{k-1}| \leq c_k$$

Then,
$$\mathbb{P}(\left|X_n - X_0\right| \geq t) \leq 2 \exp\left(-\frac{t^2}{2 \sum_{k=1}^{n} c_k^2}\right)$$

We know $X_0 = 0$, and $|X_k - X_{k-1}| \leq 1$
Azuma’s Inequality in Action

After \( i \) steps, use \( X_i \) to denote 
# of hops to origin (Manhattan distance)

How large is \( X_n \)?

Let \( X_0, X_1, \cdots \) be a martingale such that for all \( k \geq 1 \),
\[
|X_k - X_{k-1}| \leq c_k
\]

Then,
\[
\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^n c_k^2} \right)
\]

We know \( X_0 = 0 \), and \( |X_k - X_{k-1}| \leq 1 \)
\[
\mathbb{P}(|X_n| \geq c\sqrt{n}) \leq 2e^{-c^2/2}
\]
Azuma’s Inequality in Action

After $i$ steps, use $X_i$ to denote the number of hops to origin (Manhattan distance).

How large is $X_n$?

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

We know $X_0 = 0$, and $|X_k - X_{k-1}| \leq 1$

$$\mathbb{P}(|X_n| \geq c\sqrt{n}) \leq 2e^{-c^2/2}$$

Within $O(\sqrt{n \log n})$ w.h.p.
Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$
Azuma’s Inequality

Let $X_0, X_1, \ldots$ be a martingale such that for all $k \geq 1$, 

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\Pr(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

For a sequence of r.v., if in each step:
* on average make no change to current value (martingale)
* no big jump (bounded difference)
Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

For a sequence of r.v., if in each step:
* on average make no change to current value (martingale)
* no big jump (bounded difference)

Then final value does not deviate far from the initial.
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1,$

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,
$$|X_k - X_{k-1}| \leq c_k$$

Then,
$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left(- \frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

Use similar strategy as in proving Chernoff bound:
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1,$

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

Use similar strategy as in proving Chernoff bound:

(a) Apply generalized Markov’s inequality to MGF
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

Use similar strategy as in proving Chernoff bound:
(a) Apply generalized Markov’s inequality to MGF
(b)* Bound the value of MGF (use Hoeffding’s lemma)
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

Use similar strategy as in proving Chernoff bound:
(a) Apply generalized Markov’s inequality to MGF
(b)* Bound the value of MGF (use Hoeffding’s lemma)
(c) Optimize the value of MGF
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,\[|X_k - X_{k-1}| \leq c_k\]

Then,

$$\mathbb{P}(X_n - X_0 \geq t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(X_n - X_0 \geq t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$, 
\[ |X_k - X_{k-1}| \leq c_k \]

Then,

\[
P(X_n - X_0 \geq t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)
\]

\[
P(X_n - X_0 \leq -t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)
\]

W.l.o.g., assume $X_0 = 0$. (Otherwise, do the proof w.r.t. sequence $X'_i = X_i - X_0$.)
Proving Azuma’s Inequality

Let $X_0, X_1, \ldots$ be a martingale such that for all $k \geq 1$,
$$|X_k - X_{k-1}| \leq c_k$$

Then,
$$\mathbb{P}(X_n - X_0 \geq t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$
$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

W.l.o.g., assume $X_0 = 0$. (Otherwise, do the proof w.r.t. sequence $X'_i = X_i - X_0$.)

Define $Y_i = X_i - X_{i-1}$, it is easy to see $\mathbb{E}(Y_i | X_{i-1}) = 0$. 
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$ 

$\mathbb{P}(X_n - X_0 \geq t) \leq ?$
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$.

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i|X_{i-1}) = 0$.

$\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$. 

$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$. 

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq \mathbb{E}(e^{\lambda X_n}) / e^{\lambda t}$ for $\lambda > 0$
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq \mathbb{E}(e^{\lambda X_n}) e^{\lambda t}$ for $\lambda > 0$
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$

$\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y | X))$

$= \mathbb{E}
\left( e^{\lambda(Y_n + X_{n-1})} \right) = \mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}}) = \mathbb{E}\left( \mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} | X_{n-1}) \right)$
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$ \hspace{1cm} $\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$

$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$

$= \mathbb{E} \left( e^{\lambda(Y_n+X_{n-1})} \right) = \mathbb{E} \left( e^{\lambda Y_n} e^{\lambda X_{n-1}} \right) = \mathbb{E} \left( \mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} | X_{n-1}) \right)$

$= \mathbb{E} \left( e^{\lambda X_{n-1}} \mathbb{E}(e^{\lambda Y_n} | X_{n-1}) \right) \mathbb{E}(\mathbb{E}(f(X)g(X,Y) | X)) = \mathbb{E}(f(X)\mathbb{E}(g(X,Y) | X))$
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$  \quad $\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$  

for $\lambda > 0$

$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$

\[
\mathbb{E} \left( e^{\lambda(Y_n + X_{n-1})} \right) = \mathbb{E} \left( e^{\lambda Y_n} e^{\lambda X_{n-1}} \right) = \mathbb{E} \left( \mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} | X_{n-1}) \right) \\
= \mathbb{E} \left( e^{\lambda X_{n-1}} \mathbb{E}(e^{\lambda Y_n} | X_{n-1}) \right)
\]
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$

$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$

$= \mathbb{E} \left( e^{\lambda (Y_n + X_{n-1})} \right) = \mathbb{E} \left( e^{\lambda Y_n} e^{\lambda X_{n-1}} \right) = \mathbb{E} \left( \mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} | X_{n-1}) \right)$

$= \mathbb{E} \left( e^{\lambda X_{n-1}} \mathbb{E}(e^{\lambda Y_n} | X_{n-1}) \right)$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$\mathbb{E}(e^{\lambda Z}) \leq \exp \left( \frac{\lambda^2 (b-a)^2}{8} \right)$
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$

$$
\mathbb{E} \left( e^{\lambda (Y_n + X_{n-1})} \right) = \mathbb{E} \left( e^{\lambda Y_n} e^{\lambda X_{n-1}} \right) = \mathbb{E} \left( \mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} | X_{n-1}) \right)
$$

$$
= \mathbb{E} \left( e^{\lambda X_{n-1}} \mathbb{E}(e^{\lambda Y_n} | X_{n-1}) \right)
$$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$$
\mathbb{E}(e^{\lambda Z}) \leq \exp \left( \frac{\lambda^2 (b-a)^2}{8} \right)
$$

let $Z_n = (Y_n | X_{n-1})$
\( X_0, X_1, \ldots \) is a martingale, with \( X_0 = 0 \), and \( |X_k - X_{k-1}| \leq c_k \)

\( Y_i = X_i - X_{i-1} \), thus \( \mathbb{E}(Y_i|X_{i-1}) = 0 \)

\( \mathbb{P}(X_n - X_0 \geq t) \leq \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}} \) for \( \lambda > 0 \)

\[ = \mathbb{E} \left( e^{\lambda(Y_n + X_{n-1})} \right) = \mathbb{E} \left( e^{\lambda Y_n} e^{\lambda X_{n-1}} \right) = \mathbb{E} \left( \mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} | X_{n-1}) \right) \]

\[ = \mathbb{E} \left( e^{\lambda X_{n-1}} \mathbb{E}(e^{\lambda Y_n} | X_{n-1}) \right) \]

For any random variable \( Z \in [a, b] \) with \( \mathbb{E}(Z) = 0 \),

\[ \mathbb{E}(e^{\lambda Z}) \leq \exp \left( \frac{\lambda^2 (b-a)^2}{8} \right) \]

let \( Z_n = (Y_n | X_{n-1}) \)

\[ \mathbb{E}(Z_n) = \mathbb{E}(Y_n | X_{n-1}) = 0 \]
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$ 

$\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$

\[
\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}
\]

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$\mathbb{E}(e^{\lambda Z}) \leq \exp \left( \frac{\lambda^2 (b-a)^2}{8} \right)$

let $Z_n = (Y_n | X_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n | X_{n-1}) = 0$

$Z_n \in [-c_n, c_n]$
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i \mid X_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$

\[ = \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid X_{n-1})\right) \]

\[ = \mathbb{E}\left(e^{\lambda X_{n-1}} \mathbb{E}(e^{\lambda Y_n} \mid X_{n-1})\right) \]

\[ \leq \exp\left(\frac{\lambda^2 4c_n^2}{8}\right) = \exp\left(\frac{\lambda c_n^2}{2}\right) \]

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$\mathbb{E}(e^{\lambda Z}) \leq \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$

let $Z_n = (Y_n \mid X_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n \mid X_{n-1}) = 0$

$Z_n \in [-c_n, c_n]$
\( X_0, X_1, \ldots \) is a martingale, with \( X_0 = 0 \), and \( |X_k - X_{k-1}| \leq c_k \)

\[ Y_i = X_i - X_{i-1}, \text{ thus } \mathbb{E}(Y_i | X_{i-1}) = 0 \]

\[ \mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}} \text{ for } \lambda > 0 \]

\[
\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}})}{e^{\lambda t}} \leq \mathbb{E}(e^{\lambda Y_n} | X_{n-1}) \leq \mathbb{E}(e^{\lambda X_{n-1}} e^{\lambda^2 c_n^2/2}) = e^{\lambda^2 c_n^2/2} \mathbb{E}(e^{\lambda X_{n-1}})
\]

\[ \leq \exp\left(\frac{\lambda^2 c_n^2}{8}\right) = \exp\left(\frac{\lambda c_n^2}{2}\right) \]

For any random variable \( Z \in [a, b] \) with \( \mathbb{E}(Z) = 0 \),

\[ \mathbb{E}(e^{\lambda Z}) \leq \exp\left(\frac{\lambda^2(b-a)^2}{8}\right) \]

let \( Z_n = (Y_n | X_{n-1}) \)

\[ \mathbb{E}(Z_n) = \mathbb{E}(Y_n | X_{n-1}) = 0 \]

\( Z_n \in [-c_n, c_n] \)
$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | X_{i-1}) = 0$  

$\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$ for $\lambda > 0$

$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$

$= \mathbb{E}\left(e^{\lambda (Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} | X_{n-1})\right)$

$= \mathbb{E}\left(e^{\lambda X_{n-1}} \mathbb{E}(e^{\lambda Y_n} | X_{n-1})\right) \leq \mathbb{E}\left(e^{\lambda X_{n-1}} e^{\lambda^2 c_n^2/2}\right) = e^{\lambda^2 c_n^2/2} \mathbb{E}\left(e^{\lambda X_{n-1}}\right)$

$\leq e^{\lambda^2 c_n^2/2} e^{\lambda^2 c_{n-1}^2/2} \mathbb{E}(e^{\lambda X_{n-2}}) \leq \cdots \leq \left(\prod_{k=1}^{n} e^{\lambda^2 c_k^2/2}\right) \mathbb{E}(e^{\lambda X_0}) = e^{(\lambda^2/2) \sum_{k=1}^{n} c_k^2}$

$\leq \exp\left(\frac{\lambda^2 4c_n^2}{8}\right) = \exp\left(\frac{\lambda c_n^2}{2}\right)$

Let $Z_n = (Y_n | X_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n | X_{n-1}) = 0$

$Z_n \in [-c_n, c_n]$
$X_0, X_1, \ldots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i \mid X_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$

$\leq \frac{e^{(\lambda^2/2)\sum_{k=1}^{n} c_k^2}}{e^{\lambda t}} = \exp \left( -\frac{t^2}{2\sum_{k=1}^{n} c_k^2} \right)$

minimized when $\lambda = \frac{t}{\sum_{k=1}^{n} c_k^2}$

$\mathbb{E} \left( e^{\lambda(Y_n + X_{n-1})} \right) = \mathbb{E} \left( e^{\lambda Y_n} e^{\lambda X_{n-1}} \right) = \mathbb{E} \left( \mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid X_{n-1}) \right)$

$= \mathbb{E} \left( e^{\lambda X_{n-1}} \mathbb{E}(e^{\lambda Y_n} \mid X_{n-1}) \right) \leq \mathbb{E} \left( e^{\lambda X_{n-1}} e^{\lambda^2 c_n^2/2} \right) = e^{\lambda^2 c_n^2/2} \mathbb{E}(e^{\lambda X_{n-1}})$

$\leq e^{\lambda^2 c_n^2/2} e^{\lambda^2 c_{n-1}^2/2} \mathbb{E}(e^{\lambda X_{n-2}}) \leq \ldots \leq \left( \prod_{k=1}^{n} e^{\lambda^2 c_k^2/2} \right) \mathbb{E}(e^{\lambda X_0}) = e^{(\lambda^2/2)\sum_{k=1}^{n} c_k^2}$

$\leq \exp \left( \frac{\lambda^2 c_n^2}{8} \right) = \exp \left( \frac{\lambda c_n^2}{2} \right)$

let $Z_n = (Y_n \mid X_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n \mid X_{n-1}) = 0$

$Z_n \in [-c_n, c_n]$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$, $\mathbb{E}(e^{\lambda Z}) \leq \exp \left( \frac{\lambda^2(b-a)^2}{8} \right)$
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(X_n - X_0 \geq t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$
Proving Azuma’s Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(X_n - X_0 \geq t) \leq \exp\left(-\frac{t^2}{2 \sum_{k=1}^{n} c_k^2}\right)$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp\left(-\frac{t^2}{2 \sum_{k=1}^{n} c_k^2}\right)$$
Proving Azuma’s Inequality

Let \( X_0, X_1, \cdots \) be a martingale such that for all \( k \geq 1 \),
\[
|X_k - X_{k-1}| \leq c_k
\]
Then,
\[
\mathbb{P}(X_n - X_0 \geq t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)
\]
\[
\mathbb{P}(X_n - X_0 \leq -t) \leq \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)
\]

let \( X_i' = -X_i \)
\[
\mathbb{P}(X_n - X_0 \leq -t) = \mathbb{P}((-X_i') - (-X_0') \leq -t) = \mathbb{P}(X_i' - X_0' \geq t) \leq \cdots
\]
Generalized Martingales

A sequence of random variables $Y_0, Y_1, \cdots$ is a martingale with respect to the sequence $X_0, X_1, \cdots$ if for all $i \geq 0$,

- $Y_i$ is a function of $X_0, X_1, \cdots, X_i$
- $\mathbb{E}(Y_{i+1} \mid X_0, X_1, \cdots, X_i) = Y_i$
Generalized Martingales

A sequence of random variables $Y_0, Y_1, \cdots$ is a martingale with respect to the sequence $X_0, X_1, \cdots$ if for all $i \geq 0$,

- $Y_i$ is a function of $X_0, X_1, \cdots, X_i$
- $\mathbb{E}(Y_{i+1} \mid X_0, X_1, \cdots, X_i) = Y_i$

betting on a fair game

$X_i$: gain/loss of the $i^{th}$ bet

$Y_i$: wealth after the $i^{th}$ bet
Generalized Martingales

A sequence of random variables $Y_0, Y_1, \cdots$ is a martingale with respect to the sequence $X_0, X_1, \cdots$ if for all $i \geq 0$,

- $Y_i$ is a function of $X_0, X_1, \cdots, X_i$
- $\mathbb{E}(Y_{i+1} \mid X_0, X_1, \cdots, X_i) = Y_i$

betting on a fair game

$X_i$: gain/loss of the $i^{th}$ bet

$Y_i$: wealth after the $i^{th}$ bet $\leftarrow$ martingale (since game is fair)
Generalized Azuma’s Inequality

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$ such that for all $k \geq 1$,

$$|Y_k - Y_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|Y_n - Y_0| \geq t) \leq 2 \exp \left( - \frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$
Azuma’s Inequality
martingale $X_0, X_1, \ldots$
with $|X_k - X_{k-1}| \leq c_k$,
then $\mathbb{P}(|X_n - X_0| \geq t) \leq \cdots$

generalization

Generalized Azuma’s Inequality
martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
with $|Y_k - Y_{k-1}| \leq c_k$,
then $\mathbb{P}(|Y_n - Y_0| \geq t) \leq \cdots$

generalization

martingale $X_0, X_1, \ldots, X_n$
$\mathbb{E}(X_i \mid X_0, X_1, \ldots, X_{i-1}) = X_{i-1}$

martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
$Y_i = f(X_0, X_1, \ldots, X_i)$
$\mathbb{E}(Y_i \mid X_0, X_1, \ldots, X_{i-1}) = Y_{i-1}$
The Doob sequence of a function $f$ with respect to a sequence of random variables $X_1, X_2, \cdots, X_n$ is

$$Y_i = \mathbb{E} \left( f(X_1, \cdots, X_n) \mid X_1, \cdots, X_i \right)$$

In particular,

$$Y_0 = \mathbb{E} \left( f(X_1, \cdots, X_n) \right), \text{ and } Y_n = f(X_1, \cdots, X_n)$$
Doob Sequence

\[ f(\text{\$}, \text{\$}, \text{\$}, \text{\$}) \]
Doob Sequence

\[ f(\, \, , \, , \, , \, , \, ) \]

average over

no information

\[ \mathbb{E}(f) \]
Doob Sequence

\[ f(1, \ldots, \ldots) \]

randomized by

\[ \mathbb{E}(f) \longrightarrow \mathbb{E}(f|X_1) \]

no information
Doob Sequence

\[ f(1, 0), \quad \mathbb{E}(f|X_1), \quad \mathbb{E}(f|X_2) \]

randomized by

no information

\[ \mathbb{E}(f) \rightarrow \mathbb{E}(f|X_1) \rightarrow \mathbb{E}(f|X_2) \]
Doob Sequence

\[ f(1, 0, 0, \cdots) \]

randomized by

average over

no information

\[
\mathbb{E}(f) \rightarrow \mathbb{E}(f|X_1) \rightarrow \mathbb{E}(f|X_2) \rightarrow \mathbb{E}(f|X_3)
\]
Doob Sequence

\[ f(1, 0, 0, 1) \]

randomized by

\[
\begin{align*}
\mathbb{E}(f) &\rightarrow \mathbb{E}(f|X_1) \\
&\rightarrow \mathbb{E}(f|X_2) \\
&\rightarrow \mathbb{E}(f|X_3) \\
&\rightarrow \mathbb{E}(f|X_4) = f(X_4)
\end{align*}
\]
The Doob sequence of a function $f$ with respect to a sequence of random variables $X_1, X_2, \ldots, X_n$ is

$$Y_i = \mathbb{E}(f(X_1, \ldots, X_n) \mid X_1, \ldots, X_i)$$

The Doob sequence of a function $f$ is a martingale. That is,

$$\mathbb{E}(Y_i \mid X_1, \ldots, X_{i-1}) = Y_{i-1}$$
The Doob sequence of a function $f$ with respect to a sequence of random variables $X_1, X_2, \cdots, X_n$ is

$$Y_i = \mathbb{E}(f(X_1, \cdots, X_n) \mid X_1, \cdots, X_i)$$

The Doob sequence of a function $f$ is a martingale. That is,

$$\mathbb{E}(Y_i \mid X_1, \cdots, X_{i-1}) = Y_{i-1}$$

$$\mathbb{E}(Y_i \mid X_{i-1}) = \mathbb{E}(\mathbb{E}(f(X_n) \mid X_i) \mid X_{i-1})$$
Doob Martingale

The Doob sequence of a function $f$ with respect to a sequence of random variables $X_1, X_2, \cdots, X_n$ is

$$Y_i = \mathbb{E}(f(X_1, \cdots, X_n) \mid X_1, \cdots, X_i)$$

The Doob sequence of a function $f$ is a martingale. That is,

$$\mathbb{E}(Y_i \mid X_1, \cdots, X_{i-1}) = Y_{i-1}$$

$$\mathbb{E}(Y_i \mid X_{i-1}) = \mathbb{E}(\mathbb{E}(f(X_n) \mid X_i) \mid X_{i-1})$$

$$= \mathbb{E}(f(X_n) \mid X_{i-1})$$

$$\mathbb{E}(Y \mid Z) = \mathbb{E}(\mathbb{E}(Y \mid X, Z) \mid Z)$$
The Doob sequence of a function $f$ with respect to a sequence of random variables $X_1, X_2, \ldots, X_n$ is

$$Y_i = \mathbb{E}(f(X_1, \ldots, X_n) \mid X_1, \ldots, X_i)$$

The Doob sequence of a function $f$ is a martingale. That is,

$$\mathbb{E}(Y_i \mid X_1, \ldots, X_{i-1}) = Y_{i-1}$$

$$\mathbb{E}(Y_i \mid X_{i-1}) = \mathbb{E}(\mathbb{E}(f(X_n) \mid X_i) \mid X_{i-1})$$

$$= \mathbb{E}(f(X_n) \mid X_{i-1})$$

$$= Y_{i-1}$$
$G_{n,p}$
Graph parameter: $f(G)$

Example: components number, chromatic number, diameter
Graph parameter: $f(G)$

*Example*: components number, chromatic number, diameter

numbering all vertex pairs: $1, 2, 3, \ldots, \binom{n}{2}$
Graph parameter: $f(G)$

*Example*: components number, chromatic number, diameter

numbering all vertex pairs: $1, 2, 3, \ldots, \binom{n}{2}$

Define i.r.v. $I_j = \begin{cases} 1 & \text{edge } j \in G \\ 0 & \text{edge } j \notin G \end{cases}$
Graph parameter: $f(G)$

\textbf{Example}: components number, chromatic number, diameter

numbering all vertex pairs: $1, 2, 3, \ldots, \binom{n}{2}$

Define i.r.v. $I_j = \begin{cases} 1 & \text{edge } j \in G \\ 0 & \text{edge } j \notin G \end{cases} \quad Y_i = \mathbb{E}(f(G) \mid I_1, \ldots, I_i)$
**Graph parameter:** $f(G)$

*Example:* components number, chromatic number, diameter

Numbering all vertex pairs: $1, 2, 3, \ldots, \binom{n}{2}$

Define i.r.v. $I_j = \begin{cases} 1 & \text{edge } j \in G \\ 0 & \text{edge } j \notin G \end{cases}$

$Y_i = \mathbb{E}(f(G) | I_1, \ldots, I_i)$

$Y_0, Y_1, \ldots, Y_{\binom{n}{2}}$ is a Doob sequence, called *edge exposure martingale*

In particular, $Y_0 = \mathbb{E}(f(G))$, and $Y_{\binom{n}{2}} = f(G)$
Graph parameter: $f(G)$

*Example*: components number, chromatic number, diameter

numbering all vertices: 1, 2, 3, ..., $n$
Graph parameter: $f(G)$

*Example*: components number, chromatic number, diameter

numbering all vertices: $1, 2, 3, \ldots, n$

$X_i$: subgraph of $G$ induced by the first $i$ vertices
Graph parameter: $f(G)$

Example: components number, chromatic number, diameter

numbering all vertices: $1, 2, 3, \cdots, n$

$X_i$: subgraph of $G$ induced by the first $i$ vertices

$Y_i = \mathbb{E}(f(G) | X_1, \cdots, X_i)$
Graph parameter: $f(G)$

Example: components number, chromatic number, diameter

numbering all vertices: $1, 2, 3, \cdots, n$

$X_i$: subgraph of $G$ induced by the first $i$ vertices

$$Y_i = \mathbb{E}(f(G) | X_1, \cdots, X_i)$$

$Y_0, Y_1, \cdots, Y_n$ is a Doob sequence, called vertex exposure martingale.

In particular, $Y_0 = \mathbb{E}(f(G))$, and $Y_n = f(G)$
numbering all vertices: 1, 2, 3, ⋯, n

$X_i$: subgraph of $G$ induced by the first $i$ vertices

$Y_i = \mathbb{E}(\chi(G) | X_1, \cdots, X_i)$

$Y_0, Y_1, \cdots, Y_n$ is a Doob sequence (vertex exposure martingale)
In particular, $Y_0 = \mathbb{E}(\chi(G))$, and $Y_n = \chi(G)$
Generalized Azuma’s Inequality in Action

Concentration of Chromatic Number

chromatic number $\chi(G)$

$X_i$: subgraph of $G$ induced by the first $i$ vertices

$Y_i = \mathbb{E}(\chi(G) | X_1, \ldots, X_i)$

$Y_0, Y_1, \ldots, Y_n$ a Doob martingale: $Y_0 = \mathbb{E}(\chi(G))$, and $Y_n = \chi(G)$
Generalized Azuma’s Inequality in Action

Concentration of Chromatic Number

\( \chi(\mathcal{G}) \) chromatic number

\( X_i \): subgraph of \( G \) induced by the first \( i \) vertices

\( Y_i = \mathbb{E}(\chi(G) | X_1, \ldots, X_i) \)

\( Y_0, Y_1, \ldots, Y_n \) a Doob martingale: \( Y_0 = \mathbb{E}(\chi(G)) \), and \( Y_n = \chi(G) \)

Let \( Y_0, Y_1, \ldots \) be a martingale with respect to \( X_0, X_1, \ldots \) such that for all \( k \geq 1 \),

\[ |Y_k - Y_{k-1}| \leq c_k \]

Then,

\[ \mathbb{P}(|Y_n - Y_0| \geq t) \leq 2 \exp \left( - \frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right) \]
Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$ such that for all $k \geq 1$,

$$|Y_k - Y_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|Y_n - Y_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$

$X_i$: subgraph of $G$ induced by the first $i$ vertices

$Y_i = \mathbb{E}(\chi(G) | X_1, \cdots, X_i)$

$Y_0, Y_1, \cdots, Y_n$ a Doob martingale: $Y_0 = \mathbb{E}(\chi(G))$, and $Y_n = \chi(G)$

A new vertex can always be given a new color!
Generalized Azuma’s Inequality in Action

Concentration of Chromatic Number

\[ \chi(G) \]

\( X_i \): subgraph of \( G \) induced by the first \( i \) vertices

\[ Y_i = \mathbb{E}(\chi(G) | X_1, \ldots, X_i) \]

\( Y_0, Y_1, \ldots, Y_n \) a Doob martingale: \( Y_0 = \mathbb{E}(\chi(G)) \), and \( Y_n = \chi(G) \)

A new vertex can always be given a new color!

\[ |Y_i - Y_{i-1}| \leq 1 \]
Generalized Azuma's Inequality in Action

Concentration of Chromatic Number

chromatic number \(\chi(G)\)

\(X_i\): subgraph of \(G\) induced by the first \(i\) vertices

\(Y_i = \mathbb{E}(\chi(G) | X_1, \ldots, X_i)\)

\(Y_0, Y_1, \ldots, Y_n\) a Doob martingale: \(Y_0 = \mathbb{E}(\chi(G))\), and \(Y_n = \chi(G)\)

A new vertex can always be given a new color!

\[ |Y_i - Y_{i-1}| \leq 1 \]

\[
\mathbb{P}(|\chi(G) - \mathbb{E}(\chi(G))| \geq t\sqrt{n}) = \mathbb{P}(|Y_n - Y_0| \geq t\sqrt{n}) \leq 2e^{-t^2/2}
\]
Theorem [Shamir & Spencer (1987)]:
Let $G \sim G(n, p)$, then:

$$\mathbb{P}(|\chi(G) - \mathbb{E}(\chi(G))| \geq t\sqrt{n}) \leq 2e^{-t^2/2}$$
Azuma’s Inequality

Martingale $X_0, X_1, \ldots$

With $|X_k - X_{k-1}| \leq c_k$,
then $\mathbb{P}(|X_n - X_0| \geq t) \leq \cdots$

Generalized Azuma’s Inequality

Martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$

With $|Y_k - Y_{k-1}| \leq c_k$,
then $\mathbb{P}(|Y_n - Y_0| \geq t) \leq \cdots$

Martingale $X_0, X_1, \ldots, X_n$

$\mathbb{E}(X_i \mid X_0, X_1, \ldots, X_{i-1}) = X_{i-1}$

generalization

Martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$

$Y_i = f(X_0, X_1, \ldots, X_i)$

$\mathbb{E}(Y_i \mid X_0, X_1, \ldots, X_{i-1}) = Y_{i-1}$

generalization
Azuma’s Inequality

martingale $X_0, X_1, \ldots$
with $|X_k - X_{k-1}| \leq c_k$, then $\mathbb{P}(|X_n - X_0| \geq t) \leq \ldots$

generalization

Generalized Azuma’s Inequality

martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
with $|Y_k - Y_{k-1}| \leq c_k$, then $\mathbb{P}(|Y_n - Y_0| \geq t) \leq \ldots$

generalization

martingale $X_0, X_1, \ldots, X_n$
$\mathbb{E}(X_i \mid X_0, X_1, \ldots, X_{i-1}) = X_{i-1}$
generalization

martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
$Y_i = f(X_0, X_1, \ldots, X_i)$
$\mathbb{E}(Y_i \mid X_0, X_1, \ldots, X_{i-1}) = Y_{i-1}$
special case

Doob martingale $Y_0, Y_1, \ldots$
$Y_i = \mathbb{E}(f(X_0, X_1, \ldots, X_n) \mid X_0, X_1, \ldots, X_{i-1})$
Azuma’s Inequality
martingale $X_0, X_1, \ldots$
with $|X_k - X_{k-1}| \leq c_k$,
then $\mathbb{P}(|X_n - X_0| \geq t) \leq \ldots$

generalization

Generalized Azuma’s Inequality
martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
with $|Y_k - Y_{k-1}| \leq c_k$,
then $\mathbb{P}(|Y_n - Y_0| \geq t) \leq \ldots$

Doob martingale $Y_0, Y_1, \ldots$
$Y_i = \mathbb{E}(f(X_0, X_1, \ldots, X_n) \mid X_0, X_1, \ldots, X_{i-1})$

special case

martingale $X_0, X_1, \ldots, X_n$
$\mathbb{E}(X_i \mid X_0, X_1, \ldots, X_{i-1}) = X_{i-1}$
generalization

generalization

verteX exposure martingale

applied in random graphs
Azuma’s Inequality
martingale $X_0, X_1, \ldots$
with $|X_k - X_{k-1}| \leq c_k$,
then $\mathbb{P}(|X_n - X_0| \geq t) \leq \ldots$

generalization

Generalized Azuma’s Inequality
martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
with $|Y_k - Y_{k-1}| \leq c_k$,
then $\mathbb{P}(|Y_n - Y_0| \geq t) \leq \ldots$

generalization

Sample Application:
Tight Concentration of Chromatic number

Martingale $X_0, X_1, \ldots, X_n$
$\mathbb{E}(X_i \mid X_0, X_1, \ldots, X_{i-1}) = X_{i-1}$
generalization

Martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
$Y_i = f(X_0, X_1, \ldots, X_i)$
$\mathbb{E}(Y_i \mid X_0, X_1, \ldots, X_{i-1}) = Y_{i-1}$
special case

Doob martingale $Y_0, Y_1, \ldots$
$Y_i = \mathbb{E}(f(X_0, X_1, \ldots, X_n) \mid X_0, X_1, \ldots, X_{i-1})$
applied in random graphs

Vertex exposure martingale
Doob Martingale +
Generalized Azuma’s Inequality

• for a function of (potentially dependent) r.v.: 
  \[ f(X_1, X_2, \ldots, X_n) \]

• define corresponding Doob martingale: 
  \[ Y_i = \mathbb{E}(f(X_1, \ldots, X_n) \mid X_1, \ldots, X_i) \]

  in particular, \( Y_0 = \mathbb{E}(f(X_1, \ldots, X_n)) \) and \( Y_n = f(X_1, \ldots, X_n) \)

• as long as the differences \( |Y_i - Y_{i-1}| \) are bounded

• generalized Azuma’s inequality implies \( |Y_n - Y_0| \) is bounded

\[ f(X_1, \ldots, X_n) \text{ is tightly concentration to its expectation} \]
The Method of Averaged Bounded Differences

Let $X = (X_1, \ldots, X_n)$ and let $f$ be a function of $X_0, X_1, \ldots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$| \mathbb{E}(f(X) \mid X_1, \ldots, X_i) - \mathbb{E}(f(X) \mid X_1, \ldots, X_{i-1}) | \leq c_i$$

Then,

$$\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{i=1}^{n} c_i^2} \right)$$
The Method of Averaged Bounded Differences

Let $X = (X_1, \ldots, X_n)$ and let $f$ be a function of $X_0, X_1, \ldots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$\left| \mathbb{E}(f(X) \mid X_1, \ldots, X_i) - \mathbb{E}(f(X) \mid X_1, \ldots, X_{i-1}) \right| \leq c_i$$

Then,

$$\mathbb{P}(\left| f(X) - \mathbb{E}(f(X)) \right| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{i=1}^{n} c_i^2} \right)$$

Doob Martingale
The Method of Averaged Bounded Differences

Let $X = (X_1, \cdots, X_n)$ and let $f$ be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$|\mathbb{E}(f(X) \mid X_1, \cdots, X_i) - \mathbb{E}(f(X) \mid X_1, \cdots, X_{i-1})| \leq c_i$$

Then,

$$\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq 2 \exp\left(-\frac{t^2}{2 \sum_{i=1}^{n} c_i^2}\right)$$

Doob Martingale + Generalized Azuma’s Inequality
Let \( X = (X_1, \cdots, X_n) \) and let \( f \) be a function of \( X_0, X_1, \cdots, X_n \) satisfying that, for all \( 1 \leq i \leq n \),

\[
| \mathbb{E}(f(X) \mid X_1, \cdots, X_i) - \mathbb{E}(f(X) \mid X_1, \cdots, X_{i-1}) | \leq c_i
\]

Then,

\[
\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{i=1}^{n} c_i^2} \right)
\]
The Method of Averaged Bounded Differences

Let $X = (X_1, \cdots, X_n)$ and let $f$ be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$\left| \mathbb{E}(f(X) \mid X_1, \cdots, X_i) - \mathbb{E}(f(X) \mid X_1, \cdots, X_{i-1}) \right| \leq c_i$$

Then, May be hard to check!

$$\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{i=1}^{n} c_i^2} \right)$$
\[ | \mathbb{E}(f(X) \mid X_1, \ldots, X_i) - \mathbb{E}(f(X) \mid X_1, \ldots, X_{i-1}) | \leq c_i \]
Lipschitz Condition

\[ | \mathbb{E}(f(X) \mid X_1, \cdots, X_i) - \mathbb{E}(f(X) \mid X_1, \cdots, X_{i-1}) | \leq c_i \]

A function \( f(X_1, \cdots, X_n) \) satisfies the **Lipschitz condition** with constants \( c_i \) where \( 1 \leq i \leq n \), if

\[ | f(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, x_n) - f(x_1, \cdots, x_{i-1}, y_i, x_{i+1}, x_n) | \leq c_i \]
Average-case:
\[ | \mathbb{E}(f(X) \mid X_1, \ldots, X_i) - \mathbb{E}(f(X) \mid X_1, \ldots, X_{i-1}) | \leq c_i \]

Worst-case:
A function \( f(X_1, \cdots, X_n) \) satisfies the *Lipschitz condition* with constants \( c_i \) where \( 1 \leq i \leq n \), if
\[
| f(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, x_n) - f(x_1, \cdots, x_{i-1}, y_i, x_{i+1}, x_n) | \leq c_i
\]
Let $\mathbf{X} = (X_1, \cdots, X_n)$ be $n$ independent random variables and let $f(\mathbf{X})$ be a function satisfying the Lipschitz condition with constants $c_i$ where $1 \leq i \leq n$, then:

$$
P(|f(\mathbf{X}) - \mathbb{E}(f(\mathbf{X}))| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{i=1}^{n} c_i^2} \right)$$
Let $X = (X_1, \cdots, X_n)$ and let $f$ be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n,$

$$|\mathbb{E}(f(X) \mid X_1, \cdots, X_i) - \mathbb{E}(f(X) \mid X_1, \cdots, X_{i-1})| \leq c_i$$

Then,

$$\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq 2 \exp \left( - \frac{t^2}{2 \sum_{i=1}^{n} c_i^2} \right)$$

Let $X = (X_1, \cdots, X_n)$ be $n$ independent random variables and let $f(X)$ be a function satisfying the Lipschitz condition with constants $c_i$ where $1 \leq i \leq n,$ then:

$$\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq 2 \exp \left( - \frac{t^2}{2 \sum_{i=1}^{n} c_i^2} \right)$$
Let $\mathbf{X} = (X_1, \cdots, X_n)$ and let $f$ be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$|\mathbb{E}(f(\mathbf{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\mathbf{X}) \mid X_1, \cdots, X_{i-1})| \leq c_i$$

Then,

$$\mathbb{P}(|f(\mathbf{X}) - \mathbb{E}(f(\mathbf{X}))| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

Let $\mathbf{X} = (X_1, \cdots, X_n)$ be $n$ independent random variables and let $f(\mathbf{X})$ be a function satisfying the Lipschitz condition with constants $c_i$ where $1 \leq i \leq n$, then:

$$\mathbb{P}(|f(\mathbf{X}) - \mathbb{E}(f(\mathbf{X}))| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

\[ \text{Lipschitz condition} + \text{Independence} \quad \Rightarrow \quad |\mathbb{E}(f(\mathbf{X})|X_1, \cdots, X_i) - \mathbb{E}(f(\mathbf{X})|X_1, \cdots, X_{i-1})| \leq c_i \]
Azuma’s Inequality

Martingale $X_0, X_1, \ldots$

with $|X_k - X_{k-1}| \leq c_k$,
then $\mathbb{P}(|X_n - X_0| \geq t) \leq \ldots$

generalization

Generalized Azuma’s Inequality

Martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$

with $|Y_k - Y_{k-1}| \leq c_k$,
then $\mathbb{P}(|Y_n - Y_0| \geq t) \leq \ldots$

generalization

generalization

martingale $X_0, X_1, \ldots, X_n$

$\mathbb{E}(X_i \mid X_0, X_1, \ldots, X_{i-1}) = X_{i-1}$

generalization

martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$

$Y_i = f(X_0, X_1, \ldots, X_i)$

$\mathbb{E}(Y_i \mid X_0, X_1, \ldots, X_{i-1}) = Y_{i-1}$
special case

Doob martingale $Y_0, Y_1, \ldots$

$Y_i = \mathbb{E}(f(X_0, X_1, \ldots, X_n) \mid X_0, X_1, \ldots, X_{i-1})$

special case

Azuma’s Inequality
Azuma’s Inequality
martingale $X_0, X_1, \ldots$
with $|X_k - X_{k-1}| \leq c_k$,
then $\mathbb{P}(|X_n - X_0| \geq t) \leq \cdots$

Generalized Azuma’s Inequality
martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
with $|Y_k - Y_{k-1}| \leq c_k$,
then $\mathbb{P}(|Y_n - Y_0| \geq t) \leq \cdots$

The Method of Averaged Bounded Differences
$f(X)$ satisfying $|\mathbb{E}(f(X)|X_1, \ldots, X_i) - \mathbb{E}(f(X)|X_1, \ldots, X_{i-1})| \leq c_i$,
then $\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq \cdots$

martingale $X_0, X_1, \ldots, X_n$
$\mathbb{E}(X_i | X_0, X_1, \ldots, X_{i-1}) = X_{i-1}$
generalization

martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
$Y_i = f(X_0, X_1, \ldots, X_i)$
$\mathbb{E}(Y_i | X_0, X_1, \ldots, X_{i-1}) = Y_{i-1}$
special case

Doob martingale $Y_0, Y_1, \ldots$
$Y_i = \mathbb{E}(f(X_0, X_1, \ldots, X_n) | X_0, X_1, \ldots, X_{i-1})$
Azuma’s Inequality
martingale $X_0, X_1, \ldots$
with $|X_k - X_{k-1}| \leq c_k$,
then $\mathbb{P}(|X_n - X_0| \geq t) \leq \ldots$

Generalized Azuma’s Inequality
martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
with $|Y_k - Y_{k-1}| \leq c_k$,
then $\mathbb{P}(|Y_n - Y_0| \geq t) \leq \ldots$

Doob martingale $Y_0, Y_1, \ldots$
$Y_i = \mathbb{E}(f(X_0, X_1, \ldots, X_i) | X_0, X_1, \ldots, X_{i-1})$

The Method of Averaged Bounded Differences
$f(X)$ satisfying $|\mathbb{E}(f(X)|X_1, \ldots, X_i) - \mathbb{E}(f(X)|X_1, \ldots, X_{i-1})| \leq c_i$,
then $\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq \ldots$

The Method of Bounded Differences
$X = (X_1, \ldots, X_n)$ are independent r.v., $f(X)$ satisfying the Lipschitz condition,
then $\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq \ldots$

martingale $X_0, X_1, \ldots, X_n$
$\mathbb{E}(X_i | X_0, X_1, \ldots, X_{i-1}) = X_{i-1}$
generalization

martingale $Y_0, Y_1, \ldots$ w.r.t. $X_0, X_1, \ldots$
$Y_i = f(X_0, X_1, \ldots, X_i)$
$\mathbb{E}(Y_i | X_0, X_1, \ldots, X_{i-1}) = Y_{i-1}$
special case

generalization
The Method of Bounded Differences in Action:

Pattern Matching

- a random string of length \( n \)
- a pattern of length \( k \)
- # of matched substrings?
The Method of Bounded Differences in Action:

Pattern Matching

- a random string of length $n$
- a pattern of length $k$
- # of matched substrings?

an alphabet $\Sigma$ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$
The Method of Bounded Differences in Action:
Pattern Matching

- a random string of length $n$
- a pattern of length $k$
- # of matched substrings?

an alphabet $\Sigma$ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \ldots, X_n \in \Sigma$

let $Y$ be number of substrings $\pi$ in $\langle X_1, X_2, \ldots, X_n \rangle$
The Method of Bounded Differences in Action: Pattern Matching

- a random string of length \( n \)
- a pattern of length \( k \)
- \# of matched substrings?

an alphabet \( \Sigma \) with \( |\Sigma| = m \), a fixed pattern \( \pi \in \Sigma^k \)

independently and uniformly generate: \( X_1, X_2, \ldots, X_n \in \Sigma \)

let \( Y \) be number of substrings \( \pi \) in \( \langle X_1, X_2, \ldots, X_n \rangle \)

\[
\mathbb{E}(Y) = (n - k + 1) \left( \frac{1}{m} \right)^k
\]
The Method of Bounded Differences in Action: Pattern Matching

• a random string of length $n$
• a pattern of length $k$
• # of matched substrings?

an alphabet $\Sigma$ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let $Y$ be number of substrings $\pi$ in $\langle X_1, X_2, \cdots, X_n \rangle$

\[ \mathbb{E}(Y) = (n - k + 1) \left( \frac{1}{m} \right)^k \]

Deviation?
The Method of Bounded Differences in Action: Pattern Matching

- a random string of length $n$
- a pattern of length $k$
- # of matched substrings?

an alphabet $\Sigma$ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let $Y$ be number of substrings $\pi$ in $\langle X_1, X_2, \cdots, X_n \rangle$
The Method of Bounded Differences in Action:

Pattern Matching

- a random string of length $n$
- a pattern of length $k$
- # of matched substrings?

an alphabet $\Sigma$ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \ldots, X_n \in \Sigma$

let $Y$ be number of substrings $\pi$ in $\langle X_1, X_2, \ldots, X_n \rangle$

$Y = f(X_1, X_2, \ldots, X_n)$
The Method of Bounded Differences in Action:
Pattern Matching

• a random string of length \( n \)
• a pattern of length \( k \)
• # of matched substrings?

an alphabet \( \Sigma \) with \( |\Sigma| = m \), a fixed pattern \( \pi \in \Sigma^k \)

independently and uniformly generate: \( X_1, X_2, \ldots, X_n \in \Sigma \)

let \( Y \) be number of substrings \( \pi \) in \( \langle X_1, X_2, \ldots, X_n \rangle \)

\[ Y = f(X_1, X_2, \ldots, X_n) \]
The Method of Bounded Differences in Action: Pattern Matching

- a random string of length \( n \)
- a pattern of length \( k \)
- # of matched substrings?

an alphabet \( \Sigma \) with \( |\Sigma| = m \), a fixed pattern \( \pi \in \Sigma^k \)

independently and uniformly generate: \( X_1, X_2, \ldots, X_n \in \Sigma \)

let \( Y \) be number of substrings \( \pi \) in \( \langle X_1, X_2, \ldots, X_n \rangle \)

\[ Y = f(X_1, X_2, \ldots, X_n) \]
The Method of Bounded Differences in Action: Pattern Matching

- a random string of length $n$
- a pattern of length $k$
- # of matched substrings?

an alphabet $\Sigma$ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \ldots, X_n \in \Sigma$

let $Y$ be number of substrings $\pi$ in $\langle X_1, X_2, \ldots, X_n \rangle$

$Y = f(X_1, X_2, \ldots, X_n)$

changing any $X_i$ changes $f$ for at most $k$
The Method of Bounded Differences in Action:
Pattern Matching

• a random string of length $n$
• a pattern of length $k$
• # of matched substrings?

an alphabet $\Sigma$ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \ldots, X_n \in \Sigma$

let $Y$ be number of substrings $\pi$ in $\langle X_1, X_2, \ldots, X_n \rangle$

$Y = f(X_1, X_2, \ldots, X_n)$

changing any $X_i$ changes $f$ for at most $k$

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq tk\sqrt{n}) \leq 2e^{-t^2/2}$$
Concentration Inequalities

Question: probability that $X$ deviates more than $\delta$ from expectation?
Concentration Inequalities

Question: probability that $X$ deviates more than $\delta$ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$, then:

For any $\delta > 0$,

$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)(1+\delta)}\right)^\mu$$

For $0 < \delta < 1$,

$$\mathbb{P}(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)(1-\delta)}\right)^\mu$$
Concentration Inequalities

Question: probability that $X$ deviates more than $\delta$ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, then:

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$, then:

for any $t > 0$,

$$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n} (b_i - a_i)^2}\right)$$

$$\mathbb{P}(X \leq \mathbb{E}(X) - t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n} (b_i - a_i)^2}\right)$$
Concentration Inequalities

Question: probability that $X$ deviates more than $\delta$ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X_i$ for any $i$.

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X_i$ for any $i$.

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$ such that for all $k \geq 1$,

$$|Y_k - Y_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|Y_n - Y_0| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{k=1}^{n} c_k^2} \right)$$
Concentration Inequalities

Question: probability that $X$ deviates more than $\delta$ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = (X_1, \cdots, X_n)$. Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$

Let $X = (X_1, \cdots, X_n)$ and let $f$ be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$|\mathbb{E}(f(X) \mid X_1, \cdots, X_i) - \mathbb{E}(f(X) \mid X_1, \cdots, X_{i-1})| \leq c_i$$

Then,

$$\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq t) \leq 2 \exp \left( - \frac{t^2}{2 \sum_{i=1}^{n} c_i^2} \right)$$
Concentration Inequalities

Question: probability that $X$ deviates more than $\delta$ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = (X_1, \cdots, X_n)$ and let $f$ be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$, let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$.

Let $X = (X_1, \cdots, X_n)$ be $n$ independent random variables and let $f(X)$ be a function satisfying the Lipschitz condition with constants $c_i$ where $1 \leq i \leq n$, then:

$$
P(|f(X) - \mathbb{E}(f(X))| \geq t) \leq 2 \exp \left( - \frac{t^2}{2 \sum_{i=1}^{n} c_i^2 } \right)$$
Concentration Inequalities

Question: probability that $X$ deviates more than $\delta$ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = (X_1, \cdots, X_n)$ and let $f$ be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$, $f$ is a function of $X_0, X_1, \cdots, X_n$ and $X_i$.

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots, X_n$. Then:

$$\Pr(|f(X) - \mathbb{E}(f(X))| \geq t) \leq 2 \exp \left( -\frac{t^2}{2 \sum_{i=1}^{n} c_i^2} \right)$$