Martingales - Large Deviation Bound
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. $\mathbb{E}[|Z_n|] < \infty$;
3. $\mathbb{E}[Z_{n+1}|X_0, X_1, \ldots, X_n] = Z_n$;

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

Theorem

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

\[
E[Z_T] = E[Z_0]
\]

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. \( E[T] < \infty \), and there is a constant \( c \) such that \( E[|Z_{i+1} - Z_i| \mid X_1, \ldots, X_i] < c \).
Examples:

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

   What is your expected gain?

2. A couple expects to have $X$ children, $X \sim G(p)$. They expect each of the children to have a number of children distributed $G(r)$.

   What is their expected number of grandchildren?
Wald’s Equation

Theorem

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

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

Note that $T$ is not independent of $X_1, X_2, \ldots$. Corollary of the martingale stopping theorem.
Proof

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

The sequence \( Z_1, Z_2, \ldots \) is a martingale with respect to \( X_1, X_2, \ldots \).

1. \( Z_i \) is determined by \( X_1, \ldots, X_i \)
2. \( \mathbb{E}[|Z_i|] = \mathbb{E}[|\sum_{j=1}^{i} (X_j - \mathbb{E}[X])|] \leq 2i\mathbb{E}[|X|] \)
3. \( \mathbb{E}[Z_{i+1} - Z_i \mid X_0, X_1, \ldots, X_i] = \mathbb{E}[X_{i+1} - \mathbb{E}[X]] = 0 \)

\( \mathbb{E}[Z_1] = 0 \), \( \mathbb{E}[T] < \infty \), and

\[
\mathbb{E}[|Z_{i+1} - Z_i| \mid X_1, \ldots, X_i] = \mathbb{E}[|X_{i+1} - \mathbb{E}[X]|] \leq 2\mathbb{E}[X] .
\]

We can apply the martingale stopping theorem to compute

\[
\mathbb{E}[Z_T] = \mathbb{E}[Z_1] = 0 .
\]
We can apply the martingale stopping theorem to compute

\[ E[Z_T] = E[Z_1] = 0. \]

\begin{align*}
0 &= E[Z_T] = E \left[ \sum_{j=1}^{T} (X_j - E[X]) \right] = E \left[ \sum_{j=1}^{T} X_j - TE[X] \right] \\
&= E \left[ \sum_{j=1}^{T} X_j \right] - E[T] \cdot E[X] = 0,
\end{align*}
Examples

Two stages game:

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

What is your expected gain?

$Y_i =$ outcome of $i$th die in second stage.

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

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

By Wald’s equation:

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

A couple expect to have $X$ children, $X \sim G(p)$. They expect each of their children to have a number of children distributed $G(r)$. What is their expected number of grandchildren?

\[
\frac{1}{p} \cdot \frac{1}{r}
\]
Example: a $k$-run

- We flip a fair coin until we get a consecutive sequence of $k$ HEADs.
- What’s the expected number of times we flip the coin.
- A SWITCH is a HEAD followed by a TAIL.
- Let $X_1$ be the number of flips till $k$ HEADs or the first SWITCH
- Let $X_i$ be the number of flips following the $i-1$ SWITCH till $k$ HEADs or the next SWITCH ($X_i$ includes the last HEAD or TAIL).
- Let $T$ be the first $i$ with $k$ HEADs

$$E[X_i] = \sum_{j \geq 1} j 2^{-j} + \sum_{j=1}^{k-1} j 2^{-j} + (k - 1) 2^{-(k-1)}$$
• Let $X_i$ be the number of flips following the $i - 1$ SWITCH till $k$ HEADs or the next SWITCH ($X_i$ includes the last HEAD or TAIL).

• Let $T$ be the first $i$ with $k$ HEADs

• $X_i = \text{number of flips till (including) first HEAD + up to } k - 2 \text{ HEADs followed by a TAIL, or } k - 1 \text{ HEADS}$

\[ E[X_i] = \sum_{j \geq 1} j2^{-j} + \sum_{j=1}^{k-1} j2^{-j} + (k - 1)2^{-(k-1)} \]

• The probability that $X_i$ ends with $k$ HEADS is $2^{-(k-1)}$ - sequence of $k - 1$ HEADS following the first one.

\[ E[T] = 2^{k-1} \]

• The expected number of coin flips is $E[X_i]E[T]$
Hoeffding’s Bound

Theorem

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

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

Do we need independence?
Tail Inequalities

Theorem (Azuma-Hoeffding Inequality)

Let $Z_0, Z_1, \ldots, Z_n$ be a martingale (with respect to $X_1, X_2, \ldots$) such that $|Z_k - Z_{k-1}| \leq c_k$. Then, for all $t \geq 0$ and any $\lambda > 0$,

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

The following corollary is often easier to apply.

Corollary

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

$$
|X_k - X_{k-1}| \leq c .
$$

Then for all $t \geq 1$ and $\lambda > 0$,

$$
\Pr(|X_t - X_0| \geq \lambda c \sqrt{t}) \leq 2e^{-\lambda^2/2} .
$$
Application: Balls and Bins

We are throwing $m$ balls independently and uniformly at random into $n$ bins.
Let $X_i = 1$ if bin $i$ is empty after all the balls were placed, otherwise $X_i = 1$.

$$E[X_i] = Pr(X_i = 1) = \left(1 - \frac{1}{n}\right)^m$$

Let $F = \sum_{i=1}^{n} X_i$ be the number of empty bins after the $m$ balls are thrown. We know that

$$E[F] = n \left(1 - \frac{1}{n}\right)^m,$$

but the events for different bins are not independent.

Formulating the process as a (Doob) martingale we'll get

$$Pr(|F - E[F]| \geq \epsilon) \leq 2e^{-2\epsilon^2/m}$$
## Theorem (Azuma-Hoeffding Inequality)

Let $Z_0, Z_1, \ldots,$ be a martingale with respect to $X_0, X_1, X_2, \ldots,$ such that

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

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

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

Let \( X^k = X_0, \ldots, X_k \) and \( Y_i = Z_i - Z_{i-1} \).

Since \( E[Z_i \mid X^{i-1}] = Z_{i-1} \),

\[
E[Y_i \mid X^{i-1}] = E[Z_i - Z_{i-1} \mid X^{i-1}] = 0.
\]

Since \( \Pr(B_i \leq Y_i \leq B_i + c_i \mid X^{i-1}) = 1 \), by Hoeffding’s Lemma:

\[
E[e^{\beta Y_i} \mid X^{i-1}] \leq e^{\beta^2 c_i^2 / 8}.
\]

Lemma

(Hoeffding’s Lemma) Let \( X \) be a random variable such that \( \Pr(X \in [a, b]) = 1 \) and \( E[X] = 0 \). Then for every \( \lambda > 0 \),

\[
E[e^{\lambda X}] \leq e^{\lambda^2 (a-b)^2 / 8}.
\]
Proof of the Lemma

Lemma

(Hoeffding’s Lemma) Let $X$ be a random variable such that $\Pr(X \in [a, b]) = 1$ and $E[X] = 0$. Then for every $\lambda > 0$,

$$E[e^{\lambda X}] \leq e^{\lambda^2(b-a)^2/8}.$$

Since $f(x) = e^{\lambda x}$ is a convex function, for any $\alpha \in (0, 1)$ and $x \in [a, b]$,

$$f(X) \leq \alpha f(a) + (1 - \alpha)f(b).$$

Thus, for $\alpha = \frac{b-x}{b-a} \in (0, 1)$,

$$e^{\lambda x} \leq \frac{b-x}{b-a} e^{\lambda a} + \frac{x-a}{b-a} e^{\lambda b}.$$

Taking expectation, and using $E[X] = 0$, we have

$$E \left[ e^{\lambda X} \right] \leq \frac{b}{b-a} e^{\lambda a} - \frac{a}{b-a} e^{\lambda b} \leq e^{\lambda^2(b-a)^2/8}.$$
Proof of Azuma-Hoeffding Inequality

\[ E \left[ e^{\beta Y_i} \mid X_i^{i-1} \right] \leq e^{\beta^2 c_i^2 / 8} . \]

\[
E_X^n \left[ e^{\beta \sum_{i=1}^n Y_i} \right] = E_{X^{n-1}} \left[ E_X^n \left[ e^{\beta \sum_{i=1}^n Y_i} \mid X^{n-1} \right] \right] \\
= E_{X^{n-1}} \left[ e^{\beta \sum_{i=1}^{n-1} Y_i} E_X^n \left[ e^{\beta Y_n} \mid X^{n-1} \right] \right] \\
\leq e^{\beta^2 c_n^2 / 8} E_{X^{n-1}} \left[ e^{\beta \sum_{i=1}^{n-1} Y_i} \right] \\
\leq e^{\beta^2 \sum_{i=1}^n c_i^2 / 8}
\]

In the second inequality we use the fact that \( X^{n-1} \) determines the values of \( Y_1, \ldots, Y_{n-1} \).
\[ Y_i = Z_i - Z_{i-1} \text{ and } \mathbb{E}[e^{\beta \sum_{i=1}^{n} Y_i}] \leq e^{\beta^2 \sum_{i=1}^{n} c_i^2 / 8} \]

\[
\Pr(Z_t - Z_0 \geq \lambda) = \Pr\left(\sum_{i=1}^{t} Y_i \geq \lambda\right) \leq \frac{\mathbb{E}[e^{\beta \sum_{i=1}^{t} Y_i}]}{e^{\beta \lambda}}
\leq e^{-\lambda \beta} e^{\beta^2 \sum_{i=1}^{t} c_i^2 / 8}
\leq 2e^{-2\lambda^2 / (\sum_{k=1}^{t} c_k^2)},
\]

For \( \beta = \frac{4\lambda}{\sum_{i=1}^{t} c_i^2} \).

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

**Theorem (Azuma-Hoeffding Inequality)**

Let \( Z_0, Z_1, \ldots, Z_n \) be a martingale (with respect to \( X_1, X_2, \ldots \)) such that \( |Z_k - Z_{k-1}| \leq c_k \). Then, for all \( t \geq 0 \) and any \( \lambda > 0 \),

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

Assume that you play a sequence of $n$ fair games, where the bet $b_i$ in game $i$ depends on the outcome of previous games. Let $B = \max_i b_i$. The probability of winning or losing more than $\lambda$ is bounded by

$$\Pr(|Z_n| \geq \lambda) \leq 2e^{-2\lambda^2/nB^2}$$

$$\Pr(|Z_n| \geq \lambda B \sqrt{n}) \leq 2e^{-2\lambda^2}$$

$$\Pr \left(|Z_n| \geq \lambda \sqrt{\sum_{i=1}^{n} b_i^2} \right) \leq 2e^{-2\lambda^2}$$
Doob Martingale

Let $X_1, X_2, \ldots, X_n$ be sequence of random variables. Let
$Y = f(X_1, \ldots, X_n)$ be a random variable with $E[|Y|] < \infty$.

For $i = 0, 1, \ldots, n$, let

\[
\begin{align*}
Z_0 &= E[Y] = E_{X[1,n]}f(X_1, \ldots, X_n) \\
Z_i &= E_{X[i+1,n]}[Y|X_1 = x_1, X_2 = x_2, \ldots, X_i = x_i] \\
Z_n &= E[Y|X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n] = f(x_1, \ldots, x_n)
\end{align*}
\]

**Theorem**

$Z_0, Z_1, \ldots, Z_n$ is martingale with respect to $X_1, X_2, \ldots, X_n$. 

Proof

\[ Y = f(X_1, \ldots, X_n), \quad Z_0 = E[Y], \]
\[ Z_i = E_{X[i+1,n]}[Y|X_1 = x_1, \ldots, X_i = x_i], \]

\( Z_1, Z_2, \ldots, Z_n \) is a martingale if \( E[|Z_i|] = E[|Y|] < \infty \), and

\[ E_{X[i+1]}[Z_{i+1}|X_1 = x_1, \ldots, X_i = x_i] = Z_i \]

\[ E_{X[i+1]}[Z_{i+1}|x_1, x_2, \ldots, x_i] = E_{X[i+1]}[E_{X[i+2,n]}[Y|X_1, \ldots, X_{i+1}]|x_1, \ldots, x_i] \]
\[ = E_{X[i+1,n]}[Y|x_1, x_2, \ldots, x_i] \]
\[ = Z_i. \]
Simple Example

\[ Y = f(X_1, \ldots, X_n) = \sum_{i=1}^{n} X_i, \quad X_i \text{ independent } \sim U[0, 1]. \]

\[
Z_0 = E[Y] = E_{X[1,n]} f(X_1, \ldots, X_n) = E[\sum_{i=1}^{n} X_i] = n/2
\]

\[
Z_i = E_{X[i+1,n]} [Y|x_1, \ldots, x_i] = \sum_{j=1}^{i} x_j + E[\sum_{j=i}^{n} X_i] = \sum_{j=1}^{i} x_j + (n - i)/2
\]

\[
Z_n = E[Y|x_1, \ldots, x_n] = f(x_1, \ldots, x_n) = \sum_{j=1}^{n} x_j
\]

\[
E_{X_{i+1}} [Z_{i+1}|x_1, \ldots, x_i] = E_{X_{i+1}} \left[ \sum_{j=1}^{i+1} X_j + \frac{n-i-1}{2} \middle| x_1, \ldots, x_i \right]
\]

\[
= \sum_{j=1}^{i} x_j + \frac{n-i}{2} = Z_i
\]
Example: Polya’s Urn

- Start with $m$ balls, $r$ red, $m - r$ blue.
- Repeat $n$ times:
  1. Pick a ball uniformly at random, check its color and return it to the urn.
  2. If red, add a new red ball, else add a new blue ball.

Let $X_i = 1$ if we add a red ball at step $i$, else $X_i = 0$

We want to estimate the number of new red balls among the $n$ new balls, starting with ratio $r/m$

$$S_n \left( \frac{r}{m} \right) = \sum_{i=1}^{n} X_i = f(X_1, \ldots, X_n)$$

Claim: $E[S_n(\frac{r}{m})] = n \frac{r}{m}$. 
Example: Polya’s Urn

Start with \( M \) balls, \( R \) red, \( M - R \) blue. Repeat \( n \) times: pick a ball uniformly at random. If red add a red ball, else add a blue ball. 

\( X_i = 1 \) if we add a red ball in step \( i \), else \( X_i = 0 \).

\[
S_n(r/m) = \sum_{i=1}^{n} X_i = f(X_1, \ldots, X_n)
\]

Claim: \( E[S_n(r/m)] = n \frac{r}{m} \).

Proof: By induction on \( t \), that \( E[S_t] = tr/m \).

\[
E[S_{t+1} \mid S_t] = S_t + \frac{r + S_t}{m + t}
\]

\[
E[S_{t+1}] = E[E[S_{t+1} \mid S_t]] = E \left[ S_t + \frac{r + S_t}{m + t} \right]
\]

\[
= t \frac{r}{m} + \frac{r + tr/m}{m + t} = t \frac{r}{m} + \frac{r(1 + t/m)}{m(1 + t/m)} = (t + 1) \frac{r}{m}
\]
Example: Polya’s Urn

\( X_i = 1 \) if added a red ball in step \( i \), else \( X_i = 0 \),

\[ S_n(r/m) = \sum_{i=1}^{n} X_i, \text{ and } E[S_n(r/m)] = nr/m \]

Let \( Z_i = E[S_n| X_1 = x_1, \ldots, X_i = x_i] \). We prove that \( Z_1, \ldots, Z_n \) is a martingale.

\[
Z_i = E[S_n| X_1 = x_1, \ldots, X_i = x_i] = \sum_{j=1}^{i} x_j + E[S_{n-i}(\frac{r + \sum_{j=1}^{i} x_j}{m+i})]
\]

\[
= \sum_{j=1}^{i} x_j + (n - i) \frac{r + \sum_{j=1}^{i} x_j}{m+i}
\]

\[
E[Z_{i+1} | X_1, \ldots, X_i] = E[E[S_n|X_1, X_2, \ldots, X_{i+1}] | X_1 = x_1, \ldots, X_i = x_i]
\]

\[
= E \left[ \sum_{j=1}^{i} x_j + X_{i+1} + S_{n-i-1} \left( \frac{r + \sum_{j=1}^{i} x_j + X_{i+1}}{m+i+1} \right) \right]
\]
\[ Z_i = E[ S_n | X_1 = x_1, \ldots, X_i = x_i] = \sum_{j=1}^{i} x_j + (n - i) \frac{r + \sum_{j=1}^{i} x_j}{m + i} \]

\[ E[Z_{i+1} | X_1, \ldots, X_i] = E \left[ \sum_{j=1}^{i} x_j + X_{i+1} + S_{n-i-1} \left( \frac{r + \sum_{j=1}^{i} x_j + X_{i+1}}{m + i + 1} \right) \right] \]

\[ = E \left[ \sum_{j=1}^{i} x_j + X_{i+1} + (n - i - 1) \frac{r + \sum_{j=1}^{i} x_j + X_{i+1}}{m + i + 1} \right] \]

\[ = \sum_{j=1}^{i} x_j + \frac{r + \sum_{j=1}^{i} x_j}{m + i} + (n - i - 1) \frac{r + \sum_{j=1}^{i} x_j + \frac{r + \sum_{j=1}^{i} x_j}{m + i}}{m + i + 1} \]

\[ = \sum_{j=1}^{i} x_j + \frac{r + \sum_{j=1}^{i} x_j}{m + i} + (n - i - 1) \frac{m+i+1}{m+i} \left( \frac{r + \sum_{j=1}^{i} x_j}{m + i + 1} \right) = Z_i \]
Tail Inequalities: Doob Martingales

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

Random variable $Y$:
- $Y$ is a function of $X_1, X_2, \ldots, X_n$;
- $E[|Y|] < \infty$.

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

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

If we can use Azuma-Hoeffding inequality:

$$\Pr(|Z_n - Z_0| \geq \lambda) \leq \ldots.$$  

then we have,

$$\Pr(|Y - E[Y]| \geq \lambda) \leq \ldots$$

We need a bound on $|Z_i - Z_{i-1}|$. 
Example: Pattern Matching

\[ A = (a_1, a_2, \ldots, a_n) \] string of characters, each chosen independently and uniformly at random from \( \Sigma \), with \( m = |\Sigma| \).

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

\( F = \) number occurrences of \( B \) in random string \( S \).

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

Can we bound the deviation of \( F \) from its expectation?
\( F = \) number occurrences of \( B \) in random string \( A \).

\( Z_0 = \mathbb{E}[F] \) and \( Z_n = F \).

\( Z_i = \mathbb{E}[F|a_1, \ldots, a_i], \) for \( i = 1, \ldots, n \).

\( Z_0, Z_1, \ldots, Z_n \) is a Doob martingale.

Each character in \( A \) can participate in no more than \( k \) occurrences of \( B \):

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

Azuma-Hoeffding inequality (version 1):

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

Assume that \( f(X_1, X_2, \ldots, X_n) \) satisfies, for all \( 1 \leq i \leq n \),

\[
\left| f(x_1, \ldots, x_i, \ldots, x_n) - f(x_1, \ldots, y_i, \ldots, x_n) \right| \leq c_i .
\]

and \( X_1, \ldots, X_n \) are independent, then

\[
\Pr(\left| f(X_1, \ldots, X_n) - \mathbb{E}[f(X_1, \ldots, X_n)] \right| \geq \lambda) \leq 2e^{-2\lambda^2/(\sum_{k=1}^{n} c_k^2)} .
\]

[Changing the value of \( X_i \) changes the value of the function by at most \( c_i \).]
Proof

Define a Doob martingale $Z_0, Z_1, \ldots, Z_n$:

- $Z_0 = \mathbb{E}[f(X_1, \ldots, X_n)] = \mathbb{E}[f(\bar{X})]$
- $Z_i = \mathbb{E}[f(X_0, \ldots, X_n) \mid X_1, \ldots, X_i] = \mathbb{E}[f(X_i, \ldots, X_n) \mid X^i]$
- $Z_n = f(X_1, \ldots, X_n) = f(\bar{X})$

We want to prove that this martingale satisfies the conditions of

Theorem (Azuma-Hoeffding Inequality)

Let $Z_0, Z_1, \ldots, Z_n$ be a martingale with respect to $X_0, X_1, X_2, \ldots$, such that

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

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

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

If $X_1, \ldots, X_n$ are independent then for some random variable $B_k$,

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

$$Z_k - Z_{k-1} = \mathbb{E}[f(\bar{X}) \mid X^k] - \mathbb{E}[f(\bar{X}) \mid X^{k-1}].$$

Hence $Z_k - Z_{k-1}$ is bounded above by

$$\sup_x \mathbb{E}[f(\bar{X}) \mid X^{k-1}, X_k = x] - \mathbb{E}[f(\bar{X}) \mid X^{k-1}]$$

and bounded below by

$$\inf_y \mathbb{E}[f(\bar{X}) \mid X^{k-1}, X_k = y] - \mathbb{E}[f(\bar{X}) \mid X^{k-1}].$$

Thus, we need to show

$$\sup_x \mathbb{E}[f(\bar{X}) \mid X^{k-1}, X_k = x] - \inf_y \mathbb{E}[f(\bar{X}) \mid X^{k-1}, X_k = y] \leq c,$$
\[ Z_k - Z_{k-1} = \sup_{x,y} E[f(\bar{X}, X_k = x) - f(\bar{X}, X_k = y) \mid X^{k-1}] \]

Because the \( X_i \) are independent, the values for \( X_{k+1}, \ldots, X_n \) do not depend on the values of \( X_1, \ldots, X_k \).

\[
\sup_{x,y} E[f(\bar{X}, x) - f(\bar{X}, y) \mid X_1 = x_1, \ldots, X_{k-1} = x_{k-1}] = \sup_{x,y} \sum_{x_{k+1}, \ldots, x_n} \Pr((X_{k+1} = x_{k+1}) \cap \ldots \cap (X_n = x_n)) \cdot (f(x_{[1,k-1]}, x, x_{[k+1,n]} - f(x_{[1,k-1]}, y, x_{[k+1,n]}))
\]

But

\[
(f(x_{[1,k-1]}, x, x_{[k+1,n]} - f(x_{[1,k-1]}, y, x_{[k+1,n]}) \leq c_k
\]

and therefore

\[
E[f(\bar{X}, x) - f(\bar{X}, y) \mid X^{k-1}] \leq c_k
\]
We are throwing $m$ balls independently and uniformly at random into $n$ bins. Let $X_i$ be the bin that the $i$th ball falls into. Let $F$ be the number of empty bins after the $m$ balls are thrown.

\[
E[F] = n \left(1 - \frac{1}{n}\right)^m,
\]

The sequence $Z_i = E[F \mid X_1, \ldots, X_i]$ is a Doob martingale. $F = f(X_1, X_2, \ldots, X_m)$ satisfies the Lipschitz condition with bound 1, and the $X_i$'s are independent. We therefore obtain

\[
\Pr(|F - E[F]| \geq \epsilon) \leq 2e^{-2\epsilon^2/m}
\]
Example: Polya’s Urn

- Start with \( m \) balls, \( r \) red, \( m - r \) blue.
- Repeat \( n \) times:
  1. Pick a ball uniformly at random, check its color and return it to the urn.
  2. If red, add a new red ball, else add a new blue ball.

Let \( X_i = 1 \) if we add a red ball at step \( i \), else \( X_i = 0 \)
\[
S_n \left( \frac{r}{m} \right) = \sum_{i=1}^{n} X_i = f(X_1, \ldots, X_n)
\]
satisfies the Lipschitz condition with bound 1, and the \( X_i \)'s are independent.

\[
E[S_n(\frac{r}{m})] = n \frac{r}{m}.
\]

\( Z_i = E[S_n | X_1 = x_1, \ldots, X_i = x_i] \) is a Doob martingale.

\[
\Pr(|S_n - n \frac{r}{m}| \geq \epsilon) \leq 2e^{-2\epsilon^2/m}
\]
Application: Chromatic Number

Given a random graph $G$ in $G_{n,p}$, the chromatic number $\chi(G)$ is the minimum number of colors required to color all vertices of the graph so that no adjacent vertices have the same color. We use the vertex exposure martingale defined below:

Let $G_i$ be the random subgraph of $G$ induced by the set of vertices $1, \ldots, i$, let $Z_0 = E[\chi(G)]$, and let

$$Z_i = E[\chi(G) \mid G_1, \ldots, G_i].$$

Since a vertex uses no more than one new color, again we have that the gap between $Z_i$ and $Z_{i-1}$ is at most 1.

We conclude

$$\Pr(|\chi(G) - E[\chi(G)]| \geq \lambda \sqrt{n}) \leq 2e^{-2\lambda^2}.$$  

This result holds even without knowing $E[\chi(G)]$. 