Balls-into-Bins Model and Chernoff Bounds

Advanced Algorithms
Nanjing University, Fall 2018
Balls-into-Bins Model

$m$ balls
Balls-into-Bins Model

$m$ balls

$n$ bins
Balls-into-Bins Model

$m$ balls
uniformly and independently thrown into

$n$ bins
Balls-into-Bins Model

uniformly at random choose $h: [m] \rightarrow [n]$

$m$ balls
uniformly and independently thrown into

$n$ bins
Balls-into-Bins Model

uniformly at random choose $h: [m] \rightarrow [n]$

$m$ balls

uniformly and independently thrown into $n$ bins

Question: probability that each ball lands in its own bin ($h$ is 1-1)?
Balls-into-Bins Model

\[ h: [m] \to [n] \]

\( m \) balls

uniformly and independently thrown into \( n \) bins

Question: probability that each ball lands in its own bin (\( h \) is 1-1)?

Question: probability that every bin is not empty (\( h \) is onto)?
Balls-into-Bins Model

uniformly at random choose $h: [m] \rightarrow [n]$ 

$m$ balls

uniformly and independently thrown into

$n$ bins

Question: probability that each ball lands in its own bin ($h$ is 1-1)?
Question: probability that every bin is not empty ($h$ is onto)?
Question: maximum number of balls in a bin ($\max \{|h^{-1}(i)|\}$)?
Question: probability that each ball lands in its own bin ($h$ is 1-1)?
Birthday Problem

Question: probability that each ball lands in its own bin ($h$ is 1-1)?
Birthday Problem

Question: probability that each ball lands in its own bin ($h$ is 1-1)?
Birthday Problem

Question: probability that each ball lands in its own bin ($h$ is 1-1)?

$$
\prod_{i=1}^{m-1} \left(1 - \frac{i}{n}\right) \approx \prod_{i=1}^{m-1} e^{-i/n} \approx \exp \left(-\frac{m^2}{2n}\right)
$$
Birthday Problem

Question: probability that each ball lands in its own bin ($h$ is 1-1)?

$$\prod_{i=1}^{m-1} \left(1 - \frac{i}{n}\right) \approx \prod_{i=1}^{m-1} e^{-i/n} \approx \exp \left(-\frac{m^2}{2n}\right)$$

This probability is some constant when $m = \Theta(\sqrt{n})$
Question: probability that every bin is not empty ($h$ is onto)?
Question: probability that every bin is not empty ($h$ is onto)?

$$
\mathbb{P}(\text{some bin is empty}) \leq \sum_{i=1}^{n} \mathbb{P}(\text{bin } i \text{ is empty}) = n \cdot \left(1 - \frac{1}{n}\right)^m \approx n \cdot e^{-m/n}
$$
Question: probability that every bin is not empty ($h$ is onto)?

$$\mathbb{P}(\text{some bin is empty}) \leq \sum_{i=1}^{n} \mathbb{P}(\text{bin } i \text{ is empty}) = n \cdot \left(1 - \frac{1}{n}\right)^m \approx n \cdot e^{-m/n}$$

Question: how many balls we need to throw to leave no empty bins?
Coupon Collector

Question: how many balls we need to throw to leave no empty bins?
Coupon Collector

Question: how many balls we need to throw to leave no empty bins?
Coupon Collector

Question: how many balls we need to throw to leave no empty bins?

Let $X_i$ be the number of balls thrown until $i$ bins are non-empty, given $i-1$ bins are already non-empty.
Question: how many balls we need to throw to leave no empty bins?

Let $X_i$ be the number of balls thrown until $i$ bins are non-empty, given $i-1$ bins are already non-empty.

$X_i$ is a geometric r.v. with parameter $(n-i+1)/n$. 
Question: how many balls we need to throw to leave no empty bins?

Let $X_i$ be the number of balls thrown until $i$ bins are non-empty, given $i-1$ bins are already non-empty.

$X_i$ is a geometric r.v. with parameter $(n-i+1)/n$.

$$
\mathbb{E}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathbb{E}(X_i) = \sum_{i=1}^{n} \frac{n}{n - i + 1} = \sum_{i=1}^{n} \frac{n}{i} = nH_n = n \ln n + O(n)
$$
Load Balancing

Question: maximum number of balls in a bin (max\{\{|h^{-1}(i)|\}\})?
Load Balancing

Question: maximum number of balls in a bin \( \max\{|h^{-1}(i)|\} \)?

Let \( X_i \) be the number of balls in bin \( i \). That is, \( X_i = |h^{-1}(i)| \).
Load Balancing

Question: maximum number of balls in a bin \( \left( \max\{|h^{-1}(i)|\} \right) \)?

Let \( X_i \) be the number of balls in bin \( i \). That is, \( X_i = |h^{-1}(i)| \).

\[
\mathbb{E}(X_i) =
\]
Load Balancing

Question: maximum number of balls in a bin ($\max\{|h^{-1}(i)|\}$)?

Let $X_i$ be the number of balls in bin $i$. That is, $X_i = |h^{-1}(i)|$.

Let $Y_{ij}$ be i.r.v. taking value 1 iff ball $j$ lands in bin $i$.

$\mathbb{E}(X_i) =$
Load Balancing

Question: maximum number of balls in a bin (max{|h^{-1}(i)|})?

Let $X_i$ be the number of balls in bin $i$. That is, $X_i = |h^{-1}(i)|$.

Let $Y_{ij}$ be i.r.v. taking value 1 iff ball $j$ lands in bin $i$.

$$
\mathbb{E}(X_i) = \mathbb{E} \left( \sum_{j=1}^{m} Y_{ij} \right) = \sum_{j=1}^{m} \mathbb{E}(Y_{ij}) = m \cdot \frac{1}{n} = \frac{m}{n}
$$
Load Balancing

Question: maximum number of balls in a bin ($\max\{|h^{-1}(i)|\}$)?

Let $X_i$ be the number of balls in bin $i$. That is, $X_i = |h^{-1}(i)|$.

Let $Y_{ij}$ be i.r.v. taking value 1 iff ball $j$ lands in bin $i$.

$$\mathbb{E}(X_i) = \mathbb{E}\left( \sum_{j=1}^{m} Y_{ij} \right) = \sum_{j=1}^{m} \mathbb{E}(Y_{ij}) = m \cdot \frac{1}{n} = \frac{m}{n}$$

Is $\max\{\mathbb{E}(X_i)\}$ what we want?
Load Balancing

Question: maximum number of balls in a bin (max\{|h^{-1}(i)|\})?

Let $X_i$ be the number of balls in bin $i$. That is, $X_i = |h^{-1}(i)|$.

Let $Y_{ij}$ be i.r.v. taking value 1 iff ball $j$ lands in bin $i$.

$$
\mathbb{E}(X_i) = \mathbb{E}\left(\sum_{j=1}^{m} Y_{ij}\right) = \sum_{j=1}^{m} \mathbb{E}(Y_{ij}) = m \cdot \frac{1}{n} = \frac{m}{n}
$$

Is max\{\mathbb{E}(X_i)\} what we want?

For every $i$, $\mathbb{E}(X_i)$ is $m/n$, so max\{\mathbb{E}(X_i)\} is simply $m/n$.

$$
\max\{\mathbb{E}(X_i)\} = \frac{m}{n}
$$
Load Balancing

Question: maximum number of balls in a bin?

$$\max\{\mathbb{E}(X_i)\} = \frac{m}{n}$$
Load Balancing

Question: maximum number of balls in a bin?

\[ \max \{ \mathbb{E}(X_i) \} = \frac{m}{n} \]
Load Balancing

Question: maximum number of balls in a bin?

$$\max\{\mathbb{E}(X_i)\} = \frac{m}{n}$$ Something is not right...
Load Balancing

Question: maximum number of balls in a bin?

When $m = \Theta(n)$:
the max load is $O\left(\frac{\log n}{\log \log n}\right)$ with high probability.

When $m = \Omega(n \log n)$:
the max load is $O\left(\frac{m}{n}\right)$ with high probability.
Load Balancing

Question: maximum number of balls in a bin?

When $m = \Theta(n)$:
the max load is $O\left(\frac{\log n}{\log \log n}\right)$ with high probability.

When $m = \Omega(n \log n)$:
the max load is $O\left(\frac{m}{n}\right)$ with high probability.

*with high probability (w.h.p.):*
We say an event happens with high probability (with respect to $n$) if it happens with probability at least $1-1/n$. 
Load Balancing

When $m = \Theta(n)$:

the max load is $O \left( \frac{\log n}{\log \log n} \right)$ with high probability.
Load Balancing

When \( m = \Theta(n) \):

the max load is \( O \left( \frac{\log n}{\log \log n} \right) \) with high probability.

\[
\mathbb{P} (\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P} (X_i \geq t)
\]
Load Balancing

When $m = \Theta(n)$:

the max load is $O\left(\frac{\log n}{\log \log n}\right)$ with high probability.

\[
\mathbb{P}(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \geq t) \leq \frac{1}{n}
\]
Load Balancing

When $m = \Theta(n)$:

the max load is $O\left(\frac{\log n}{\log \log n}\right)$ with high probability.

\[
\mathbb{P} \left( \exists i : X_i \geq t \right) \leq \sum_{i=1}^{n} \mathbb{P} \left( X_i \geq t \right) \leq \frac{1}{n}
\]

So we need $\mathbb{P} \left( X_i \geq t \right) \leq \frac{1}{n^2}$
Load Balancing

When \( m = \Theta(n) \):
the max load is \( O \left( \frac{\log n}{\log \log n} \right) \) with high probability.

\[
\mathbb{P} \left( \exists i : X_i \geq t \right) \leq \sum_{i=1}^{n} \mathbb{P} \left( X_i \geq t \right) \leq \frac{1}{n}
\]

So we need \( \mathbb{P} \left( X_i \geq t \right) \leq \frac{1}{n^2} \)

\[
\mathbb{P}(X_i \geq t) \leq \binom{m}{t} \left( \frac{1}{n} \right)^t \leq \left( \frac{em}{t} \right)^t \left( \frac{1}{n} \right)^t
\]
Load Balancing

When \( m = \Theta(n) \):
the max load is \( O \left( \frac{\log n}{\log \log n} \right) \) with high probability.

\[
P(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} P(X_i \geq t) \leq \frac{1}{n}
\]

So we need \( P(X_i \geq t) \leq \frac{1}{n^2} \)

\[
P(X_i \geq t) \leq \binom{m}{t} \left( \frac{1}{n} \right)^t \leq \left( \frac{em}{t} \right)^t \left( \frac{1}{n} \right)^t
\]

\[
= \left( \frac{e}{t} \right)^t \quad \text{let } m=n
\]
Load Balancing

When \( m = \Theta(n) \):

the max load is \( O \left( \frac{\log n}{\log \log n} \right) \) with high probability.

\[
\mathbb{P} (\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P} (X_i \geq t) \leq \frac{1}{n}
\]

So we need \( \mathbb{P} (X_i \geq t) \leq \frac{1}{n^2} \)

\[
\mathbb{P}(X_i \geq t) \leq \binom{m}{t} \left( \frac{1}{n} \right)^t \leq \left( \frac{em}{t} \right)^t \left( \frac{1}{n} \right)^t
\]

\[
= \left( \frac{e}{t} \right)^t = \left( \frac{e \ln \ln n}{3 \ln n} \right)^{\frac{3 \ln n}{\ln \ln n}} \quad \text{let } t = 3 \ln(n)/\ln\ln(n)
\]
Load Balancing

When \( m = \Theta(n) \):

the max load is \( O \left( \frac{\log n}{\log \log n} \right) \) with high probability.

\[
\mathbb{P} (\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P} (X_i \geq t) \leq \frac{1}{n}
\]

So we need \( \mathbb{P}(X_i \geq t) \leq \frac{1}{n^2} \)

\[
\mathbb{P}(X_i \geq t) \leq \binom{m}{t} \left( \frac{1}{n} \right)^t \leq \left( \frac{em}{t} \right)^t \left( \frac{1}{n} \right)^t
\]

\[
= \left( \frac{e}{t} \right)^t = \left( \frac{e \ln \ln n}{3 \ln n} \right)^{\frac{3 \ln n}{\ln \ln n}}
\]

for sufficiently large \( n \)

\[
\leq \left( \frac{\ln \ln n}{\ln n} \right)^{\frac{3 \ln n}{\ln \ln n}} = (e^{\ln \ln n - \ln \ln n})^{\frac{3 \ln n}{\ln \ln n}} = e^{-3 \ln n + o(\ln n)} \leq e^{-2 \ln n} = \frac{1}{n^2}
\]
Load Balancing

When $m = \Omega(n \log n)$:
the max load is $O\left(\frac{m}{n}\right)$ with high probability.
Load Balancing

When $m = \Omega(n \log n)$:
the max load is $O\left(\frac{m}{n}\right)$ with high probability.

$$\mathbb{P}(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \geq t) \leq \frac{1}{n}$$
Load Balancing

When \( m = \Omega(n \log n) \):
the max load is \( O\left(\frac{m}{n}\right) \) with high probability.

\[
P(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} P(X_i \geq t) \leq \frac{1}{n}
\]

So we need \( P(X_i \geq t) \leq \frac{1}{n^2} \)
Load Balancing

When \( m = \Omega(n \log n) \):
the max load is \( O\left(\frac{m}{n}\right) \) with high probability.

\[
\mathbb{P}\left( \exists i : X_i \geq t \right) \leq \sum_{i=1}^{n} \mathbb{P}\left( X_i \geq t \right) \leq \frac{1}{n}
\]

So we need \( \mathbb{P}\left( X_i \geq t \right) \leq \frac{1}{n^2} \)

\[
\mathbb{P}(X_i \geq t) \leq \binom{m}{t} \left(\frac{1}{n}\right)^t \leq \left(\frac{em}{t}\right)^t \left(\frac{1}{n}\right)^t
\]
Load Balancing

When $m = \Omega(n \log n)$: the max load is $O\left(\frac{m}{n}\right)$ with high probability.

\[
P(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} P(X_i \geq t) \leq \frac{1}{n}
\]

So we need $P(X_i \geq t) \leq \frac{1}{n^2}$

\[
P(X_i \geq t) \leq \binom{m}{t} \left(\frac{1}{n}\right)^t \leq \left(\frac{em}{t}\right)^t \left(\frac{1}{n}\right)^t
\]

\[
= \left(\frac{e \log n}{t}\right)^t \quad \text{let } m = n\log(n)
\]
Load Balancing

When \( m = \Omega(n \log n) \): the max load is \( O\left(\frac{m}{n}\right) \) with high probability.

\[
\mathbb{P}(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \geq t) \leq \frac{1}{n}
\]

So we need \( \mathbb{P}(X_i \geq t) \leq \frac{1}{n^2} \)

\[
\mathbb{P}(X_i \geq t) \leq \binom{m}{t} \left(\frac{1}{n}\right)^t \leq \left(\frac{em}{t}\right)^t \left(\frac{1}{n}\right)^t
\]

\[
= \left(\frac{e \lg n}{t}\right)^t = \left(\frac{e}{4}\right)^{4 \lg n}
\]

let \( t = 4m/n = 4\lg(n) \)
Load Balancing

When \( m = \Omega(n \log n) \): the max load is \( O\left(\frac{m}{n}\right) \) with high probability.

\[
P(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} P(X_i \geq t) \leq \frac{1}{n}
\]

So we need \( P(X_i \geq t) \leq \frac{1}{n^2} \)

\[
P(X_i \geq t) \leq \binom{m}{t} \left(\frac{1}{n}\right)^t \leq \left(\frac{em}{t}\right)^t \left(\frac{1}{n}\right)^t
\]

\[
= \left(\frac{e \log n}{t}\right)^t = \left(\frac{e}{4}\right)^{4 \log n}
\]

let \( t = 4m/n = 4\log(n) \)

\[
\leq \left(\frac{1}{2}\right)^{2 \log n} = \frac{1}{n^2}
\]
Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

“uniformly at random choose $h$: $[m] \rightarrow [n]$”

Question: maximum number of balls in a bin ($\max\{|h^{-1}(i)|\}$)?

When $m = \Theta(n)$:
the max load is $O\left(\frac{\log n}{\log \log n}\right)$ with high probability.

When $m = \Omega(n \log n)$:
the max load is $O\left(\frac{m}{n}\right)$ with high probability.
Concentration

balls into bins

coin flipping
Concentration

balls into bins

coin flipping

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

Herman Chernoff
Chernoff Bounds

Life in Los Angeles

Herman Chernoff
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 $\delta > 0$,

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

for $0 < \delta < 1$,

$$
P(X \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}
$$
(Convenient) Chernoff Bounds

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 $0 < \delta < 1$,
\[
\begin{align*}
\mathbb{P}(X \geq (1 + \delta)\mu) &\leq \exp \left( -\frac{\mu \delta^2}{3} \right) \\
\mathbb{P}(X \leq (1 - \delta)\mu) &\leq \exp \left( -\frac{\mu \delta^2}{2} \right)
\end{align*}
\]

for $t \geq 2e\mu$,
\[
\mathbb{P}(X \geq t) \leq 2^{-t}
\]
Power of the Chernoff Bounds

\[ \mathbb{P}(X \geq (1 + \delta)\mu) \text{ when } \mu = 50 \]
Power of the Chernoff Bounds

\[ \Pr(X \geq (1 + \delta)\mu) \text{ when } \mu = 50 \]
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

**Question:** maximum number of balls in a bin?

$X_i$: load of bin $i$

$Y_{ij}$: i.r.v. taking value 1 iff ball $j$ lands in bin $i$
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

$X_i$: load of bin $i$

$Y_{ij}$: i.r.v. taking value 1 iff ball $j$ lands in bin $i$

$$
\mu = \mathbb{E}(X_i) = \mathbb{E} \left( \sum_{j=1}^{m} Y_{ij} \right) = \sum_{j=1}^{m} \mathbb{E}(Y_{ij}) = \frac{m}{n}
$$

$X_i \sim \text{Binomial} \left( m, \frac{1}{n} \right)$
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

**Question:** maximum number of balls in a bin?

$X_i$: load of bin $i$

$Y_{ij}$: i.r.v. taking value 1 iff ball $j$ lands in bin $i$

$$\mu = \mathbb{E}(X_i) = \mathbb{E} \left( \sum_{j=1}^{m} Y_{ij} \right) = \sum_{j=1}^{m} \mathbb{E}(Y_{ij}) = \frac{m}{n}$$

$X_i \sim \text{Binomial} \left( m, \frac{1}{n} \right)$

For $m = n$, $\mu = 1$

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

implies
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

$X_i$: load of bin $i$
$Y_{ij}$: i.r.v. taking value 1 iff ball $j$ lands in bin $i$

$$
\mu = \mathbb{E}(X_i) = \mathbb{E} \left( \sum_{j=1}^{m} Y_{ij} \right) = \sum_{j=1}^{m} \mathbb{E}(Y_{ij}) = \frac{m}{n} \quad X_i \sim \text{Binomial} \left( m, \frac{1}{n} \right)
$$

For $m = n$, $\mu = 1$

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

implies

$$
\mathbb{P}(X_i \geq t) \leq \frac{e^t}{et^t} \leq \frac{1}{e} \left( \frac{\ln \ln n}{\ln n} \right)^{\frac{e \ln n}{\ln \ln n}} \leq \frac{1}{n^2} \quad \text{when } t \geq \frac{e \ln n}{\ln \ln n} \text{ and } n \text{ sufficiently large}
$$
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

$X_i$: load of bin $i$

$Y_{ij}$: i.r.v. taking value 1 iff ball $j$ lands in bin $i$

$$\mu = \mathbb{E}(X_i) = \mathbb{E}\left(\sum_{j=1}^{m} Y_{ij}\right) = \sum_{j=1}^{m} \mathbb{E}(Y_{ij}) = \frac{m}{n} \quad X_i \sim \text{Binomial}\left(m, \frac{1}{n}\right)$$

For $m = n, \mu = 1$

$$\mathbb{P}(X_i \geq (1 + \delta)\mu) \leq \left(\frac{e^{\delta}}{(1 + \delta)(1 + \delta)}\right)^{\mu} \quad \text{implies}$$

$$\mathbb{P}(X_i \geq t) \leq \frac{e^t}{et} \leq \frac{1}{e} \left(\frac{\ln \ln n}{\ln n}\right)^{\frac{e \ln n}{\ln n}} \leq \frac{1}{n^2} \quad \text{when } t \geq \frac{e \ln n}{\ln \ln n} \text{ and } n \text{ sufficiently large}$$

$$\mathbb{P}(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \geq t) \leq \frac{1}{n}$$
Chernoff Bounds in Action: 
Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

$X_i$: load of bin $i$
$Y_{ij}$: i.r.v. taking value 1 iff ball $j$ lands in bin $i$

$$
\mu = \mathbb{E}(X_i) = \mathbb{E}\left(\sum_{j=1}^{m} Y_{ij}\right) = \sum_{j=1}^{m} \mathbb{E}(Y_{ij}) = \frac{m}{n}
$$

$X_i \sim \text{Binomial}\left(m, \frac{1}{n}\right)$

For $m = n$, $\mu = 1$

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

implies

$$
\mathbb{P}(X_i \geq t) \leq \frac{e^t}{et^t} \leq \frac{1}{e} \left(\frac{\ln \ln n}{\ln n}\right)^{\frac{e \ln n}{\ln n}} \leq \frac{1}{n^2}
$$

when $t \geq \frac{e \ln n}{\ln \ln n}$ and $n$ sufficiently large

$$
\mathbb{P}(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \geq t) \leq \frac{1}{n}
$$

when $m = \Theta(n)$ max load is $O\left(\frac{\log n}{\log \log n}\right)$ w.h.p.
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

$$\mu = \mathbb{E}(X_i) = \frac{m}{n} \quad X_i \sim \text{Binomial} \ (m, \frac{1}{n})$$

For $m \geq n \ln n$, $\mu \geq \ln n$
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

\[ \mu = \mathbb{E}(X_i) = \frac{m}{n} \quad X_i \sim \text{Binomial } (m, \frac{1}{n}) \]

For $m \geq n \ln n$, $\mu \geq \ln n$

\[ \mathbb{P}(X_i \geq t) \leq 2^{-t} \quad \text{when } \quad t \geq 2e\mu \quad \text{implies} \]
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

$$\mu = \mathbb{E}(X_i) = \frac{m}{n} \quad X_i \sim \text{Binomial} \left( m, \frac{1}{n} \right)$$

For $m \geq n \ln n$, $\mu \geq \ln n$

$$\mathbb{P}(X_i \geq t) \leq 2^{-t} \quad \text{when} \quad t \geq 2e\mu \quad \text{implies}$$

$$\mathbb{P}(X_i \geq 2e\mu) \leq 2^{-2e\mu} \leq 2^{-2e\ln n} \leq \frac{1}{n^2}$$
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

$$\mu = \mathbb{E}(X_i) = \frac{m}{n} \quad X_i \sim \text{Binomial} \left( m, \frac{1}{n} \right)$$

For $m \geq n \ln n$, $\mu \geq \ln n$

$$\mathbb{P}(X_i \geq t) \leq 2^{-t} \quad \text{when} \quad t \geq 2e\mu \quad \text{implies}$$

$$\mathbb{P}(X_i \geq 2e\mu) \leq 2^{-2e\mu} \leq 2^{-2e \ln n} \leq \frac{1}{n^2}$$

$$\mathbb{P}(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \geq t) \leq \frac{1}{n}$$
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

\[
\mu = \mathbb{E}(X_i) = \frac{m}{n} \quad X_i \sim \text{Binomial} \left( m, \frac{1}{n} \right)
\]

For $m \geq n \ln n$, $\mu \geq \ln n$

\[
\mathbb{P}(X_i \geq t) \leq 2^{-t} \quad \text{when} \quad t \geq 2e\mu \quad \text{implies}
\]

\[
\mathbb{P}(X_i \geq 2e\mu) \leq 2^{-2e\mu} \leq 2^{-2e \ln n} \leq \frac{1}{n^2}
\]

\[
\mathbb{P} \left( \exists i : X_i \geq t \right) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \geq t) \leq \frac{1}{n}
\]

When $m = \Omega(n \log n)$, max load is $O \left( \frac{m}{n} \right)$ w.h.p.
Chernoff Bounds in Action: Load Balancing

“m balls are thrown into n bins uniformly and independently at random”

Question: maximum number of balls in a bin?

\[ \mu = \mathbb{E}(X_i) = \frac{m}{n} \quad X_i \sim \text{Binomial} \left( m, \frac{1}{n} \right) \]

For \( m \geq n \ln n \), \( \mu \geq \ln n \)

\[ \mathbb{P}(X_i \geq t) \leq 2^{-t} \quad \text{when} \quad t \geq 2e\mu \quad \text{imply} \]

\[ \mathbb{P}(X_i \geq 2e\mu) \leq 2^{-2e\mu} \leq 2^{-2e \ln n} \leq \frac{1}{n^2} \]

\[ \mathbb{P}(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \geq t) \leq \frac{1}{n} \]

\[ \mathbb{P}(X_i \leq (1 - \delta)\mu) \leq \exp \left( -\frac{\mu\delta^2}{2} \right) \quad \text{imply} \]

\[ \mathbb{P}(\exists i : X_i \leq \mu/2) \leq \frac{1}{n} \quad \text{when} \quad u \geq 16\ln n \]
Chernoff Bounds in Action: Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

$$\mu = \mathbb{E}(X_i) = \frac{m}{n} \quad X_i \sim \text{Binomial} \left( m, \frac{1}{n} \right)$$

For $m \geq n \ln n$, $\mu \geq \ln n$

$$\mathbb{P}(X_i \geq t) \leq 2^{-t} \quad \text{when} \quad t \geq 2e\mu \quad \text{implies}$$

$$\mathbb{P}(X_i \geq 2e\mu) \leq 2^{-2e\mu} \leq 2^{-2e\ln n} \leq \frac{1}{n^2}$$

$$\mathbb{P}(\exists i : X_i \geq t) \leq \sum_{i=1}^{n} \mathbb{P}(X_i \geq t) \leq \frac{1}{n}$$

when $m = \Omega(n \log n)$ max load is $O \left( \frac{m}{n} \right)$ w.h.p.

$$\mathbb{P}(X_i \leq (1 - \delta)\mu) \leq \exp \left( -\frac{\mu\delta^2}{2} \right) \quad \text{implies}$$

$$\mathbb{P}(\exists i : X_i \leq \mu/2) \leq \frac{1}{n} \quad \text{when} \quad u \geq 16 \ln n$$

when $m = \Omega(n \log n)$ min load is $\Omega \left( \frac{m}{n} \right)$ w.h.p.
Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

When $m = \Theta(n)$:
The max load is $O\left(\frac{\log n}{\log \log n}\right)$ with high probability.

When $m = \Omega(n \log n)$:
The max load is $O\left(\frac{m}{n}\right)$ with high probability.
Load Balancing

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

When $m = \Theta(n)$:
the max load is $O\left(\frac{\log n}{\log \log n}\right)$ with high probability.

When $m = \Omega(n \log n)$:
the max load is $O\left(\frac{m}{n}\right)$ with high probability.

When $m = \Omega(n \log n)$:
each bin’s load is $\Theta\left(\frac{m}{n}\right)$ with high probability.
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 $0 < \delta < 1$,

$$\mathbb{P}(X \leq (1 - \delta)\mu) \leq \left( \frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}} \right)^\mu$$
Generalization of Markov’s Inequality

For any $X$, for $h : X \to \mathbb{R}^+$, for any $t > 0$,

$$\mathbb{P}(h(X) \geq t) \leq \frac{\mathbb{E}(h(X))}{t}$$
Moment Generating Functions

The moment generating function of $X$ is:

$$M(\lambda) = \mathbb{E}(e^{\lambda X})$$
Moment Generating Functions

The moment generating function of $X$ is:

$$M(\lambda) = \mathbb{E}(e^{\lambda X})$$

by Taylor’s expansion:

$$\mathbb{E}(e^{\lambda X}) = \mathbb{E} \left( \sum_{k=0}^{\infty} \frac{X^k}{k!} \lambda^k \right) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \mathbb{E}(X^k)$$
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

$\mathbb{P}(X \geq (1 + \delta)\mu) \leq ?$ for $\delta > 0$
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq ? \quad \text{for} \quad \delta > 0$$

$$\mathbb{P}(X \geq (1 + \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}}$$

$\lambda > 0$, and

generalized Markov’s inequality
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

$\mathbb{P}(X \geq (1 + \delta)\mu) \leq? \quad \text{for} \quad \delta > 0$

$\mathbb{P}(X \geq (1 + \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}}$

$\lambda > 0$, and

generalized Markov’s inequality
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq? \quad \text{for} \quad \delta > 0$$

$$\mathbb{P}(X \geq (1 + \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}}$$

$\lambda > 0$, and
generalized Markov’s inequality

$$= \mathbb{E}(e^{\lambda \sum_{i=1}^{n} X_i}) = \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda X_i}\right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i})$$

independence of $X_i$

not linearity of expectation
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = E(X)$

$P(X \geq (1 + \delta)\mu) \leq ?$ for $\delta > 0$

$P(X \geq (1 + \delta)\mu) = P(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{E(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}}$

$\lambda > 0$, and

generalized Markov’s inequality

$= E(e^{\lambda \sum_{i=1}^{n} X_i}) = E\left(\prod_{i=1}^{n} e^{\lambda X_i}\right) = \prod_{i=1}^{n} E(e^{\lambda X_i})$

independence of $X_i$

not linearity of expectation
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

$\mathbb{P}(X \geq (1+\delta)\mu) \leq \ ? \text{ for } \delta > 0$

$\mathbb{P}(X \geq (1+\delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}}$

\[ \lambda > 0, \text{ and } \]
\[ \text{generalized Markov's inequality} \]

$\mathbb{E}(e^{\lambda \sum_{i=1}^{n} X_i}) = \mathbb{E}\left( \prod_{i=1}^{n} e^{\lambda X_i} \right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i})$

\[ \text{independence of } X_i \]
\[ \text{not linearity of expectation} \]

assume $\mathbb{P}(X_i = 1) = p_i$, then $\mu = \sum_{i=1}^{n} p_i$
independent \( X_1, X_2, \cdots, X_n \in \{0, 1\} \), let \( X = \sum_{i=1}^{n} X_i \), and \( \mu = \mathbb{E}(X) \)

\[ \mathbb{P}(X \geq (1 + \delta)\mu) \leq \quad \text{for } \delta > 0 \]

\[ \mathbb{P}(X \geq (1 + \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}} \]

\[ \lambda > 0, \text{ and} \]

generalized Markov’s inequality

\[ \mathbb{E}(e^{\lambda \sum_{i=1}^{n} X_i}) = \mathbb{E}\left( \prod_{i=1}^{n} e^{\lambda X_i} \right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i}) \]

independence of \( X_i \)

not linearity of expectation

assume \( \mathbb{P}(X_i = 1) = p_i \), then \( \mu = \sum_{i=1}^{n} p_i \)

\[ = \mathbb{P}(X_i = 1) \cdot e^{\lambda} + \mathbb{P}(X_i = 0) \cdot e^{0} = p_i e^{\lambda} + (1 - p_i) = 1 + p_i(e^{\lambda} - 1) \leq e^{p_i(e^{\lambda} - 1)} \]
independent \( X_1, X_2, \cdots, X_n \in \{0, 1\} \), let \( X = \sum_{i=1}^{n} X_i \), and \( \mu = \mathbb{E}(X) \)

\[ \mathbb{P}(X \geq (1 + \delta)\mu) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda (1 + \delta)\mu}} \]

\( \lambda > 0 \), and generalized Markov’s inequality

\[ \mathbb{E}(e^{\lambda \sum_{i=1}^{n} X_i}) = \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda X_i}\right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i}) \]

not linearity of expectation

assume \( \mathbb{P}(X_i = 1) = p_i \), then \( \mu = \sum_{i=1}^{n} p_i \)

\[ \mathbb{P}(X_i = 1) \cdot e^{\lambda} + \mathbb{P}(X_i = 0) \cdot e^{0} = p_i e^{\lambda} + (1 - p_i) = 1 + p_i(e^{\lambda} - 1) \leq e^{p_i(e^{\lambda} - 1)} \quad 1 + y \leq e^{y} \]
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq? \quad \text{for} \quad \delta > 0$$

$$\mathbb{P}(X \geq (1 + \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}}$$

$\lambda > 0$, and generalized Markov’s inequality

$$e^{\lambda \sum_{i=1}^{n} X_i} = \mathbb{E}(\prod_{i=1}^{n} e^{\lambda X_i}) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i})$$

$\leq \prod_{i=1}^{n} e^{p_i(e^{\lambda} - 1)} = e^{\mu(e^{\lambda} - 1)}$

independence of $X_i$

not linearity of expectation

assume $\mathbb{P}(X_i = 1) = p_i$, then $\mu = \sum_{i=1}^{n} p_i$

$$= \mathbb{P}(X_i = 1) \cdot e^{\lambda} + \mathbb{P}(X_i = 0) \cdot e^{0} = p_i e^{\lambda} + (1 - p_i) = 1 + p_i(e^{\lambda} - 1) \leq e^{p_i(e^{\lambda} - 1)} \quad 1 + y \leq e^{y}$$
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq? \quad \text{for} \quad \delta > 0$$

$$\mathbb{P}(X \geq (1 + \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}}$$

$$\leq \frac{e^{u(e^{\lambda}-1)}}{e^{u\lambda(1+\delta)}} = \left(\frac{e^{(e^{\lambda}-1)}}{e^{\lambda(1+\delta)}}\right)^\mu$$

$$= \mathbb{E}(e^{\lambda \sum_{i=1}^{n} X_i}) = \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda X_i}\right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i})$$

$$\leq \prod_{i=1}^{n} e^{p_i(e^{\lambda}-1)} = e^{\mu(e^{\lambda}-1)}$$

assume $\mathbb{P}(X_i = 1) = p_i$, then $\mu = \sum_{i=1}^{n} p_i$

$$= \mathbb{P}(X_i = 1) \cdot e^{\lambda} + \mathbb{P}(X_i = 0) \cdot e^{0} = p_i e^{\lambda} + (1 - p_i) = 1 + p_i(e^{\lambda} - 1) \leq e^{p_i(e^{\lambda}-1)} \quad \text{for} \quad 1 + y \leq e^y$$
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

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

$$\leq \frac{e^{u(e^{\lambda}-1)}}{e^{u\lambda(1+\delta)}} = \left(\frac{e^{(e^{\lambda}-1)}}{e^{\lambda(1+\delta)}}\right)^\mu \leq \left(\frac{e^{\delta}}{(1 + \delta)^{1+\delta}}\right)^\mu \text{ minimized when } \lambda = \ln(1 + \delta)$$

$$= \mathbb{E}(e^{\lambda \sum_{i=1}^{n} X_i}) = \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda X_i}\right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i})$$

$$\leq \prod_{i=1}^{n} e^{p_i(e^{\lambda}-1)} = e^{\mu(e^{\lambda}-1)}$$

assume $\mathbb{P}(X_i = 1) = p_i$, then $\mu = \sum_{i=1}^{n} p_i$

$$= \mathbb{P}(X_i = 1) \cdot e^{\lambda} + \mathbb{P}(X_i = 0) \cdot e^{0} = p_ie^{\lambda} + (1 - p_i) = 1 + p_i(e^{\lambda} - 1) \leq e^{p_i(e^{\lambda}-1)} \quad 1 + y \leq e^{y}$$
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq? \quad \text{for } \delta > 0$$

$$\mathbb{P}(X \geq (1 + \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}} \leq \frac{e^\lambda}{(1 + \delta)^{1+\delta}}$$

\(\lambda > 0\), and generalized Markov’s inequality minimized when \(\lambda = \ln(1 + \delta)\)

(a) apply Markov’s inequality to moment generating function

$$= \mathbb{E}(e^{\lambda \sum_{i=1}^{n} X_i}) = \mathbb{E} \left( \prod_{i=1}^{n} e^{\lambda X_i} \right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i})$$

independence of $X_i$

not linearity of expectation

$$\leq \prod_{i=1}^{n} e^{p_i(e^\lambda - 1)} = e^{\mu(e^\lambda - 1)}$$

assume $\mathbb{P}(X_i = 1) = p_i$, then $\mu = \sum_{i=1}^{n} p_i$

$$= \mathbb{P}(X_i = 1) \cdot e^\lambda + \mathbb{P}(X_i = 0) \cdot e^0 = p_i e^\lambda + (1 - p_i) = 1 + p_i (e^\lambda - 1) \leq e^{p_i(e^\lambda - 1)}$$

$1 + y \leq e^y$
independent $X_1, X_2, \cdots, X_n \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$

$$\mathbb{P}(X \geq (1 + \delta)\mu) \leq? \quad \text{for} \quad \delta > 0$$

$$\mathbb{P}(X \geq (1 + \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}}$$

(a) apply Markov’s inequality to moment generating function

$$\leq \frac{e^{u(e^{\lambda X} - 1)}}{e^{u\lambda(1+\delta)}} = \left(\frac{e^{(e^{\lambda} - 1)\frac{u}{\lambda(1+\delta)}}}{e^{\lambda(1+\delta)}}\right) \leq \left(\frac{e^{1+\delta}}{(1+\delta)^{1+\delta}}\right)$$

minimized when $\lambda = \ln(1 + \delta)$

$$= \mathbb{E}(e^{\lambda \sum_{i=1}^{n} X_i}) = \mathbb{E}\left(\prod_{i=1}^{n} e^{\lambda X_i}\right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i})$$

not linearity of expectation

indeedence of $X_i$

(b) bound the value of the moment generating function

$$\leq \prod_{i=1}^{n} e^{p_i(e^{\lambda} - 1)} = e^{\mu(e^{\lambda} - 1)}$$

assume $\mathbb{P}(X_i = 1) = p_i$, then $\mu = \sum_{i=1}^{n} p_i$

$$= \mathbb{P}(X_i = 1) \cdot e^{\lambda} + \mathbb{P}(X_i = 0) \cdot e^{0} = p_i e^{\lambda} + (1 - p_i) = 1 + p_i(e^{\lambda} - 1) \leq e^{p_i(e^{\lambda} - 1)} \quad 1 + y \leq e^{y}$$
independent \( X_1, X_2, \cdots, X_n \in \{0, 1\} \), let \( X = \sum_{i=1}^{n} X_i \), and \( \mu = \mathbb{E}(X) \)

\[ \mathbb{P}(X \geq (1 + \delta)\mu) \leq? \quad \text{for} \quad \delta > 0 \]

\[ \mathbb{P}(X \geq (1 + \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1+\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1+\delta)\mu}} \]

\( \lambda > 0 \), and generalized Markov’s inequality

\( \lambda = \ln(1 + \delta) \)

\( \text{(a) apply Markov’s inequality to moment generating function} \)

\[ \leq \frac{e^{u(e^{\lambda}-1)}}{e^{u\lambda(1+\delta)}} = \left( \frac{e^{(e^{\lambda}-1)}}{e^{\lambda(1+\delta)}} \right) \leq \left( \frac{e^{\lambda}}{(1 + \delta)^{1+\delta}} \right) \]

\( \text{minimized when} \)

\( c = \ln(1 + \delta) \)

\( \text{(c) optimize the bound of the moment generating function} \)

\[ = \mathbb{E}(e^{\lambda \sum_{i=1}^{n} X_i}) = \mathbb{E}\left( \prod_{i=1}^{n} e^{\lambda X_i} \right) = \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i}) \]

\( \text{independence of} \ X_i \)

\( \text{not linearity of expectation} \)

\[ \leq \prod_{i=1}^{n} e^{p_i(e^{\lambda}-1)} = e^{\mu(e^{\lambda}-1)} \]

\( \text{(b) bound the value of the moment generating function} \)

assume \( \mathbb{P}(X_i = 1) = p_i \), then \( \mu = \sum_{i=1}^{n} p_i \)

\[ = \mathbb{P}(X_i = 1) \cdot e^{\lambda} + \mathbb{P}(X_i = 0) \cdot e^{0} = p_i e^{\lambda} + (1 - p_i) = 1 + p_i(e^{\lambda} - 1) \leq e^{p_i(e^{\lambda} - 1)} \]

\[ 1 + y \leq e^{y} \]
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 $0 < \delta < 1$,

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

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 $\delta > 0$,

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

for $0 < \delta < 1$,

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

for any $\lambda < 0$,

$$\mathbb{P}(X \leq (1 - \delta)\mu) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(1-\delta)\mu}) \leq \frac{\mathbb{E}(e^{\lambda X})}{e^{\lambda(1-\delta)\mu}} \leq \cdots$$
Hoeffding’s Inequality

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 $t > 0$,

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

$$\mathbb{P}(X \leq \mathbb{E}(X) - t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n} (b_i-a_i)^2}\right)$$
For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in \{0, 1\}$, let $X = \sum_{i=1}^{n} X_i$, then:

for any $t > 0,$

\[
P(X \geq \mathbb{E}(X) + t) \leq \exp \left( -\frac{2t^2}{n} \right)
\]

\[
P(X \leq \mathbb{E}(X) - t) \leq \exp \left( -\frac{2t^2}{n} \right)
\]
Hoeffding’s Lemma

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

\[ \mathbb{E}(e^{\lambda X}) \leq \exp \left( \frac{\lambda^2 (b-a)^2}{8} \right) \]
independent $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$

$\Pr(X \geq \mathbb{E}(X) + t) \leq ?$ for $t > 0$
independent $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$

$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \ ?$  for  $t > 0$

let $Y_i = X_i - \mathbb{E}(X_i)$, and $Y = \sum_{i=1}^{n} Y_i = X - \mathbb{E}(X)$
independent $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$

$P(X \geq \mathbb{E}(X) + t) \leq ?$ for $t > 0$

let $Y_i = X_i - \mathbb{E}(X_i)$, and $Y = \sum_{i=1}^{n} Y_i = X - \mathbb{E}(X)$ thus $\mathbb{E}(Y_i) = 0$, and $\mathbb{E}(Y) = 0$
independent $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$

$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \frac{\mathbb{E}(e^{\lambda Y})}{e^{\lambda t}}$ for $t > 0$

let $Y_i = X_i - \mathbb{E}(X_i)$, and $Y = \sum_{i=1}^{n} Y_i = X - \mathbb{E}(X)$ thus $\mathbb{E}(Y_i) = 0$, and $\mathbb{E}(Y) = 0$

$\mathbb{P}(X \geq \mathbb{E}(X) + t) = \mathbb{P}(Y \geq t) = \mathbb{P}(e^{\lambda Y} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda Y})}{e^{\lambda t}}$ \hspace{1cm} $\lambda > 0$, and
generalized Markov's inequality
independent \( X_1, X_2, \cdots, X_n \) where \( X_i \in [a_i, b_i] \), let \( X = \sum_{i=1}^{n} X_i \)

\[
\Pr(X \geq \mathbb{E}(X) + t) \leq \quad \text{for } \quad t > 0
\]

let \( Y_i = X_i - \mathbb{E}(X_i) \), and \( Y = \sum_{i=1}^{n} Y_i = X - \mathbb{E}(X) \)  

thus \( \mathbb{E}(Y_i) = 0 \), and \( \mathbb{E}(Y) = 0 \)

\[
\Pr(X \geq \mathbb{E}(X) + t) = \Pr(Y \geq t) = \Pr(e^{\lambda Y} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda Y})}{e^{\lambda t}} \quad \lambda > 0, \text{ and}
\]

generalized Markov’s inequality

\[
= e^{-\lambda t} \mathbb{E} \left( \prod_{i=1}^{n} e^{\lambda Y_i} \right) = e^{-\lambda t} \prod_{i=1}^{n} \mathbb{E}(e^{\lambda Y_i})
\]

indepedence of \( Y_i \)
independent $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$

$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \frac{1}{e^{\lambda t}}$ for $t > 0$

let $Y_i = X_i - \mathbb{E}(X_i)$, and $Y = \sum_{i=1}^{n} Y_i = X - \mathbb{E}(X)$ thus $\mathbb{E}(Y_i) = 0$, and $\mathbb{E}(Y) = 0$

$\mathbb{P}(X \geq \mathbb{E}(X) + t) = \mathbb{P}(Y \geq t) = \mathbb{P}(e^{\lambda Y} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda Y})}{e^{\lambda t}}$ \hspace{1cm} $\lambda > 0$, and
generalized Markov’s inequality

$= e^{-\lambda t} \mathbb{E} \left( \prod_{i=1}^{n} e^{\lambda Y_i} \right) = e^{-\lambda t} \prod_{i=1}^{n} \mathbb{E}(e^{\lambda Y_i})$ \hspace{1cm} independence of $Y_i$

$\leq e^{-\lambda t} \prod_{i=1}^{n} \exp \left( \frac{\lambda^2(b_i-a_i)^2}{8} \right)$ \hspace{1cm} Hoeffding’s lemma
independent $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$

$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq? \text{ for } t > 0$

let $Y_i = X_i - \mathbb{E}(X_i)$, and $Y = \sum_{i=1}^{n} Y_i = X - \mathbb{E}(X)$ thus $\mathbb{E}(Y_i) = 0$, and $\mathbb{E}(Y) = 0$

$\mathbb{P}(X \geq \mathbb{E}(X) + t) = \mathbb{P}(Y \geq t) = \mathbb{P}(e^{\lambda Y} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda Y})}{e^{\lambda t}}$ $\lambda > 0$, and generalized Markov’s inequality

$= e^{-\lambda t} \mathbb{E} \left( \prod_{i=1}^{n} e^{\lambda Y_i} \right) = e^{-\lambda t} \prod_{i=1}^{n} \mathbb{E}(e^{\lambda Y_i})$ independence of $Y_i$

$\leq e^{-\lambda t} \prod_{i=1}^{n} \exp \left( \frac{\lambda^2(b_i-a_i)^2}{8} \right)$ Hoeffding’s lemma

$= \exp \left( -\lambda t + \frac{\lambda^2}{8} \sum_{i=1}^{n} (b_i - a_i)^2 \right)$
independent $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^{n} X_i$

$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \frac{4t}{\lambda}$ for $t > 0$

let $Y_i = X_i - \mathbb{E}(X_i)$, and $Y = \sum_{i=1}^{n} Y_i = X - \mathbb{E}(X)$ thus $\mathbb{E}(Y_i) = 0$, and $\mathbb{E}(Y) = 0$

$\mathbb{P}(X \geq \mathbb{E}(X) + t) = \mathbb{P}(Y \geq t) = \mathbb{P}(e^{\lambda Y} \geq e^{\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda Y})}{e^{\lambda t}}$ for $\lambda > 0$, and generalized Markov’s inequality

$= e^{-\lambda t} \mathbb{E} \left( \prod_{i=1}^{n} e^{\lambda Y_i} \right) = e^{-\lambda t} \prod_{i=1}^{n} \mathbb{E}(e^{\lambda Y_i})$ independence of $Y_i$

$\leq e^{-\lambda t} \prod_{i=1}^{n} \exp \left( \frac{\lambda^2(b_i-a_i)^2}{8} \right)$ Hoeffding’s lemma

$= \exp \left( -\lambda t + \frac{\lambda^2}{8} \sum_{i=1}^{n} (b_i - a_i)^2 \right)$

$\leq \exp \left( -\frac{2t^2}{\sum_{i=1}^{n} (b_i-a_i)^2} \right)$ minimized when $\lambda = \frac{4t}{\sum_{i=1}^{n} (b_i-a_i)^2}$
Hoeffding’s Inequality

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 $t > 0$,
\[
\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \exp \left( - \sum_{i=1}^{n} \frac{2t^2}{(b_i - a_i)^2} \right)
\]
\[
\mathbb{P}(X \leq \mathbb{E}(X) - t) \leq \exp \left( - \sum_{i=1}^{n} \frac{2t^2}{(b_i - a_i)^2} \right)
\]
Hoeffding’s Inequality

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 $t > 0$,

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

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

for any $\lambda < 0$,

$$
\mathbb{P}(X \leq \mathbb{E}(X) - t) = \mathbb{P}(Y \leq -t) = \mathbb{P}(e^{\lambda Y} \geq e^{-\lambda t}) \leq \frac{\mathbb{E}(e^{\lambda Y})}{e^{-\lambda t}} \leq \cdots
$$
Hoeffding’s Inequality in Action: Randomized Quicksort

sort $n$ distinct elements using QuickSort

choose pivot uniformly at random in each recursive call of QuickSort
Hoeffding’s Inequality in Action: Randomized Quicksort

sort $n$ distinct elements using QuickSort

choose pivot uniformly at random in each recursive call of QuickSort

expected cost under adversarial input: $\Theta(n \log n)$
Hoeffding’s Inequality in Action: Randomized Quicksort

Sort $n$ distinct elements using QuickSort

Choose pivot uniformly at random in each recursive call of QuickSort

Expected cost under adversarial input:
$\Theta(n \log n)$

Worst case cost under any input:
$\Theta(n^2)$
Hoeffding’s Inequality in Action: Randomized Quicksort

sort \( n \) distinct elements using QuickSort

choose pivot \textit{uniformly at random} in each recursive call of QuickSort

expected cost under \textit{adversarial} input: \( \Theta(n \log n) \)

worst case cost under \textit{any} input: \( \Theta(n^2) \)

Question: probability that cost is \( \omega(n \log n) \)?
Hoeffding’s Inequality in Action: Randomized Quicksort

Question: probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \log n$ with high probability.
Hoeffding’s Inequality in Action: Randomized Quicksort

Question: probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \log n$ with high probability.

Let $P_i$ be a path from root to the $i^{th}$ leaf, there are $n' \leq n$ such paths.
Hoeffding’s Inequality in Action: Randomized Quicksort

Question: probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \log n$ with high probability.

Let $P_i$ be a path from root to the $i^{th}$ leaf, there are $n' \leq n$ such paths

Cost of the algorithm is at most:
$n \cdot \max\{|P_i|\}$
Hoeffding’s Inequality in Action: Randomized Quicksort

Question: probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \log n$ with high probability.

Let $P_i$ be a path from root to the $i^{th}$ leaf, there are $n' \leq n$ such paths.

Cost of the algorithm is at most:

$n \cdot \max\{|P_i|\}$
Hoeffding’s Inequality in Action: Randomized Quicksort

Question: probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \lg n$ with high probability.

\[
\mathbb{P}(\exists P_i : |P_i| \geq 30 \log n) \leq \frac{1}{n} \iff \mathbb{P}(|P_i| \geq 30 \log n) \leq \frac{1}{n^2}
\]

Let $P_i$ be a path from root to the $i^{th}$ leaf, there are $n' \leq n$ such paths.

Cost of the algorithm is at most:

\[n \cdot \max\{|P_i|\}\]
Hoeffding’s Inequality in Action: Randomized Quicksort

Question: probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \lg n$ with high probability.

Let $P_i$ be a path from root to the $i^{th}$ leaf, there are $n' \leq n$ such paths

Cost of the algorithm is at most:

$n \cdot \max\{|P_i|\}$
Hoeffding’s Inequality in Action: Randomized Quicksort

Question: probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \lg n$ with high probability.

\[
\mathbb{P}(\exists P_i : |P_i| \geq 30 \lg n) \leq \frac{1}{n} \iff \mathbb{P}(|P_i| \geq 30 \lg n) \leq \frac{1}{n^2}
\]

for the $j^{th}$ node in $P_i$, call it *good* if it partitions length $l$ array into two parts each of length $\geq l/3$

call it *bad* otherwise

Let $P_i$ be a path from root to the $i^{th}$ leaf, there are $n' \leq n$ such paths

Cost of the algorithm is at most:

$n \cdot \max\{|P_i|\}$
Hoeffding’s Inequality in Action: Randomized Quicksort

Question: probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \log n$ with high probability.

$$\mathbb{P}(\exists P_i : |P_i| \geq 30 \log n) \leq \frac{1}{n} \iff \mathbb{P}(|P_i| \geq 30 \log n) \leq \frac{1}{n^2}$$

for the $j^{\text{th}}$ node in $P_i$, call it good if it partitions length $l$ array into two parts each of length $\geq l/3$
call it bad otherwise

let $X_{ij}$ be i.r.v. taking value 1 iff the $j^{\text{th}}$ node in $P_i$ is bad
let $X_i = \sum_{j=1}^{|P_i|} X_{ij}$, we know $\mathbb{E}(X_i) = \frac{2}{3}|P_i|$

Let $P_i$ be a path from root to the $i^{\text{th}}$ leaf, there are $n' \leq n$ such paths

Cost of the algorithm is at most:
$$n \cdot \max\{|P_i|\}$$
Hoeffding’s Inequality in Action: Randomized Quicksort

**Question:** probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \log n$ with high probability.

\[
\mathbb{P}(\exists P_i : |P_i| \geq 30 \log n) \leq \frac{1}{n} \iff \mathbb{P}(|P_i| \geq 30 \log n) \leq \frac{1}{n^2}
\]

for the $j^{th}$ node in $P_i$, call it *good* if it partitions length $l$ array into two parts each of length $\geq l/3$
call it *bad* otherwise

let $X_{ij}$ be i.r.v. taking value 1 iff the $j^{th}$ node in $P_i$ is bad
let $X_i = \sum_{j=1}^{|P_i|} X_{ij}$, we know $\mathbb{E}(X_i) = \frac{2}{3} |P_i|$

we know there are at most $2 \log n$ good nodes in $P_i$

Let $P_i$ be a path from root to the $i^{th}$ leaf, there are $n' \leq n$ such paths

Cost of the algorithm is at most:

\[n \cdot \max\{|P_i|\}\]
Hoeffding’s Inequality in Action: Randomized Quicksort

Question: probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \lg n$ with high probability.

\[
\mathbb{P}(\exists P_i : |P_i| \geq 30 \lg n) \leq \frac{1}{n} \iff \mathbb{P}(|P_i| \geq 30 \lg n) \leq \frac{1}{n^2}
\]

for the $j^{th}$ node in $P_i$,
call it good if it partitions length $l$ array into two parts each of length $\geq l/3$
call it bad otherwise

let $X_{ij}$ be i.r.v. taking value 1 iff the $j^{th}$ node in $P_i$ is bad
let $X_i = \sum_{j=1}^{|P_i|} X_{ij}$, we know $\mathbb{E}(X_i) = \frac{2}{3} |P_i|

we know there are at most $2 \lg n$ good nodes in $P_i$

$\mathbb{P}(|P_i| = 30 \lg n) \leq \mathbb{P}(X_i \geq 28 \lg n) = \mathbb{P}(X_i > \mathbb{E}(X_i) + 8 \lg n)$

Let $P_i$ be a path from root to the $i^{th}$ leaf, there are $n' \leq n$ such paths

Cost of the algorithm is at most:

$n \cdot \max\{|P_i|\}$
Hoeffding’s Inequality in Action: Randomized Quicksort

**Question:** probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \lg n$ with high probability.

Let $P_i$ be a path from root to the $i^{th}$ leaf, there are $n' \leq n$ such paths.

Cost of the algorithm is at most:

$$n \cdot \max\{|P_i|\}$$

$$\mathbb{P}(\exists P_i : |P_i| \geq 30 \lg n) \leq \frac{1}{n} \iff \mathbb{P}(|P_i| \geq 30 \lg n) \leq \frac{1}{n^2}$$

for the $j^{th}$ node in $P_i$,
call it *good* if it partitions length $l$ array into two parts each of length $\geq l/3$
call it *bad* otherwise

let $X_{ij}$ be i.r.v. taking value 1 iff the $j^{th}$ node in $P_i$ is bad
let $X_i = \sum_{j=1}^{|P_i|} X_{ij}$, we know $\mathbb{E}(X_i) = \frac{2}{3} |P_i|$

we know there are at most $2 \lg n$ good nodes in $P_i$

$$\mathbb{P}(|P_i| = 30 \lg n) \leq \mathbb{P}(X_i \geq 28 \lg n) = \mathbb{P}(X_i > \mathbb{E}(X_i) + 8 \lg n) \leq \exp \left(- \frac{2(8 \lg n)^2}{30 \lg n} \right) \leq \frac{1}{n^3}$$
**Hoeffding’s Inequality in Action: Randomized Quicksort**

**Question:** probability that cost is $\omega(n \log n)$?

The cost will be at most $30n \lg n$ with high probability.

Let $P_i$ be a path from root to the $i$th leaf, there are $n' \leq n$ such paths.

Cost of the algorithm is at most:

$$n \cdot \max\{|P_i|\}$$

$$\mathbb{P}(\exists P_i : |P_i| \geq 30 \lg n) \leq \frac{1}{n} \iff \mathbb{P}(|P_i| \geq 30 \lg n) \leq \frac{1}{n^2}$$

for the $j$th node in $P_i$,
call it *good* if it partitions length $l$ array into two parts each of length $\geq l/3$
call it *bad* otherwise

let $X_{ij}$ be i.r.v. taking value 1 iff the $j$th node in $P_i$ is bad
let $X_i = \sum_{j=1}^{|P_i|} X_{ij}$, we know $\mathbb{E}(X_i) = \frac{2}{3} |P_i|$

we know there are at most $2 \lg n$ good nodes in $P_i$

$$\mathbb{P}(|P_i| = 30 \lg n) \leq \mathbb{P}(X_i \geq 28 \lg n) = \mathbb{P}(X_i > \mathbb{E}(X_i) + 8 \lg n) \leq \exp\left(-\frac{2(8 \lg n)^2}{30 \lg n}\right) \leq \frac{1}{n^3}$$

$$\mathbb{P}(|P_i| \geq 30 \lg n) \leq n \cdot \mathbb{P}(|P_i| = 30 \lg n) \leq \frac{1}{n^2}$$
Question: probability that $X$ deviates more than $\delta$ from expectation?
(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 $t > 0$, 
$$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \exp \left( - \sum_{i=1}^{n} \frac{2t^2}{(b_i-a_i)^2} \right)$$

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

“$m$ balls are thrown into $n$ bins uniformly and independently at random”

Question: maximum number of balls in a bin?

When $m = \Theta(n)$:

the max load is $O \left( \frac{\log n}{\log \log n} \right)$ with high probability.
200 Balls into 200 Bins

previous strategy

new strategy
“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

Question: maximum number of balls in a bin?

When $m = n$:
the max load is $O(\log \log n)$ with high probability.
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

Question: maximum number of balls in a bin?

When $m = n$:

the max load is $O(\log \log n)$ with high probability.

$O\left(\frac{\log n}{\log \log n}\right)$ versus $O(\log \log n)$, exponential gap
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has $O(\log \log n)$ balls, w.h.p.
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has $O(\log \log n)$ balls, w.h.p.

Assume there are at most $\beta_i$ bins each containing at least $i$ balls in the end
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has $O(\log \log n)$ balls, w.h.p.

Assume there are at most $\beta_i$ bins each containing at least $i$ balls in the end

Probability that a ball increases # of bins containing at least $i + 1$ balls?
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has $O(\log \log n)$ balls, w.h.p.

Assume there are at most $\beta_i$ bins each containing at least $i$ balls in the end

Probability that a ball increases # of bins containing at least $i + 1$ balls? at most $\left(\frac{\beta_i}{n}\right)^2$
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has $O(\log \log n)$ balls, w.h.p.

Assume there are at most $\beta_i$ bins each containing at least $i$ balls in the end

Probability that a ball increases # of bins containing at least $i + 1$ balls? at most $\left(\frac{\beta_i}{n}\right)^2$

In expectation, after $n$ balls, # of bins containing at least $i + 1$ balls is
Power of Two Choices

“\(m\) balls are thrown into \(n\) bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has \(O(\log \log n)\) balls, w.h.p.

Assume there are at most \(\beta_i\) bins each containing at least \(i\) balls in the end

Probability that a ball increases \# of bins containing at least \(i+1\) balls? at most \(\left(\frac{\beta_i}{n}\right)^2\)

In expectation, after \(n\) balls, \# of bins containing at least \(i+1\) balls is at most \(n \left(\frac{\beta_i}{n}\right)^2 = \frac{\beta_i^2}{n}\)
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has $O(\log \log n)$ balls, w.h.p.

Assume there are at most $\beta_i$ bins each containing at least $i$ balls in the end

Probability that a ball increases # of bins containing at least $i + 1$ balls? at most $\left( \frac{\beta_i}{n} \right)^2$

In expectation, after $n$ balls, # of bins containing at least $i + 1$ balls is at most $n \left( \frac{\beta_i}{n} \right)^2 = \frac{\beta_i^2}{n}$

$\beta_{i+1} = \frac{\beta_i^2}{n}$
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has $O(\log \log n)$ balls, w.h.p.

Assume there are at most $\beta_i$ bins each containing at least $i$ balls in the end

Probability that a ball increases # of bins containing at least $i + 1$ balls? at most $\left(\frac{\beta_i}{n}\right)^2$

In expectation, after $n$ balls, # of bins containing at least $i + 1$ balls is at most $n \left(\frac{\beta_i}{n}\right)^2 = \frac{\beta_i^2}{n}$

$$\beta_{i+1} = \frac{\beta_i^2}{n}$$

$$\beta_4 \leq \frac{n}{4}$$
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has $O(\log \log n)$ balls, w.h.p.

Assume there are at most $\beta_i$ bins each containing at least $i$ balls in the end

Probability that a ball increases # of bins containing at least $i + 1$ balls? at most $\left( \frac{\beta_i}{n} \right)^2$

In expectation, after $n$ balls, # of bins containing at least $i + 1$ balls is at most $n \left( \frac{\beta_i}{n} \right)^2 = \frac{\beta_i^2}{n}$

$\beta_{i+1} = \frac{\beta_i^2}{n} \implies \beta_i \leq \frac{n}{(4)^{2^{i-4}}}$

$\beta_4 \leq \frac{n}{4}$
Power of Two Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has $O(\log \log n)$ balls, w.h.p.

Assume there are at most $\beta_i$ bins each containing at least $i$ balls in the end

Probability that a ball increases # of bins containing at least $i + 1$ balls? at most $\left(\frac{\beta_i}{n}\right)^2$

In expectation, after $n$ balls, # of bins containing at least $i + 1$ balls is at most $n \left(\frac{\beta_i}{n}\right)^2 = \frac{\beta_i^2}{n}$

$\beta_{i+1} = \frac{\beta_i^2}{n}$

$\Rightarrow \beta_i \leq \frac{n}{(4)^{2^{i-4}}}$

$\Rightarrow \beta_i \leq 1$ when $i \geq \lg \lg n$

$\beta_4 \leq \frac{n}{4}$
Power of Two Choices

“\( m \) balls are thrown into \( n \) bins in the following manner: for each ball, choose two bins uniformly and independently at random, then place the ball in the less loaded bin”

the max loaded bin has \( O(\log \log n) \) balls, w.h.p.

Assume there are at most \( \beta_i \) bins each containing at least \( i \) balls in the end

Probability that a ball increases # of bins containing at least \( i + 1 \) balls? at most \( \left( \frac{\beta_i}{n} \right)^2 \)

In expectation, after \( n \) balls, # of bins containing at least \( i + 1 \) balls is at most \( n \left( \frac{\beta_i}{n} \right)^2 = \frac{\beta_i^2}{n} \)

\[
\beta_{i+1} = \frac{\beta_i^2}{n} \quad \Rightarrow \quad \beta_i \leq \frac{n}{(4)^{2^{i-4}}} \quad \Rightarrow \quad \beta_i \leq 1 \text{ when } i \geq \lg \lg n
\]
Power of $d$ Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose $d \geq 2$ bins uniformly and independently at random, then place the ball in the least loaded bin”
Power of $d$ Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose $d \geq 2$ bins uniformly and independently at random, then place the ball in the least loaded bin”

the max loaded bin has $O(\log^{(d)} n)$ balls, w.h.p.
Power of $d$ Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose $d \geq 2$ bins uniformly and independently at random, then place the ball in the least loaded bin”

the max loaded bin has $O(\log^d n)$ balls, w.h.p.
Power of $d$ Choices

“$m$ balls are thrown into $n$ bins in the following manner: for each ball, choose $d \geq 2$ bins uniformly and independently at random, then place the ball in the least loaded bin”

the max loaded bin has $O(\log^d n)$ balls, w.h.p.

the max loaded bin has $O\left(\frac{\log \log n}{\log d}\right)$ balls, w.h.p.