

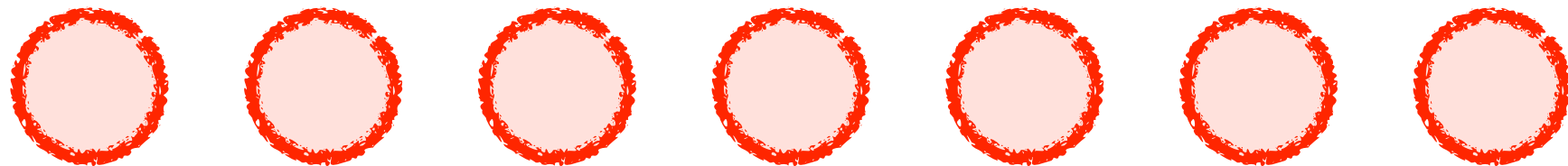
Advanced Algorithms

Concentration of Measure

尹一通 Nanjing University, 2026 Fall

Balls into Bins

(Coupon Collector)



uniform & independent

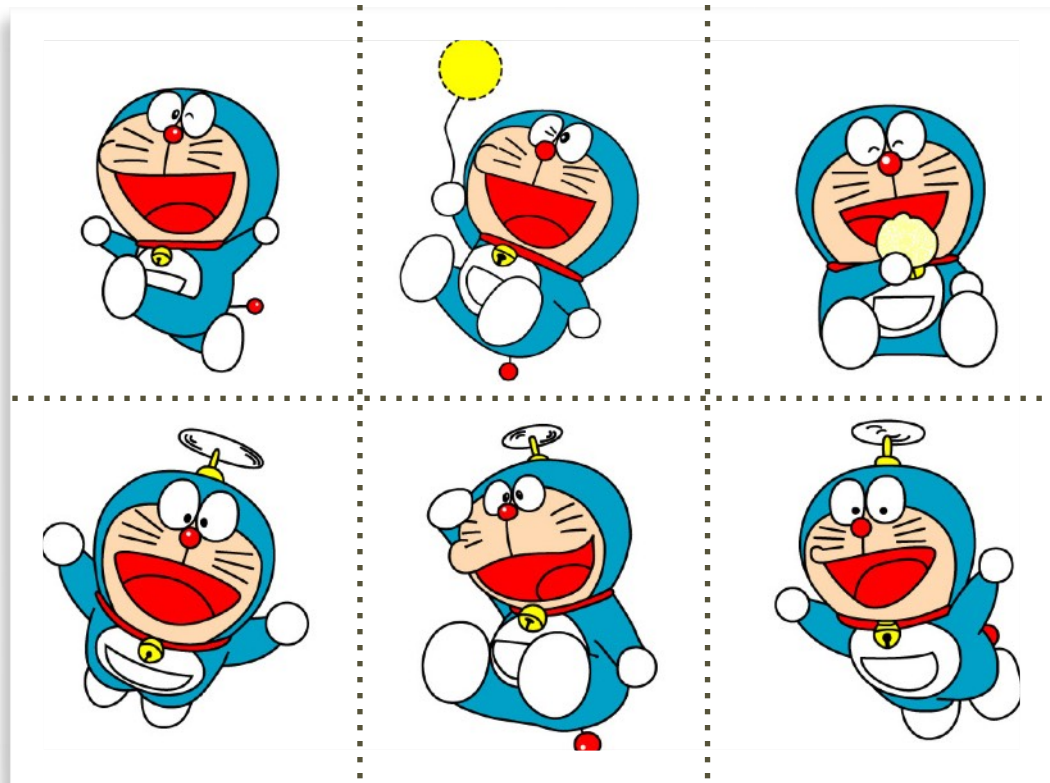
n bins



surjection (cover all bins)

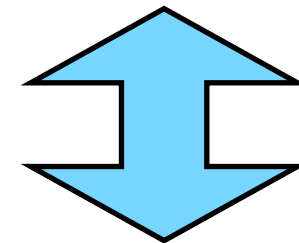
Coupon Collector

coupons in cookie box



each box comes with a
uniformly random coupon

number of boxes bought
to collect all n coupons

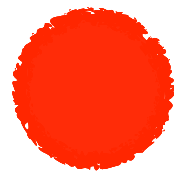


number of balls thrown to
cover all n bins

Coupon Collector

X : number of balls thrown to make
all the n bins nonempty

$$X = \sum_{i=1}^n X_i$$



$$X_i = 4$$

bins



$i - 1$

X_i is **geometric**!

with $p_i = 1 - \frac{i-1}{n}$

$$\mathbb{E}[X_i] = \frac{1}{p_i} = \frac{n}{n - i + 1}$$

Coupon Collector

X : number of balls thrown to make all the n bins nonempty

X_i : number of balls thrown while there are exactly $(i-1)$ nonempty bins

$$X = \sum_{i=1}^n X_i$$

$$\mathbb{E}[X_i] = \frac{1}{p_i} = \frac{n}{n - i + 1}$$

$$\mathbb{E}[X] = \sum_{i=1}^n \mathbb{E}[X_i] \quad \text{linearity of expectations}$$

$$= \sum_{i=1}^n \frac{n}{n - i + 1}$$

$$= n \sum_{i=1}^n \frac{1}{i}$$

$$= nH(n)$$

Harmonic number

expected $n \ln n + O(n)$ balls

Coupon Collector

X : number of balls
thrown to make all the
 n bins nonempty

Theorem: For $c > 0$,
 $\Pr[X \geq n \ln n + cn] \leq e^{-c}$

Proof: For one bin, it misses all balls with probability

$$\begin{aligned} \left(1 - \frac{1}{n}\right)^{n \ln n + cn} &= \left(1 - \frac{1}{n}\right)^{n(\ln n + c)} \\ &< e^{-(\ln n + c)} \\ &< \frac{1}{ne^c} \end{aligned}$$

Coupon Collector

X : number of balls
thrown to make all the
 n bins nonempty

Theorem: For $c > 0$,
 $\Pr[X \geq n \ln n + cn] \leq e^{-c}$

Proof: For one bin, it misses all balls with probability
$$< \frac{1}{ne^c}$$

union bound!

$$\begin{aligned} \Pr[\exists \text{ a bin misses all balls}] &\leq n \Pr[\text{first bin misses all bins}] \\ &< e^{-c} \end{aligned}$$

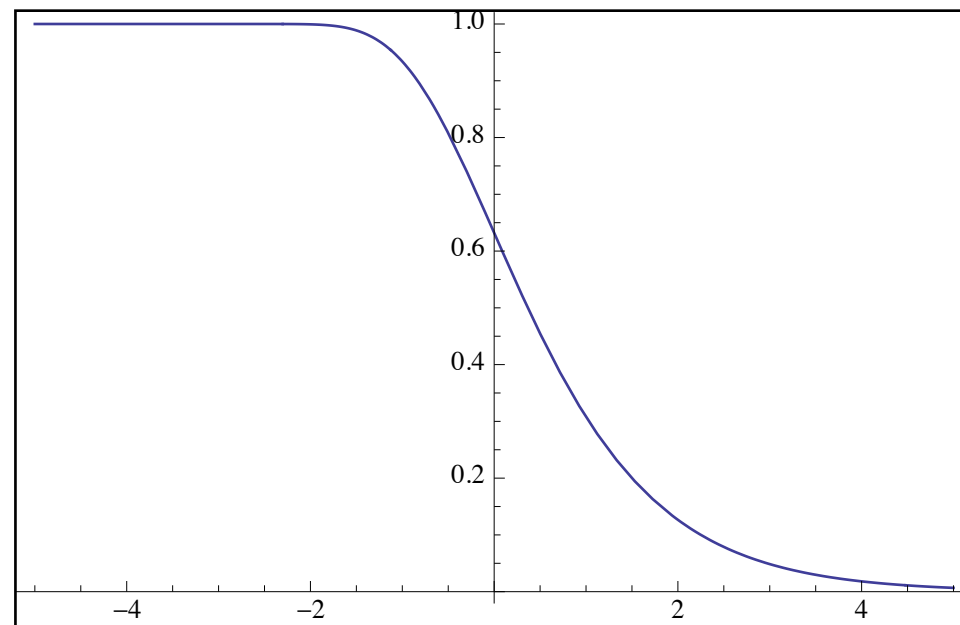
Coupon Collector

X : number of balls
thrown to make all the
 n bins nonempty

Theorem: For $c > 0$,
 $\Pr[X \geq n \ln n + cn] \leq e^{-c}$

a sharp threshold:

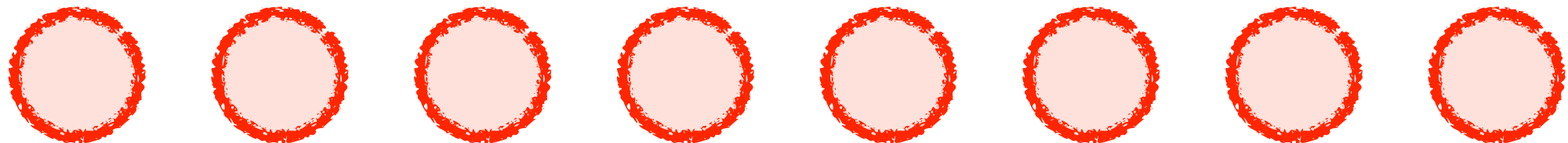
$$\lim_{n \rightarrow \infty} \Pr[X \geq n \ln n + cn] = 1 - e^{-e^{-c}}$$



Balls into Bins

(Occupancy Problem)

m balls



n bins



loads
of bins

X_1

X_2

X_3

• • • • •

X_n

maximum load?

Occupancy Problem

m balls are thrown into n bins.

X_i : number of balls in the i -th bin

$$\sum_{i=1}^n X_i = m \xrightarrow{\text{linearity of expectation}} \sum_{i=1}^n \mathbb{E}[X_i] = m$$

By symmetry: $X_1, \dots, X_n$ are identically distributed

$$\left. \begin{array}{l} \sum_{i=1}^n X_i = m \\ \sum_{i=1}^n \mathbb{E}[X_i] = m \\ X_1, \dots, X_n \text{ are identically distributed} \end{array} \right\} \forall i : \mathbb{E}[X_i] = \frac{m}{n}$$

$$\max_{1 \leq i \leq n} \mathbb{E}[X_i] = \frac{m}{n}$$

Occupancy Problem

m balls are thrown into n bins.

X_i : number of balls in the i -th bin

$$\max_{1 \leq i \leq n} \mathbb{E}[X_i] = \frac{m}{n}$$

Theorem: When $m = n$, the maximum load

$$\max_{1 \leq i \leq n} X_i = O\left(\frac{\log n}{\log \log n}\right) \text{ w.h.p.}$$

w.h.p. (with high probability): with probability $1 - O(1/n)$

m balls are thrown into n bins.

X_i : number of balls in the i -th bin

$$\Pr \left[\max_{1 \leq i \leq n} X_i \geq L \right] = \Pr \left[\exists 1 \leq i \leq n \text{ s.t. } X_i \geq L \right] \stackrel{\text{union bound}}{\leq} n \Pr [X_i \geq L]$$

$$\Pr [X_i \geq L] \leq \Pr [\exists L \text{ balls thrown into bin } i]$$

union bound

$$\leq \sum_{S \in \binom{[m]}{L}} \Pr [\text{all balls in } S \text{ are thrown into bin } i]$$

$$= \binom{m}{L} \frac{1}{n^L} \leq \frac{m^L}{L! n^L} \leq \left(\frac{em}{n} \right)^L \frac{1}{L^L}$$

Stirling's approximation: $L! \geq \left(\frac{L}{e} \right)^L$

m balls are thrown into n bins.

X_i : number of balls in the i -th bin

$$\Pr \left[\max_{1 \leq i \leq n} X_i \geq L \right] \leq n \Pr [X_i \geq L] \leq \frac{1}{n}$$

$$\Pr [X_i \geq L] \leq \left(\frac{em}{n} \right)^L \frac{1}{L^L}$$

$$\left\{ \begin{array}{ll} \text{(when } m = n) & \leq \left(\frac{e}{L} \right)^L \leq \frac{1}{n^2} \\ \text{(when } m \geq n \ln n) & \leq \left(\frac{L}{e} \right)^L \frac{1}{L^L} = e^{-L} \leq \frac{1}{n^2} \end{array} \right. \quad \text{for } L = \frac{3 \ln n}{\ln \ln n}$$

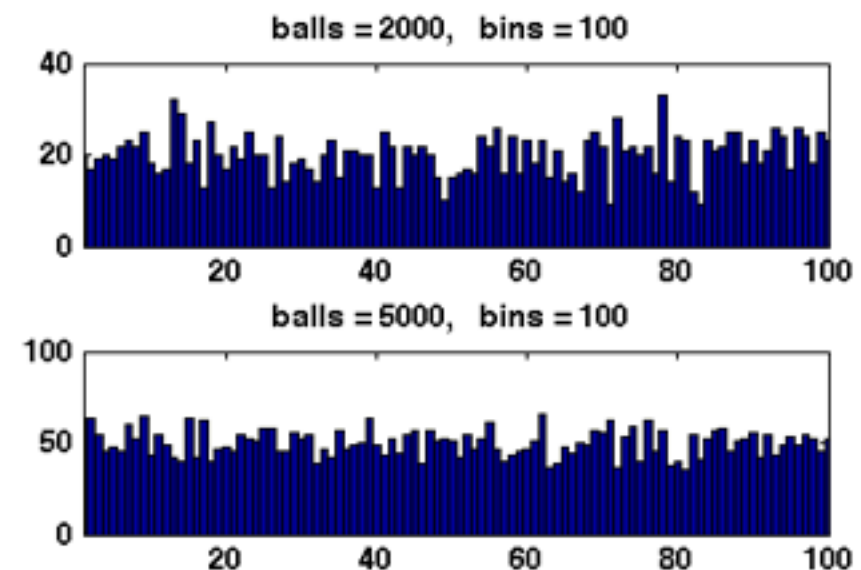
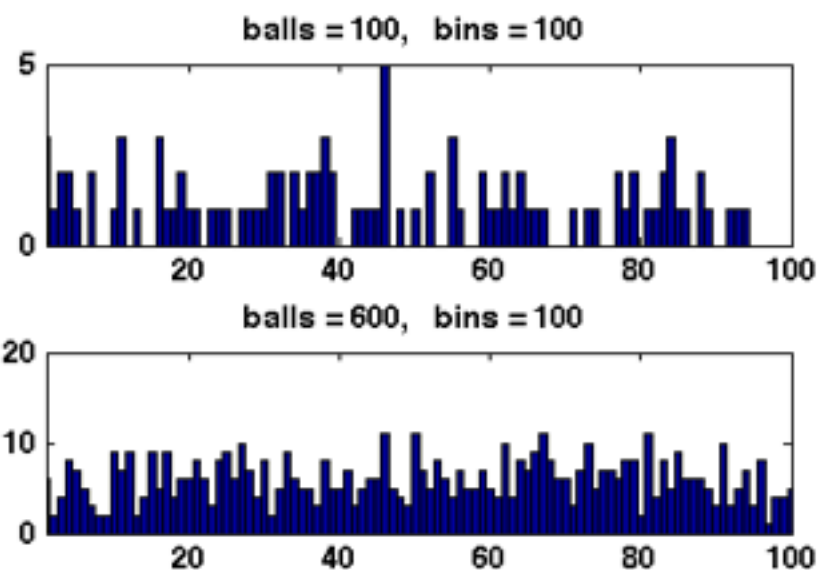
$$\text{for } L = \frac{e^2 m}{n} \geq e^2 \ln n$$

Occupancy Problem

- m balls are thrown into n bins uniformly and independently:

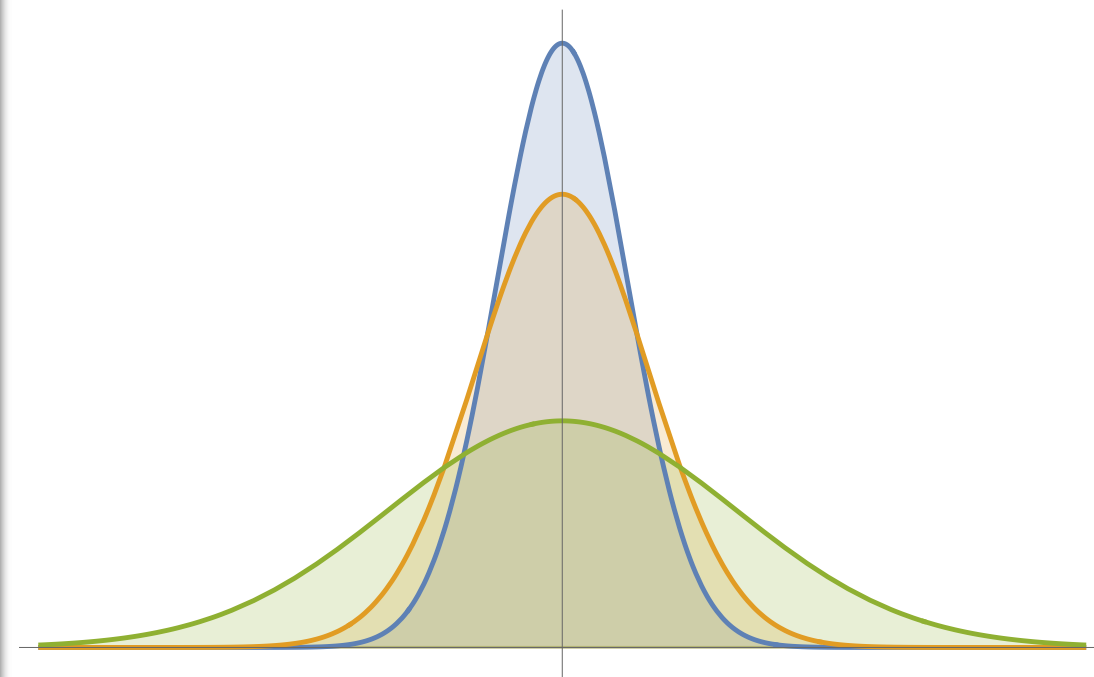
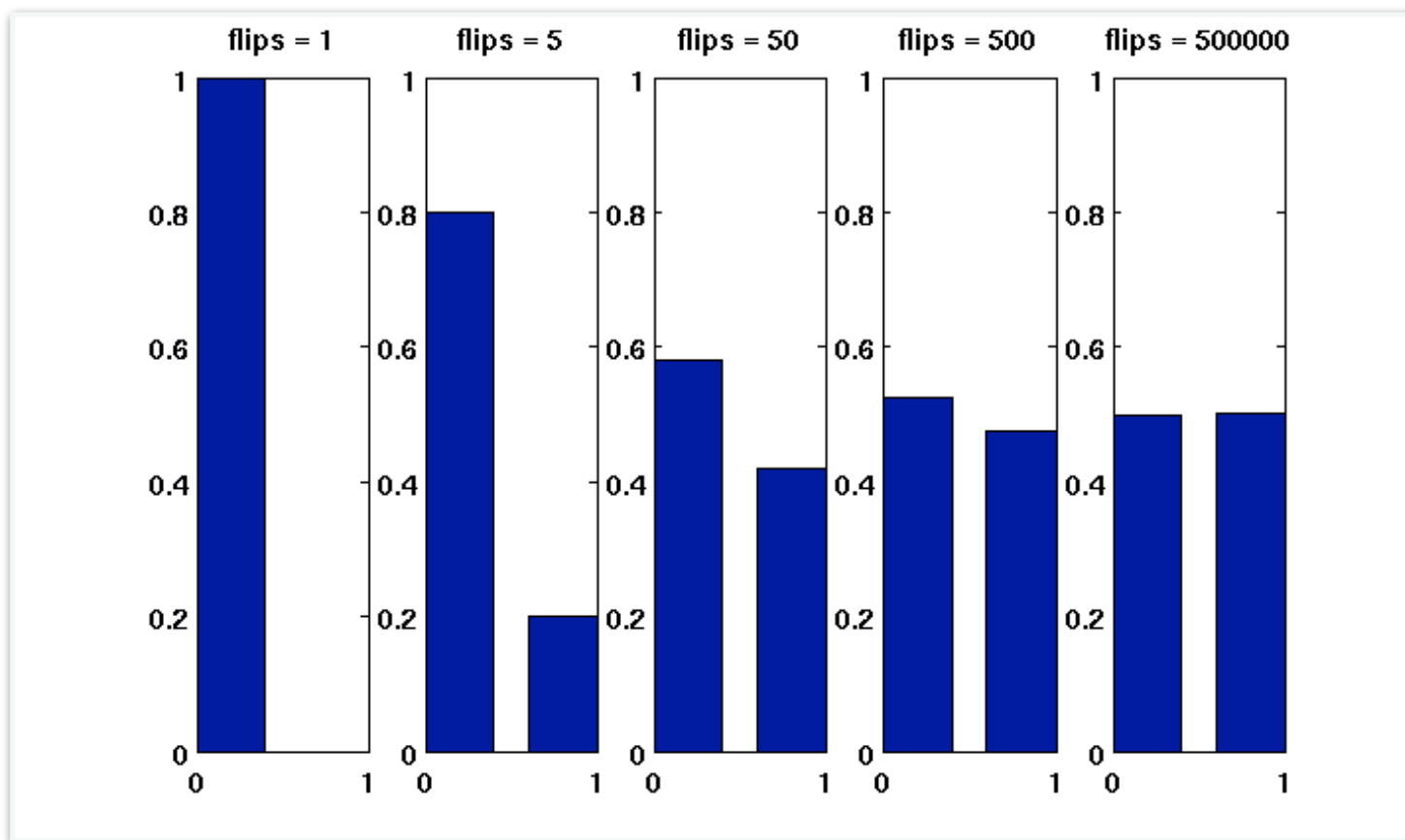
Theorem: With high probability, the maximum load is

$$\begin{cases} O\left(\frac{\log n}{\log \log n}\right) & \text{when } m = n \\ O\left(\frac{m}{n}\right) & \text{when } m \geq n \ln n \end{cases}$$



Measure Concentration

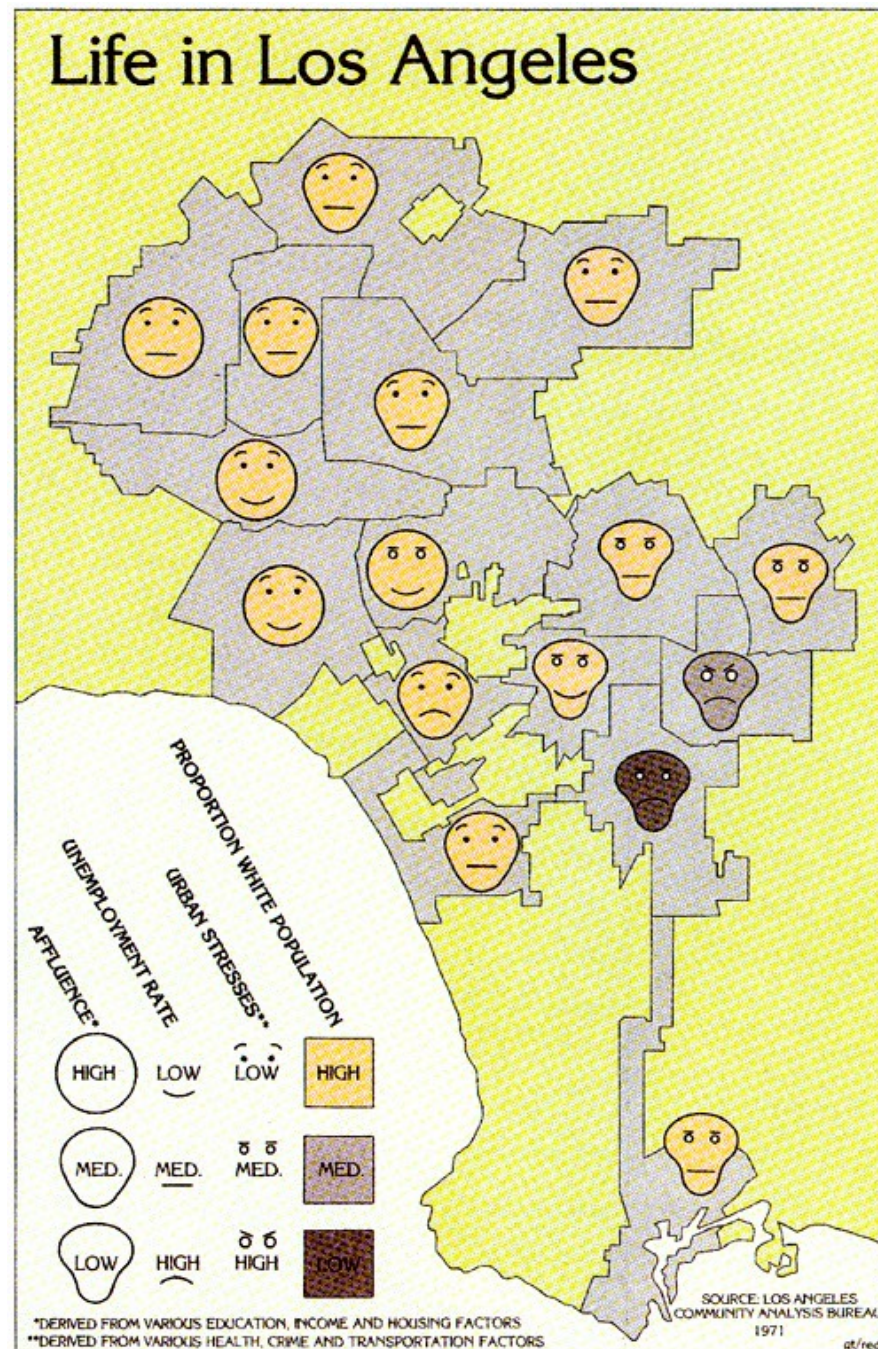
- Flip a coin for many times:



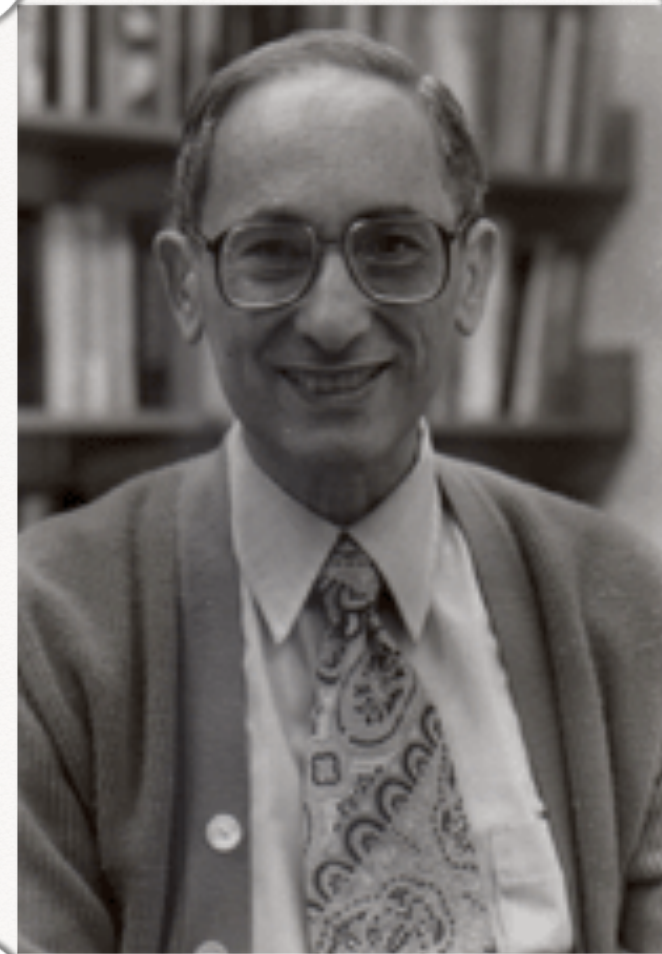
Chernoff-Hoeffding Bounds

Chernoff Bound

(Bernstein Inequalities)



Chernoff face



Herman Chernoff

Chernoff Bound

Chernoff Bound:

For *independent* $X_1, \dots, X_n \in \{0,1\}$ with

$$X = \sum_{i=1}^n X_i \text{ and } \mu = \mathbb{E}[X]$$

For any $\delta > 0$,

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

For any $0 < \delta < 1$,

$$\Pr[X \leq (1 - \delta)\mu] \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}} \right)^\mu$$

Chernoff Bound

Chernoff Bound:

For *independent* $X_1, \dots, X_n \in \{0,1\}$ with

$$X = \sum_{i=1}^n X_i \text{ and } \mu = \mathbb{E}[X]$$

For any $0 < \delta < 1$,

$$\Pr[X \geq (1 + \delta)\mu] \leq \exp\left(-\frac{\mu\delta^2}{3}\right)$$

$$\Pr[X \leq (1 - \delta)\mu] \leq \exp\left(-\frac{\mu\delta^2}{2}\right)$$

For $t \geq 2e\mu$:

$$\Pr[X \geq t] \leq 2^{-t}$$

Balls into Bins

m balls are thrown into n bins.

X_i : number of balls in the i -th bin

$$X_i = \sum_{j=1}^m X_{ij} \quad \text{where } X_{ij} = \begin{cases} 1 & \text{with prob. } \frac{1}{n} \\ 0 & \text{with prob. } 1 - \frac{1}{n} \end{cases}$$

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

Chernoff Bound: For $\delta > 0$,

$$\Pr [X_i \geq (1 + \delta)\mu] \leq \left(\frac{e^\delta}{(1 + \delta)^{(1+\delta)}} \right)^\mu$$

m balls are thrown into n bins.

X_i : number of balls in the i -th bin

$$\mu = \mathbb{E}[X_i] = \frac{m}{n}$$

Chernoff Bound: For $\delta > 0$,

$$\Pr [X_i \geq (1 + \delta)\mu] \leq \left(\frac{e^\delta}{(1 + \delta)^{(1+\delta)}} \right)^\mu$$

- When $m = n$: $\mu = 1$

$$\Pr [X_i \geq L] \leq \frac{e^L}{eL^L} \leq \frac{1}{n^2} \quad \text{for } L = \frac{e \ln n}{\ln \ln n}$$

- union bound:

$$\Pr \left[\max_{1 \leq i \leq n} X_i \geq L \right] \leq n \Pr [X_i \geq L] \leq \frac{1}{n}$$

Max load is $O\left(\frac{\log n}{\log \log n}\right)$ w.h.p.

m balls are thrown into n bins.

X_i : number of balls in the i -th bin

$$\mu = \mathbb{E}[X_i] = \frac{m}{n}$$

Chernoff Bound: For $L \geq 2e\mu$,
$$\Pr [X_i \geq L] \leq 2^{-L}$$

- When $m \geq n \ln n$: $\mu \geq \ln n$

$$\Pr \left[X_i \geq \frac{2em}{n} \right] = \Pr [X_i \geq 2e\mu] \leq 2^{-2e\mu} \leq 2^{-2e \ln n} \leq \frac{1}{n^2}$$

- union bound:

$$\Pr \left[\max_{1 \leq i \leq n} X_i \geq \frac{2em}{n} \right] \leq n \Pr \left[X_i \geq \frac{2em}{n} \right] \leq \frac{1}{n}$$

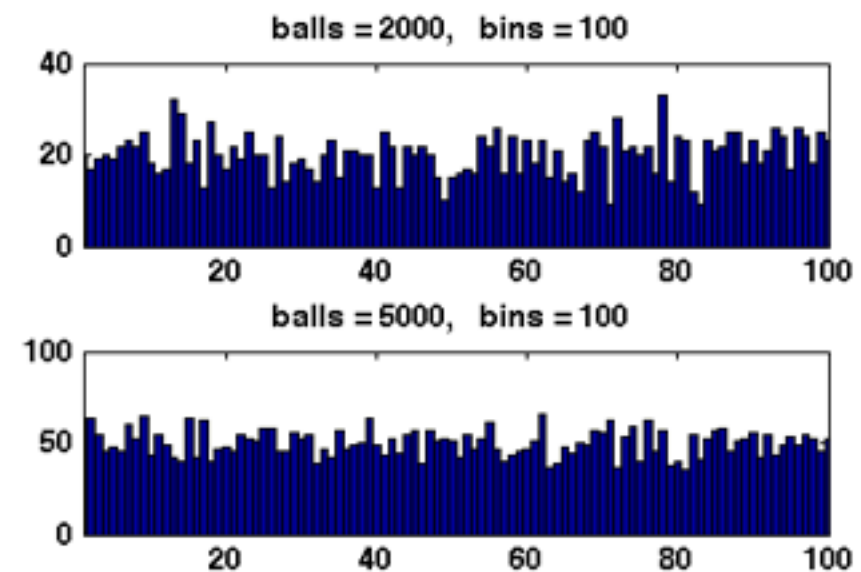
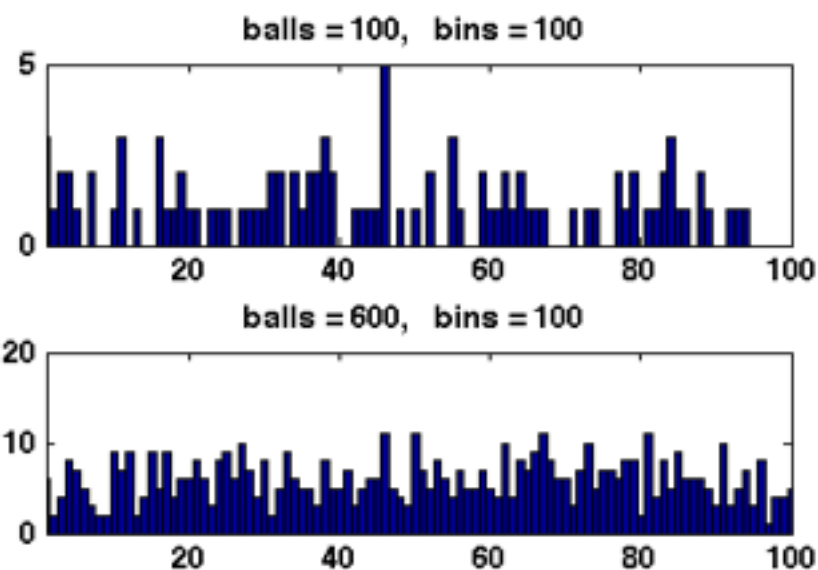
Max load is $O\left(\frac{m}{n}\right)$ w.h.p.

Balls into Bins

- m balls are thrown into n bins uniformly and independently:

Theorem: With high probability, the maximum load is

$$\begin{cases} O\left(\frac{\log n}{\log \log n}\right) & \text{when } m = n \\ O\left(\frac{m}{n}\right) & \text{when } m \geq n \ln n \end{cases}$$



Chernoff Bound

Chernoff Bound:

For *independent* $X_1, \dots, X_n \in \{0,1\}$ with

$$X = \sum_{i=1}^n X_i \text{ and } \mu = \mathbb{E}[X]$$

For any $\delta > 0$,

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

For any $0 < \delta < 1$,

$$\Pr[X \leq (1 - \delta)\mu] \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}} \right)^\mu$$

Markov's Inequality

Markov's Inequality

For *nonnegative* random variable X , for any $t > 0$,

$$\Pr[X \geq t] \leq \frac{\mathbb{E}[X]}{t}$$

Corollary

For random variable X and *nonnegative-valued* function f , for any $t > 0$,

$$\Pr[f(X) \geq t] \leq \frac{\mathbb{E}[f(X)]}{t}$$

Moment Generating Function

Moment generating function (MGF):

The MGF of a random variable X is defined as

$$M(\lambda) = \mathbb{E} \left[e^{\lambda X} \right].$$

- Taylor's expansion:

$$\mathbb{E} \left[e^{\lambda X} \right] = \mathbb{E} \left[\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} X^k \right] = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \mathbb{E} \left[X^k \right]$$

- Independent $X_1, \dots, X_n \in \{0,1\}$ with

$$X = \sum_{i=1}^n X_i \text{ and } \mu = \mathbb{E}[X]$$

- Markov for MGF:** (for any $\lambda > 0$) (Markov's inequality)

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

- Bound MGF:**

$$\mathbb{E}[e^{\lambda X}] = \mathbb{E}\left[\prod_{i=1}^n e^{\lambda X_i}\right] \stackrel{\text{(independence)}}{=} \prod_{i=1}^n \mathbb{E}[e^{\lambda X_i}] \leq \prod_{i=1}^n e^{(e^\lambda - 1)p_i} \stackrel{(\mu = \sum_{i=1}^n p_i)}{=} e^{(e^\lambda - 1)\mu}$$

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

(where $p_i = \Pr[X_i = 1]$)

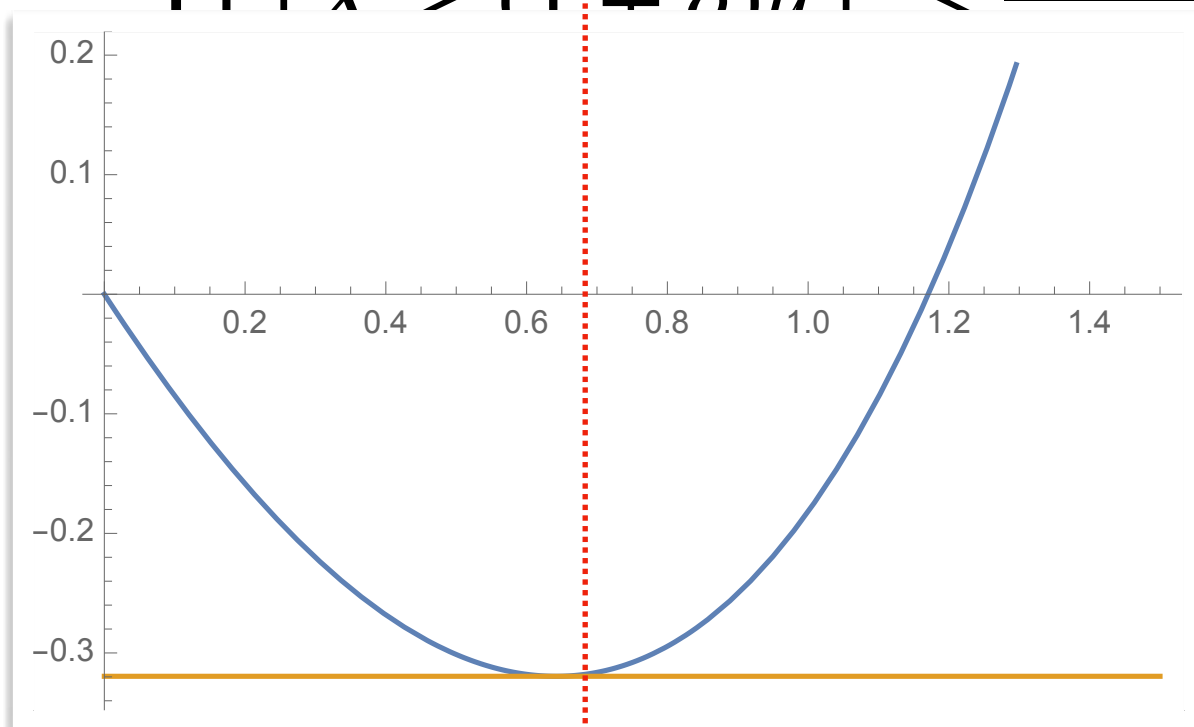
- Independent $X_1, \dots, X_n \in \{0,1\}$ with

$$X = \sum_{i=1}^n X_i \text{ and } \mu = \mathbb{E}[X]$$

- Markov for MGF:** (for any $\lambda > 0$)

$$\Pr[X > (1 + \delta)\mu] < \frac{\mathbb{E}[e^{\lambda X}]}{\delta^\lambda \mu^\lambda} \leq e^{(e^\lambda - 1 - \lambda(1 + \delta))\mu} = \left(\frac{e^\delta}{(1 + \delta)^{1 + \delta}} \right)^\mu$$

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



$$= \prod_{i=1}^n \mathbb{E}[e^{\lambda X_i}] \leq e^{(e^\lambda - 1)\mu}$$

- Optimization:**

$(e^\lambda - 1 - \lambda(1 + \delta))$ achieves Min at stationary point $\lambda = \ln(1 + \delta)$

- Independent $X_1, \dots, X_n \in \{0,1\}$ with

$$X = \sum_{i=1}^n X_i \text{ and } \mu = \mathbb{E}[X]$$

- Markov for MGF:** (for any $\lambda > 0$)

$$\Pr [X \geq (1 + \delta)\mu] \leq \frac{\mathbb{E} [e^{\lambda X}]}{e^{\lambda(1+\delta)\mu}} \leq e^{(e^\lambda - 1 - \lambda(1+\delta))\mu} = \left(\frac{e^\delta}{(1 + \delta)^{(1+\delta)}} \right)^\mu$$

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

- Bound MGF:**

$$\mathbb{E} [e^{\lambda X}] = \mathbb{E} \left[\prod_{i=1}^n e^{\lambda X_i} \right] = \prod_{i=1}^n \mathbb{E} [e^{\lambda X_i}] \leq e^{(e^\lambda - 1)\mu}$$

- Optimization:**

$(e^\lambda - 1 - \lambda(1 + \delta))$ achieves Min at stationary point $\lambda = \ln(1 + \delta)$

Chernoff Bound

Chernoff Bound:

For *independent* $X_1, \dots, X_n \in \{0,1\}$ with

$$X = \sum_{i=1}^n X_i \text{ and } \mu = \mathbb{E}[X]$$

For any $\delta > 0$,

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

Chernoff Bound

Chernoff Bound:

For *independent* $X_1, \dots, X_n \in \{0,1\}$ with

$$X = \sum_{i=1}^n X_i \text{ and } \mu = \mathbb{E}[X]$$

For any $0 < \delta < 1$,

$$\Pr[X \leq (1 - \delta)\mu] \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}} \right)^\mu$$

(for any $\lambda < 0$)

$$\Pr[X \leq (1 - \delta)\mu] \leq \Pr[e^{\lambda X} \geq e^{\lambda(1-\delta)\mu}] \leq \frac{\mathbb{E}[e^{\lambda X}]}{e^{\lambda(1-\delta)\mu}} \leq e^{(e^\lambda - 1 - \lambda(1-\delta))\mu}$$

(for $\lambda = \ln(1 - \delta)$) $= \left(\frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}} \right)^\mu$

Chernoff Bound:

For *independent* or *negatively associated* $X_1, \dots, X_n \in \{0,1\}$

$$\text{with } X = \sum_{i=1}^n X_i \text{ and } \mu = \mathbb{E}[X]$$

For any $\delta > 0$,

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

For any $0 < \delta < 1$,

$$\Pr[X \leq (1 - \delta)\mu] \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{(1-\delta)}} \right)^\mu$$

For *negatively associated* $X_1, \dots, X_n \in \{0,1\}$:

$$\mathbb{E} [e^{\lambda X}] = \mathbb{E} \left[\prod_{i=1}^n e^{\lambda X_i} \right] \leq \prod_{i=1}^n \mathbb{E} [e^{\lambda X_i}]$$

Chernoff-Hoeffding Bound

Chernoff Bound:

For $X = \sum_{i=1}^n X_i$, where $X_1, \dots, X_n \in \{0,1\}$ are *independent*
(or *negatively associated*),
for any $t > 0$:

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

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

A party of $O(\sqrt{n \log n})$ can manipulate a vote w.h.p. against n voters who neither care (uniform) nor communicate (independent).

(two-sided) Error Reduction

- Decision problem $f: \{0,1\}^* \rightarrow \{0,1\}$.
- Monte Carlo randomized algorithm $\mathcal{A}$ with *two-sided* error:
 - $\forall x \in \{0,1\}^*: \Pr(\mathcal{A}(x) = f(x)) \geq \frac{1}{2} + p$
- $\mathcal{A}^n$: **independently** run $\mathcal{A}$ for n times, return **majority** of the n outputs

$$\Pr(\mathcal{A}^n(x) \neq f(x)) = \Pr\left(X \leq \frac{n}{2}\right) = \Pr(X \leq \mathbb{E}[X] - pn) \leq \exp(-2p^2n) \leq \delta$$

$$\text{where } X = \sum_{i=1}^n X_i \text{ and } X_i = I[\mathcal{A}(x) = f(x) \text{ in } i\text{th run}]$$

$$\text{when } n \geq \frac{1}{2p^2} \ln \frac{1}{\delta}$$

The Median Trick

- Computation problem $f : \{0,1\}^* \rightarrow \mathbb{R}$
- Randomized approximation algorithm $\mathcal{A}$: $\forall x \in \{0,1\}^*$,

$$\Pr(\mathcal{A}(x) \in (1 \pm \epsilon)f(x)) = \Pr((1 - \epsilon)f(x) \leq \mathcal{A}(x) \leq (1 + \epsilon)f(x)) \geq \frac{1}{2} + p$$

- $\mathcal{A}^n$: **independently** run $\mathcal{A}$ for n times, return **median** of the n outputs
 - Let $X_i = I[\mathcal{A}(x) \in (1 \pm \epsilon)f(x) \text{ in the } i\text{th run of } \mathcal{A}(x)] \implies \mathbb{E}[X_i] \geq 1/2 + p$

- **Observation:** $\mathcal{A}^n(x) \in (1 \pm \epsilon)f(x)$ if $X = \sum_{i=1}^n X_i > \frac{n}{2}$

$$\Pr(\mathcal{A}(x) \notin (1 \pm \epsilon)f(x)) \leq \Pr\left(X \leq \frac{n}{2}\right) \leq \Pr(X \leq \mathbb{E}[X] - np) \leq e^{-2p^2n} \leq \delta$$

$$\text{when } n \geq \frac{1}{2p^2} \ln \frac{1}{\delta}$$

Chernoff-Hoeffding Bound

Chernoff Bound:

For $X = \sum_{i=1}^n X_i$, where $X_1, \dots, X_n \in \{0,1\}$ are *independent*
(or *negatively associated*),

for any $t > 0$:

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

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

Chernoff-Hoeffding Bound

Hoeffding Bound:

For $X = \sum_{i=1}^n X_i$, where $X_i \in [a_i, b_i]$, $1 \leq i \leq n$, are *independent*
(or *negatively associated*),

for any $t > 0$:

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

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

$$X = \sum_{i=1}^n X_i, \text{ where } X_i \in [a_i, b_i] \text{ for every } 1 \leq i \leq n$$

$$\text{let } \begin{cases} Y = X - \mathbb{E}[X] \\ Y_i = X_i - \mathbb{E}[X_i] \end{cases} \Rightarrow \begin{cases} \mathbb{E}[Y] = \mathbb{E}[Y_i] = 0 \\ Y = \sum_{i=1}^n Y_i \end{cases}$$

(for $\lambda > 0$)

(neg. assoc.)_n

$$\Pr[X - \mathbb{E}[X] \geq t] = \Pr[Y \geq t] \leq e^{-\lambda t} \mathbb{E}[e^{\lambda Y}] \leq e^{-\lambda t} \prod_{i=1}^n \mathbb{E}[e^{\lambda Y_i}]$$

Hoeffding's Lemma: For any $Y \in [a, b]$ with $\mathbb{E}[Y] = 0$,

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

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

$$\text{when } \lambda = \frac{4t}{\sum_{i=1}^n (b_i - a_i)^2}$$

$$X = \sum_{i=1}^n X_i, \text{ where } X_i \in [a_i, b_i] \text{ for every } 1 \leq i \leq n$$

$$\text{let } \begin{cases} Y = X - \mathbb{E}[X] \\ Y_i = X_i - \mathbb{E}[X_i] \end{cases} \Rightarrow \begin{cases} \mathbb{E}[Y] = \mathbb{E}[Y_i] = 0 \\ Y = \sum_{i=1}^n Y_i \end{cases}$$

(for $\lambda < 0$) (neg. assoc.) n

$$\Pr[X - \mathbb{E}[X] \leq -t] = \Pr[Y \leq -t] \leq e^{\lambda t} \mathbb{E}[e^{\lambda Y}] \leq e^{\lambda t} \prod_{i=1}^n \mathbb{E}[e^{\lambda Y_i}]$$

Hoeffding's Lemma: For any $Y \in [a, b]$ with $\mathbb{E}[Y] = 0$,

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

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

$$\text{when } \lambda = \frac{-4t}{\sum_{i=1}^n (b_i - a_i)^2}$$

Chernoff-Hoeffding Bound

Hoeffding Bound:

For $X = \sum_{i=1}^n X_i$, where $X_i \in [a_i, b_i]$, $1 \leq i \leq n$, are *independent*
(or *negatively associated*),

for any $t > 0$:

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

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

Sub-Gaussian Random Variables

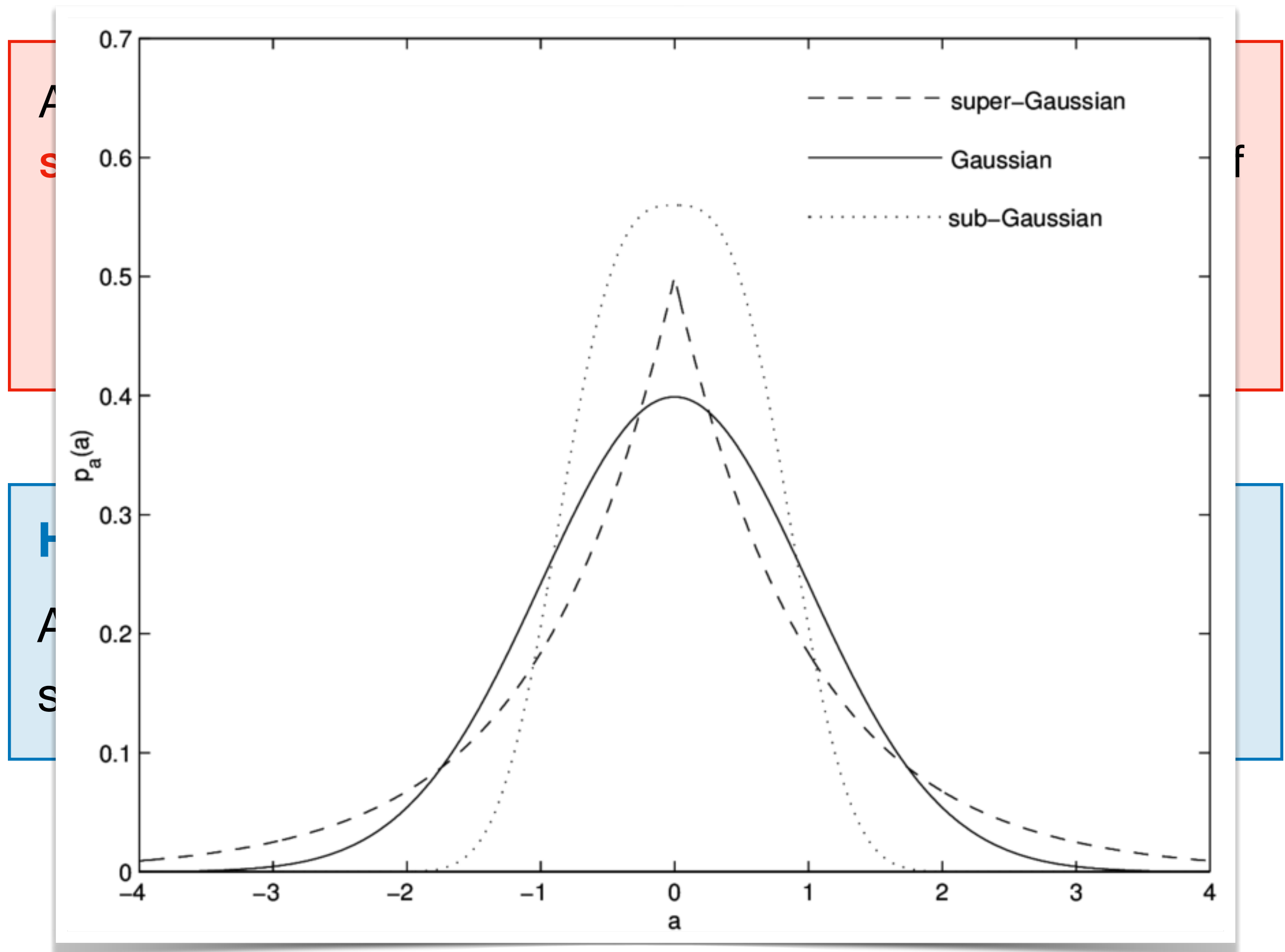
A *centered* ($\mathbb{E}[Y] = 0$) random variable Y is said to be **sub-Gaussian with variance factor ν** (denoted $Y \in \mathcal{G}(\nu)$) if

$$\mathbb{E} [e^{\lambda Y}] \leq \exp \left(\frac{\lambda^2 \nu}{2} \right)$$

Hoeffding's Lemma:

Any centered bounded random variable $Y \in [a, b]$ is sub-Gaussian with variance factor $(b - a)^2/4$.

Sub-Gaussian Random Variables



Sub-Gaussian Random Variables

Chernoff-Hoeffding:

For $Y = \sum_{i=1}^n Y_i$, where $Y_i \in \mathcal{G}(\nu_i)$, $1 \leq i \leq n$, are *independent*

(or *negatively associated*) and centered (i.e. $\mathbb{E}[Y_i] = 0$)

for any $t > 0$:

$$\Pr [Y \geq t] \leq \exp \left(-\frac{t^2}{2 \sum_{i=1}^n \nu_i} \right)$$

$$\Pr [Y \leq -t] \leq \exp \left(-\frac{t^2}{2 \sum_{i=1}^n \nu_i} \right)$$

The Method of Bounded Differences

The Method of Bounded Differences

McDiarmid's Inequality:

For independent $X_1, X_2, \dots, X_n$, if n -variate function f satisfies the **Lipschitz condition**: for every $1 \leq i \leq n$,

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

for any possible $x_1, \dots, x_n$ and y_i ,

then for any $t > 0$:

$$\Pr \left[\left| f(X_1, \dots, X_n) - \mathbb{E}[f(X_1, \dots, X_n)] \right| \geq t \right] \leq 2 \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

- Chernoff: sum of Boolean variables, 1-Lipschitz
- Hoeffding: sum of $[a_i, b_i]$ -bounded variables, $(b_i - a_i)$ -Lipschitz

Balls into Bins

m balls are thrown into n bins
 Y : number of empty bins

$$Y_i = \begin{cases} 1 & \text{bin } i \text{ is empty} \\ 0 & \text{otherwise} \end{cases}$$

$$Y = \sum_{i=1}^n Y_i \quad \mathbb{E}[Y_i] = \Pr[\text{bin } i \text{ is empty}] = \left(1 - \frac{1}{n}\right)^m$$

linearity of expectation:

$$\mathbb{E}[Y] = \sum_{i=1}^n \mathbb{E}[Y_i] = n \left(1 - \frac{1}{n}\right)^m$$

$$\Pr \left[|Y - \mathbb{E}[Y]| \geq t \right] < ?$$

Y_i 's are dependent

Balls into Bins

m balls are thrown into n bins
 Y : number of empty bins

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

X_j : the index of the bin into which the j -th ball is thrown

$X_1, \dots, X_m \in [n]$ are uniform and **independent**

$Y = f(X_1, \dots, X_m) = n - \left| \{X_1, \dots, X_m\} \right|$ is **1-Lipschitz**

$$\Pr \left[\left| Y - \mathbb{E}[Y] \right| \geq t \right] = \Pr \left[\left| f(X_1, \dots, X_m) - \mathbb{E}[f(X_1, \dots, X_m)] \right| \geq t \right]$$

(**McDiarmid's inequality**) $\leq 2 \exp \left(-\frac{t^2}{2m} \right)$

Pattern Matching

uniform random string $X \in \Sigma^n$ with alphabet size $|\Sigma| = m$

fixed pattern $\pi \in \Sigma^k$

Y : number of substrings of X matching the pattern π

$$Y_i = \begin{cases} 1 & \text{if } X_i X_{i+1} \cdots X_{i+k-1} = \pi \\ 0 & \text{otherwise} \end{cases} \quad Y = \sum_{i=1}^{n-k+1} Y_i$$

$$\mathbb{E}[Y_i] = \Pr[X_i X_{i+1} \cdots X_{i+k-1} = \pi] = \frac{1}{m^k}$$

linearity of expectation:

$$\mathbb{E}[Y] = \sum_{i=1}^{n-k+1} \mathbb{E}[Y_i] = \frac{n-k+1}{m^k}$$

Pattern Matching

uniform random string $X \in \Sigma^n$ with alphabet size $|\Sigma| = m$

fixed pattern $\pi \in \Sigma^k$

Y : number of substrings of X matching the pattern π

$$\mathbb{E}[Y] = \frac{n - k + 1}{m^k} \quad X_1, \dots, X_n \in \Sigma \text{ are independent}$$

$$Y = f_\pi(X_1, \dots, X_n) = \sum_{i=1}^{n-k+1} I[X_i X_{i+1} \cdots X_{i+k-1} = \pi] \text{ is } k\text{-Lipschitz}$$

$$\Pr \left[|Y - \mathbb{E}[Y]| \geq t \right] = \Pr \left[|f_\pi(X_1, \dots, X_n) - \mathbb{E}[f_\pi(X_1, \dots, X_n)]| \geq t \right]$$

$$\text{(McDiarmid's inequality)} \leq 2 \exp \left(-\frac{t^2}{2nk^2} \right)$$

Sprinkling Points on Hypercube

uniform random point $X \in \{0,1\}^n$ in hypercube

fixed subset $S \subseteq \{0,1\}^n$

Y : shortest *Hamming* distance from X to S

Hamming distance: $H(x, y) = \sum_{i=1}^n |x_i - y_i|$ for $x, y \in \{0,1\}^n$

$Y = \min_{y \in S} H(X, y) = f_S(X_1, \dots, X_n)$ is **1-Lipschitz**

$$\Pr \left[|Y - \mathbb{E}[Y]| \geq t \right] = \Pr \left[|f_S(X_1, \dots, X_n) - \mathbb{E}[f_S(X_1, \dots, X_n)]| \geq t \right]$$

(**McDiarmid's inequality**) $\leq 2 \exp \left(-\frac{t^2}{2n} \right)$

Sprinkling Points on Hypercube

uniform random point $X \in \{0,1\}^n$ in hypercube

fixed subset $S \subseteq \{0,1\}^n$

Y : shortest *Hamming* distance from X to S

Hamming distance: $H(x, y) = \sum_{i=1}^n |x_i - y_i|$ for $x, y \in \{0,1\}^n$

$$\Pr \left[|Y - \mathbb{E}[Y]| \geq \sqrt{2cn \ln n} \right] \leq 2n^{-c}$$

the distance to S is pretty much the same from pretty much everywhere
(unless S is very big)

Bloom Filters (Bloom 1970)

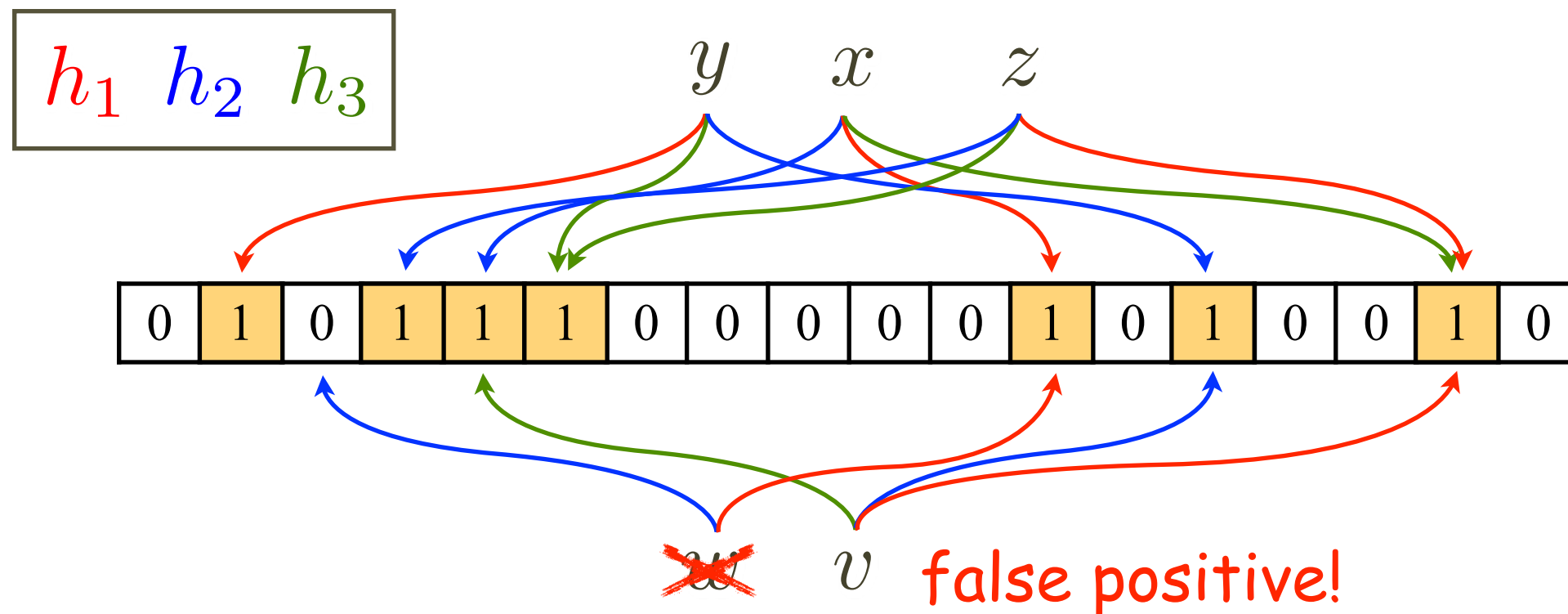
- uniform & independent hash function $h_1, \dots, h_k : U \rightarrow [m]$

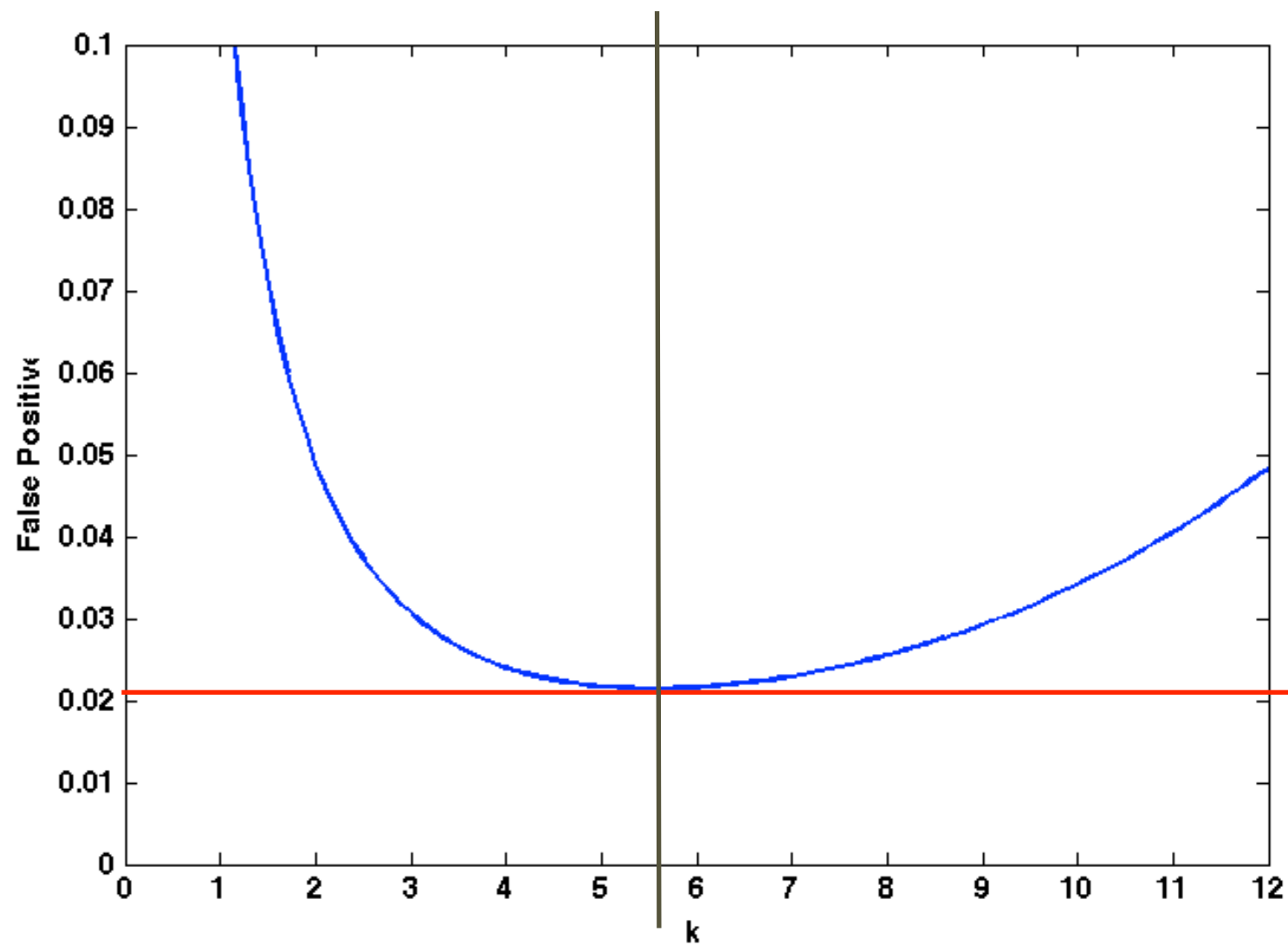
Data Structure: bit array $A \in \{0,1\}^m$

A is initialized to all 0's;

for each $x_i \in S$: set $A[h_j(x_i)] = 1$ for all $1 \leq j \leq k$;

Query x : “yes” iff $A[h_j(x)] = 1$ for all $1 \leq j \leq k$





$y: x \in U$

$\dots, h_k : U \rightarrow [m]$

all $1 \leq j \leq k;$

$\leq j \leq k$

- $x \notin S$: false positive

$$\Pr \left[\forall 1 \leq j \leq k : A[h_j(x)] = 1 \right]$$

choose $k = c \ln 2$
 $m = cn$

$$\begin{aligned} \text{heuristic} &= \left(\Pr \left[A[h_j(x)] = 1 \right] \right)^k = \left(1 - \Pr \left[A[h_j(x)] = 0 \right] \right)^k \\ &\leq (1 - (1 - 1/m)^{kn})^k \approx (1 - e^{-kn/m})^k = 2^{-c \ln 2} \leq (0.6185)^c \end{aligned}$$

- After inserting n keys: for each $1 \leq i \leq m$

$$\Pr [A[i] = 1] = 1 - \Pr [A[i] = 0] = 1 - (1 - 1/m)^{kn}$$

$$\Rightarrow \mathbb{E} \left[\sum_{i=1}^m A[i] \right] = m(1 - (1 - 1/m)^{kn})$$

- $x \notin S$: **false positive**

$$\Pr \left[\forall 1 \leq j \leq k : A[h_j(x)] = 1 \right] = \left(\frac{\sum_{i=1}^m A[i]}{m} \right)^k \quad \text{conditioned on any } A \in \{0,1\}^m$$

$$= \sum_{s=1}^m \left(\frac{s}{m} \right)^k \Pr \left[\sum_{i=1}^m A[i] = s \right]$$

$$\leq \left(1 - (1 - 1/m)^{kn} + \frac{t}{m} \right)^k + \Pr \left[\left| \sum_{i=1}^m A[i] - \mathbb{E} \left[\sum_{i=1}^m A[i] \right] \right| > t \right]$$

(McDiarmid's inequality)

$$\leq 2 \exp \left(-\frac{t^2}{2kn} \right)$$

- After inserting n keys: for each $1 \leq i \leq m$

$$\Pr [A[i] = 1] = 1 - \Pr [A[i] = 0] = 1 - (1 - 1/m)^{kn}$$

$$\Rightarrow \mathbb{E} \left[\sum_{i=1}^m A[i] \right] = m(1 - (1 - 1/m)^{kn})$$

- $x \notin S$: false positive

$$\Pr \left[\forall 1 \leq j \leq k : A[h_j(x)] = 1 \right]$$

$$\leq \left(1 - (1 - 1/m)^{kn} + \frac{t}{m} \right)^k + 2 \exp \left(-\frac{t^2}{2kn} \right)$$

$$\leq \left(1 - (1 - 1/m)^{kn} + o(1) \right)^k + n^{-100}$$

$$\approx \left(1 - e^{-kn/m} \right)^k + n^{-100}$$

$$= 2^{-c \ln 2} + n^{-100} \leq (0.6185)^c + n^{-100}$$

choose

$$t = \sqrt{200kn \ln(2n)}$$

choose $k = c \ln 2$

where $m = cn$

The Method of Bounded Differences

McDiarmid's Inequality:

For independent $X_1, X_2, \dots, X_n$, if n -variate function f satisfies the **Lipschitz condition**: for every $1 \leq i \leq n$,

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

for any possible $x_1, \dots, x_n$ and y_i ,

then for any $t > 0$:

$$\Pr \left[\left| f(X_1, \dots, X_n) - \mathbb{E}[f(X_1, \dots, X_n)] \right| \geq t \right] \leq 2 \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

Every Lipschitz function is well approximated by a constant function under product measures.

Martingale

Martingale:

A sequence of random variables $X_0, X_1, \dots$ is a **martingale** if for all $t > 0$,

$$\mathbb{E} [X_t \mid X_0, X_1, \dots, X_{t-1}] = X_{t-1}.$$

- For random variable X and event A : (discrete probability)

$$\mathbb{E}[X \mid A] = \sum_x x \Pr[X = x \mid A]$$

- For random variables X and Y (not necessarily independent):

$f(y) = \mathbb{E}[X \mid Y = y]$ is well-defined

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



Martingale:

A sequence of random variables $X_0, X_1, \dots$ is a **martingale** if for all $t > 0$,

$$\mathbb{E} [X_t \mid X_0, X_1, \dots, X_{t-1}] = X_{t-1}.$$

- **Fair gambling game:** Given the capitals up until time $t - 1$, the expected change to the capital after the t -th bet is 0.

A sequence of random variables $X_0, X_1, \dots$ is:
a **super-martingale** if for all $t > 0$,

$$\mathbb{E} [X_t \mid X_0, X_1, \dots, X_{t-1}] \leq X_{t-1}$$

a **sub-martingale** if for all $t > 0$,

$$\mathbb{E} [X_t \mid X_0, X_1, \dots, X_{t-1}] \geq X_{t-1}$$

Martingale (Generalized)

Martingale (Generalized Version):

A sequence of random variables $Y_0, Y_1, \dots$ is a **martingale with respect to** $X_0, X_1, \dots$ if for all $t \geq 0$,

- Y_t is a function of $X_0, \dots, X_t$
- $\mathbb{E} [Y_{t+1} \mid X_0, X_1, \dots, X_t] = Y_t$

- A fair gambling game:
 - X_i : outcome (win/loss) of the i -th betting
 - Y_i : capital after the i -th betting

Martingale (Generalized)

Martingale (Generalized Version):

A sequence of random variables $Y_0, Y_1, \dots$ is a **martingale with respect to** $X_0, X_1, \dots$ if for all $t \geq 0$,

- Y_t is a function of $X_0, \dots, X_t$
- $\mathbb{E} [Y_{t+1} \mid X_0, X_1, \dots, X_t] = Y_t$

- A probability space: $(\Omega, \mathcal{F}, \text{Pr})$
- A filtration of σ -algebras $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots$ s.t. for all $t \geq 0$:
 - Y_t is $\mathcal{F}_t$ -measurable
 - $\mathbb{E} [Y_{t+1} \mid \mathcal{F}_t] = Y_t$

Azuma's Inequality

Azuma's Inequality:

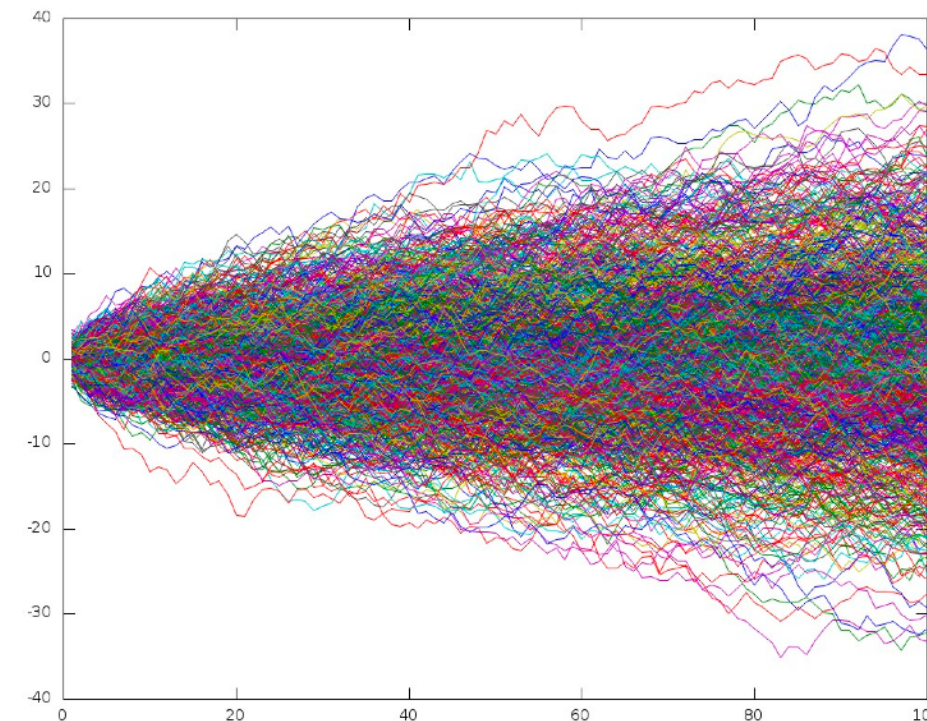
For martingale $Y_0, Y_1, \dots$ (with respect to $X_0, X_1, \dots$) satisfying:

$$\forall i \geq 0, \left| Y_i - Y_{i-1} \right| \leq c_i$$

for any $n \geq 1$ and $t > 0$:

$$\Pr \left[\left| Y_n - Y_0 \right| \geq t \right] \leq 2 \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

- Your capital does not change too fast if:
 - the game is fair (martingale)
 - payoff for each gambling is bounded



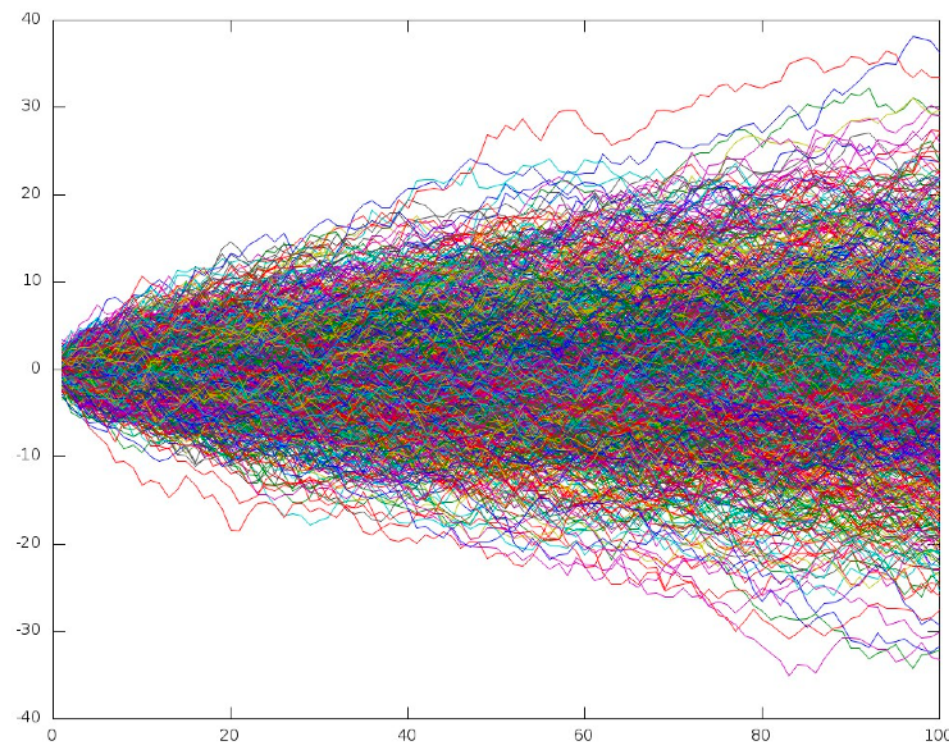
Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$$

$$Y_0 = \mathbb{E} [f(X_1, \dots, X_n)] \quad \text{-----} \rightarrow \quad f(X_1, \dots, X_n) = Y_n$$

no information full information



Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$$



$$\mathbb{E}[f] = Y_0,$$

averaged over

Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$$

randomized by

$$f(\textcircled{1}, \textcircled{\$}, \textcircled{\$}, \textcircled{\$}, \textcircled{\$}, \textcircled{\$})$$

averaged over

$$\mathbf{E}[f] = Y_0, \quad Y_1,$$

Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$$

randomized by

$$f(\overbrace{1, 0}, \underbrace{\$, \$, \$, \$}_{\text{averaged over}})$$

$$\mathbf{E}[f] = Y_0, \quad Y_1, \quad Y_2,$$

Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$$

randomized by

$$f(\underbrace{1, 0, 0}_{\text{deterministic}}, \underbrace{\text{coin}, \text{coin}, \text{coin}}_{\text{randomized}})$$

averaged over

$$\mathbb{E}[f] = Y_0, \quad Y_1, \quad Y_2, \quad Y_3,$$

Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$$

randomized by

$$f(\underbrace{1, 0, 0, 1}_{\text{randomized by}}, \underbrace{\text{coin}, \text{coin}}_{\text{averaged over}})$$

averaged over

$$\mathbf{E}[f] = Y_0, \quad Y_1, \quad Y_2, \quad Y_3, \quad Y_4,$$

Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$$

randomized by

$$f(\textcircled{1}, \textcircled{0}, \textcircled{0}, \textcircled{1}, \textcircled{0}, \textcircled{\$})$$

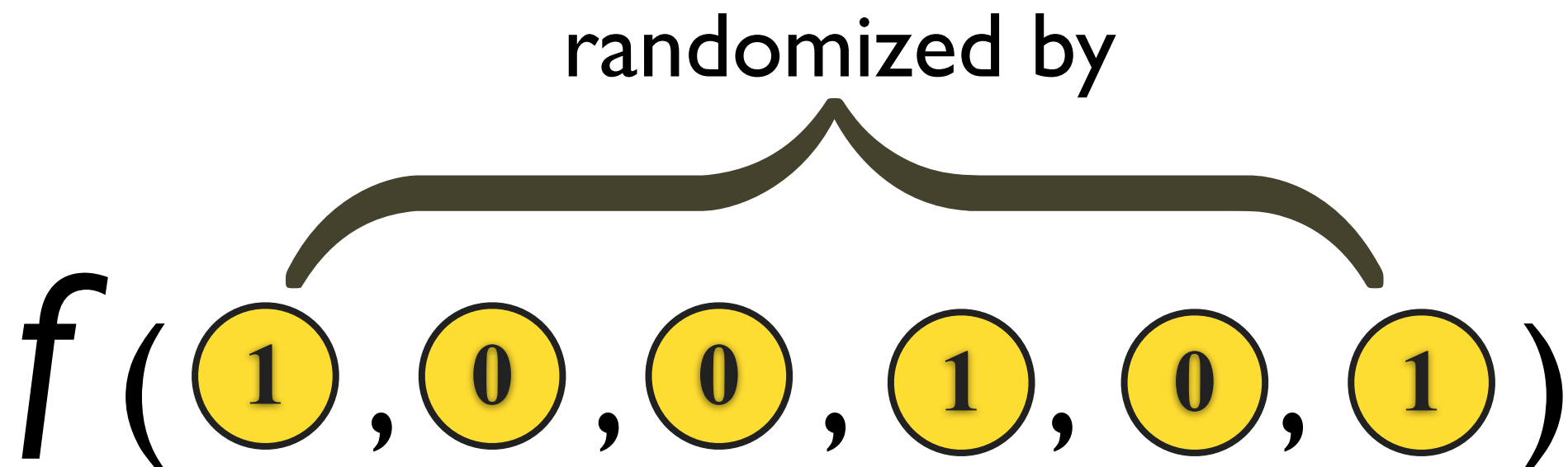
averaged over

$$\mathbf{E}[f] = Y_0, \quad Y_1, \quad Y_2, \quad Y_3, \quad Y_4, \quad Y_5,$$

Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$$



no information

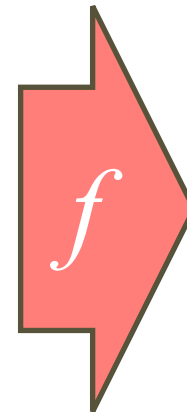
full information

$$\mathbf{E}[f] = Y_0, \quad Y_1, \quad Y_2, \quad Y_3, \quad Y_4, \quad Y_5, \quad Y_6 = f$$

Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$$



Doob Martingale

A **Doob sequence** $Y_0, Y_1, \dots, Y_n$ of an n -variate function f with respect to a random vector $(X_1, \dots, X_n)$ is:

$$\forall 0 \leq i \leq n, \quad Y_i = \mathbb{E} \left[f(X_1, \dots, X_n) \mid X_1, \dots, X_i \right]$$

Theorem:

The Doob sequence $Y_0, Y_1, \dots, Y_n$ is a martingale w.r.t. $X_1, \dots, X_n$.

- $\forall 0 \leq i \leq n, Y_i$ is a function of $X_1, \dots, X_i$
- $\mathbb{E} \left[Y_i \mid X_1, \dots, X_{i-1} \right]$
 $= \mathbb{E} \left[\mathbb{E} \left[f(X_1, \dots, X_n) \mid X_1, \dots, X_i \right] \mid X_1, \dots, X_{i-1} \right]$
 $= \mathbb{E} \left[f(X_1, \dots, X_n) \mid X_1, \dots, X_{i-1} \right] = Y_{i-1}$

The Method of Bounded Differences

The Method of Bounded Differences:

For n -variate function f on random vector $X = (X_1, \dots, X_n)$ satisfying the **Lipschitz condition**: for every $1 \leq i \leq n$

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

for any $t > 0$:

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

The Method of Bounded Differences

The Method of Bounded Differences:

For n -variate function f on random vector $X = (X_1, \dots, X_n)$ satisfying the **Lipschitz condition**: for every $1 \leq i \leq n$

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

for any $t > 0$: Y_i Y_{i-1}

$$\Pr \left[\left| \underbrace{f(X)}_{Y_n} - \underbrace{\mathbb{E}[f(X)]}_{Y_0} \right| \geq t \right] \leq \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

Doob martingale: $Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$

Azuma's Inequality

Azuma's Inequality:

For martingale $Y_0, Y_1, \dots$ (with respect to $X_0, X_1, \dots$) satisfying:

$$\forall i \geq 0, \left| Y_i - Y_{i-1} \right| \leq c_i$$

for any $n \geq 1$ and $t > 0$:

$$\Pr \left[\left| Y_n - Y_0 \right| \geq t \right] \leq 2 \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

The Method of Bounded Differences

The Method of Bounded Differences:

For n -variate function f on random vector $X = (X_1, \dots, X_n)$ satisfying the **Lipschitz condition**: for every $1 \leq i \leq n$

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

for any $t > 0$: Y_i Y_{i-1}

$$(\text{Azuma}) \Pr \left[\left| \underbrace{f(X)}_{Y_n} - \underbrace{\mathbb{E}[f(X)]}_{Y_0} \right| \geq t \right] \leq \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

Doob martingale: $Y_i = \mathbb{E} [f(X_1, \dots, X_n) \mid X_1, \dots, X_i]$

The Method of Bounded Differences

The Method of Bounded Differences:

For n -variate function f on random vector $X = (X_1, \dots, X_n)$ satisfying the **average-case Lipschitz condition**: for every $1 \leq i \leq n$

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

for any $t > 0$:

usually difficult to verify

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

worst-case Lipschitz: for every $1 \leq i \leq n$,

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

for any possible $x_1, \dots, x_n$ and y_i

The Method of Bounded Differences

McDiarmid's Inequality:

For independent $X_1, X_2, \dots, X_n$, if n -variate function f satisfies the **Lipschitz condition**: for every $1 \leq i \leq n$,

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

for any possible $x_1, \dots, x_n$ and y_i ,

then for any $t > 0$:

$$\Pr \left[\left| f(X_1, \dots, X_n) - \mathbb{E}[f(X_1, \dots, X_n)] \right| \geq t \right] \leq 2 \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

worst-case Lipschitz condition + independent $X_1, \dots, X_n$ $\implies$ **average-case Lipschitz condition**

Martingale Concentration

Azuma's Inequality:

For martingale $Y_0, Y_1, \dots$ (with respect to $X_0, X_1, \dots$) satisfying:

$$\forall i \geq 0, \left| Y_i - Y_{i-1} \right| \leq c_i$$

for any $n \geq 1$ and $t > 0$:

$$\Pr \left[\left| Y_n - Y_0 \right| \geq t \right] \leq 2 \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

Difference: $D_i = Y_i - Y_{i-1}$ $S_n = \sum_{i=1}^n D_i = Y_n - Y_0$

Martingale difference: D_i is a function of $X_0, \dots, X_i$

$$\begin{aligned} \mathbb{E} \left[D_i \mid X_0, \dots, X_{i-1} \right] &= \mathbb{E} \left[Y_i - Y_{i-1} \mid X_0, \dots, X_{i-1} \right] \\ &= \mathbb{E} \left[Y_i \mid X_0, \dots, X_{i-1} \right] - \mathbb{E} \left[Y_{i-1} \mid X_0, \dots, X_{i-1} \right] \\ &= Y_{i-1} - Y_{i-1} = 0 \end{aligned}$$

- **Martingale property:**

- D_i is a function of $X_0, \dots, X_i$ and $\mathbb{E} [D_i \mid X_0, \dots, X_{i-1}] = 0$

- **Bounded differences:** $\forall i \geq 1, |D_i| \leq c_i$

$$S_n = \sum_{i=1}^n D_i$$

$$\textbf{Azuma: } \Pr \left[|S_n| \geq t \right] \leq 2 \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

$$(\text{for } \lambda > 0) \quad \Pr [S_n \geq t] = \Pr [e^{\lambda S_n} \geq e^{\lambda t}] \leq e^{-\lambda t} \mathbb{E} [e^{\lambda S_n}]$$

$$\begin{aligned} \mathbb{E} [e^{\lambda S_n}] &= \mathbb{E} \left[\mathbb{E} [e^{\lambda S_n} \mid X_0, \dots, X_{n-1}] \right] = \mathbb{E} \left[\mathbb{E} [e^{\lambda(S_{n-1} + D_n)} \mid X_0, \dots, X_{n-1}] \right] \\ &= \mathbb{E} \left[\mathbb{E} [e^{\lambda S_{n-1}} \cdot e^{\lambda D_n} \mid X_0, \dots, X_{n-1}] \right] = \mathbb{E} \left[e^{\lambda S_{n-1}} \cdot \mathbb{E} [e^{\lambda D_n} \mid X_0, \dots, X_{n-1}] \right] \end{aligned}$$

- **Martingale property:**

- D_i is a function of $X_0, \dots, X_i$ and $\mathbb{E} [D_i \mid X_0, \dots, X_{i-1}] = 0$

- **Bounded differences:** $\forall i \geq 1, |D_i| \leq c_i$

$$S_n = \sum_{i=1}^n D_i$$

$$\mathbb{E} [e^{\lambda S_n}] = \mathbb{E} \left[e^{\lambda S_{n-1}} \cdot \mathbb{E} [e^{\lambda D_n} \mid X_0, \dots, X_{n-1}] \right] \leq e^{\lambda^2 c_n^2 / 2} \cdot \mathbb{E} [e^{\lambda S_{n-1}}]$$

Hoeffding's Lemma: For any $Z \in [a, b]$ with $\mathbb{E}[Z] = 0$,

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

$$Z = (D_n \mid X_0, \dots, X_{n-1}) \quad \implies \quad \mathbb{E} [e^{\lambda D_n} \mid X_0, \dots, X_{n-1}] \leq e^{\lambda^2 c_n^2 / 2}$$

- **Martingale property:**

- D_i is a function of $X_0, \dots, X_i$ and $\mathbb{E} [D_i \mid X_0, \dots, X_{i-1}] = 0$

- **Bounded differences:** $\forall i \geq 1, |D_i| \leq c_i$

$$S_n = \sum_{i=1}^n D_i$$

$$\mathbb{E} [e^{\lambda S_n}] \leq e^{\lambda^2 c_n^2 / 2} \cdot \mathbb{E} [e^{\lambda S_{n-1}}] \leq \exp \left(\frac{\lambda^2}{2} \sum_{i=1}^n c_i^2 \right)$$

(for $\lambda > 0$)

$$\Pr [S_n \geq t] \leq e^{-\lambda t} \mathbb{E} [e^{\lambda S_n}] \leq \exp \left(\frac{\lambda^2}{2} \sum_{i=1}^n c_i^2 - \lambda t \right) \leq \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$$

when $\lambda = \frac{t}{\sum_{i=1}^n c_i^2}$

Azuma: $\Pr [S_n \geq t] \leq \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$

- **Martingale property:**

- D_i is a function of $X_0, \dots, X_i$ and $\mathbb{E} [D_i \mid X_0, \dots, X_{i-1}] = 0$

- **Bounded differences:** $\forall i \geq 1, |D_i| \leq c_i$

$$S_n = \sum_{i=1}^n D_i$$

Azuma: $\Pr [S_n \leq -t] \leq \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right)$

(for $\lambda < 0$) $\Pr [S_n \leq -t] = \Pr [e^{\lambda S_n} \geq e^{-\lambda t}] \leq e^{\lambda t} \mathbb{E} [e^{\lambda S_n}]$

$$\leq \exp \left(\frac{\lambda^2}{2} \sum_{i=1}^n c_i^2 + \lambda t \right) \leq \exp \left(-\frac{t^2}{2 \sum_{i=1}^n c_i^2} \right) \quad \text{when } \lambda = \frac{-t}{\sum_{i=1}^n c_i^2}$$

$$\mathbb{E} [e^{\lambda S_n}] \leq e^{\lambda^2 c_n^2 / 2} \cdot \mathbb{E} [e^{\lambda S_{n-1}}] \leq \exp \left(\frac{\lambda^2}{2} \sum_{i=1}^n c_i^2 \right)$$

Poisson Approximation

Poisson Tails

Poisson random variable $X \sim \text{Pois}(\mu)$:

$$\Pr[X = k] = \frac{e^{-\mu} \mu^k}{k!} \text{ for } k = 0, 1, 2, \dots$$

Theorem: For $Y \sim \text{Pois}(\mu)$,

$$k > \mu \implies \Pr[X > k] < e^{-\mu} \left(\frac{e\mu}{k} \right)^k$$

$$k < \mu \implies \Pr[X < k] < e^{-\mu} \left(\frac{e\mu}{k} \right)^k$$

Poisson random variable $X \sim \text{Pois}(\mu)$:

$$\Pr[X = k] = \frac{e^{-\mu} \mu^k}{k!} \text{ for } k = 0, 1, 2, \dots$$

MGF: $\mathbb{E} [e^{\lambda X}] = \sum_{k=0}^{\infty} \frac{e^{-\mu} \mu^k}{k!} e^{\lambda k} = e^{\mu(e^\lambda - 1)} \sum_{k=0}^{\infty} \frac{e^{-\mu e^\lambda} (\mu e^\lambda)^k}{k!} = e^{\mu(e^\lambda - 1)}$

Theorem: For $Y \sim \text{Pois}(\mu)$,

$$k > \mu \implies \Pr[X > k] < e^{-\mu} \left(\frac{e\mu}{k} \right)^k$$

(for $\lambda > 0$)

$$\Pr[X > k] = \Pr [e^{\lambda X} > e^{\lambda k}] < \frac{\mathbb{E} [e^{\lambda X}]}{e^{\lambda k}} = e^{\mu(e^\lambda - 1) - \lambda k} = e^{-\mu} \left(\frac{e\mu}{k} \right)^k$$

when $\lambda = \ln(k/\mu) > 0$

Poisson random variable $X \sim \text{Pois}(\mu)$:

$$\Pr[X = k] = \frac{e^{-\mu} \mu^k}{k!} \text{ for } k = 0, 1, 2, \dots$$

MGF: $\mathbb{E} [e^{\lambda X}] = \sum_{k=0}^{\infty} \frac{e^{-\mu} \mu^k}{k!} e^{\lambda k} = e^{\mu(e^{\lambda}-1)} \sum_{k=0}^{\infty} \frac{e^{-\mu e^{\lambda}} (\mu e^{\lambda})^k}{k!} = e^{\mu(e^{\lambda}-1)}$

Theorem: For $Y \sim \text{Pois}(\mu)$,

$$k < \mu \implies \Pr[X < k] < e^{-\mu} \left(\frac{e\mu}{k} \right)^k$$

(for $\lambda < 0$)

$$\Pr[X < k] = \Pr [e^{\lambda X} > e^{\lambda k}] < \frac{\mathbb{E} [e^{\lambda X}]}{e^{\lambda k}} = e^{\mu(e^{\lambda}-1)-\lambda k} = e^{-\mu} \left(\frac{e\mu}{k} \right)^k$$

when $\lambda = \ln(k/\mu) < 0$

Poisson Heuristics

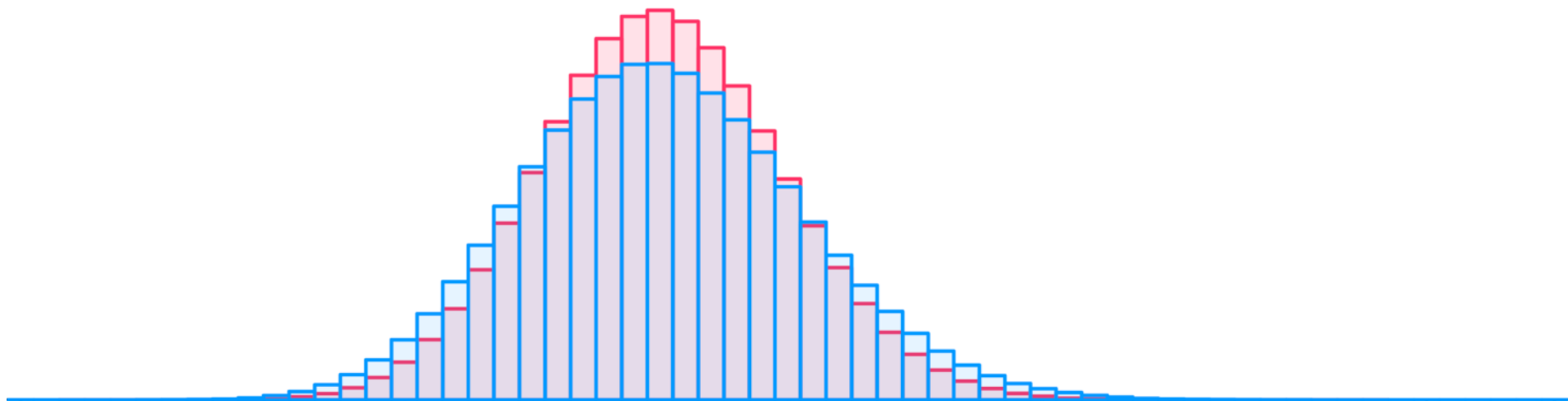
m balls are thrown into n bins.

X_i : number of balls in the i -th bin

- $X_1, \dots, X_n$ are correlated binomial random variables:

$$X_1, \dots, X_n \sim \text{Bin}(m, 1/n) \text{ subject to } \sum_{i=1}^n X_i = m$$

- *i.i.d.* **Poisson** random variables $Y_1, \dots, Y_n \sim \text{Pois}(m/n)$



Poisson Heuristics

m balls are thrown into n bins.

X_i : number of balls in the i -th bin

- **Heuristics:** treat loads of bins as *i.i.d.* $Y_1, \dots, Y_n \sim \text{Pois}(m/n)$

Poisson random variable $Y \sim \text{Pois}(\lambda)$:

$$\Pr[Y = k] = \frac{e^{-\lambda} \lambda^k}{k!} \text{ for } k = 0, 1, 2, \dots$$

(when $m = n \ln n + cn$)

Coupon collector:

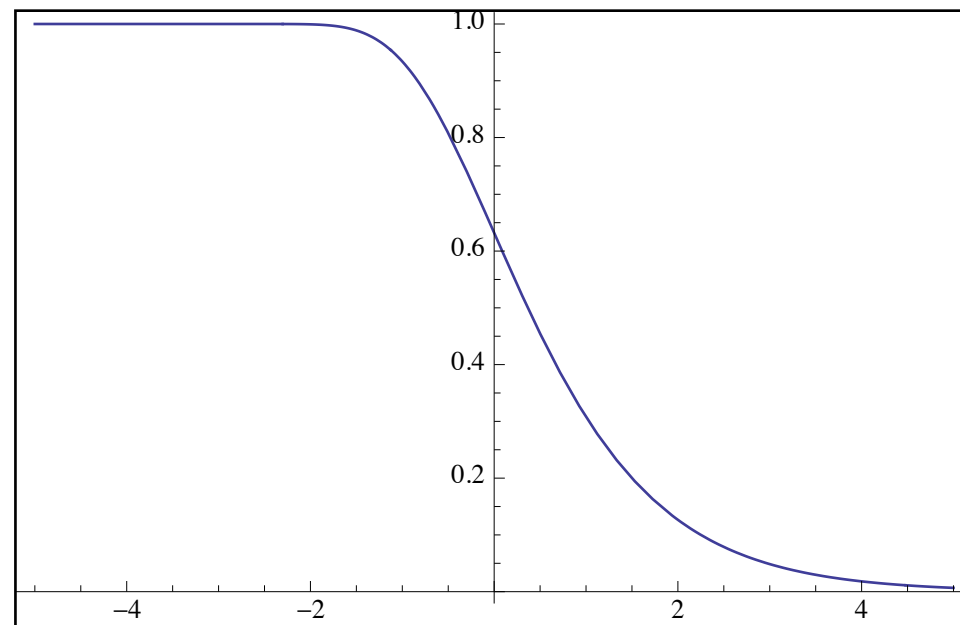
$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] = \left(1 - e^{-\frac{m}{n}} \right)^n = \left(1 - \frac{e^{-c}}{n} \right)^n \rightarrow e^{-e^{-c}}$$

Coupon Collector

X : number of balls thrown to make
all the n bins nonempty

a sharp threshold:

$$\lim_{n \rightarrow \infty} \Pr[X \geq n \ln n + cn] = 1 - e^{-e^{-c}}$$



Poisson Heuristics

m balls are thrown into n bins.

X_i : number of balls in the i -th bin

- **Heuristics:** treat loads of bins as *i.i.d.* $Y_1, \dots, Y_n \sim \text{Pois}(m/n)$

Poisson random variable $Y \sim \text{Pois}(\lambda)$:

$$\Pr[Y = k] = \frac{e^{-\lambda} \lambda^k}{k!} \text{ for } k = 0, 1, 2, \dots$$

**Occupancy
problem:**

$$\Pr \left[\max_{1 \leq i \leq n} Y_i < L \right] = \left(\Pr[Y_i < L] \right)^n \leq \left(1 - \Pr[Y_i = L] \right)^n$$
$$\text{(when } m = n) = \left(1 - \frac{1}{eL!} \right)^n \leq e^{-n/(eL!)} \leq \frac{1}{n^2} \quad \text{for } L = \frac{\ln n}{\ln \ln n}$$

since $L! \leq e\sqrt{L} (L/e)^L$

Occupancy Problem

n balls are thrown into n bins.

X_i : number of balls in the i -th bin

Theorem:

With high probability, the maximum load is

$$\max_{1 \leq i \leq n} X_i = \Omega \left(\frac{\log n}{\log \log n} \right)$$

Poisson Approximation

- loads of n bins receiving m balls: $X_1, \dots, X_n$
- i.i.d.* **Poisson** random variables $Y_1, \dots, Y_n \sim \text{Pois}(\lambda)$

Theorem: $\forall m_1, \dots, m_n \in \mathbb{N}$ s.t. $\sum_{i=1}^n m_i = m$

$$\Pr \left[\bigwedge_{i=1}^n X_i = m_i \right] = \Pr \left[\bigwedge_{i=1}^n Y_i = m_i \mid \sum_{i=1}^n Y_i = m \right]$$

$$\Pr \left[\bigwedge_{i=1}^n X_i = m_i \right] = \frac{\binom{m}{m_1, \dots, m_n}}{n^m} = \frac{m!}{m_1! \cdots m_n! n^m}$$

multinomial coefficient

$$\Pr \left[\bigwedge_{i=1}^n Y_i = m_i \mid \sum_{i=1}^n Y_i = m \right] = \frac{\Pr \left[\bigwedge_{i=1}^n Y_i = m_i \right]}{\Pr \left[\sum_{i=1}^n Y_i = m \right]} = \frac{\prod_{i=1}^n e^{-\lambda} \frac{\lambda^{m_i}}{m_i!}}{e^{-n\lambda} \frac{(n\lambda)^m}{m!}} = \frac{m!}{m_1! \cdots m_n! n^m}$$

$$m = n \ln n + cn$$

Thm: *i.i.d.* $Y_1, \dots, Y_n \sim \text{Pois}(\frac{m}{n})$ and $Y = \sum_{i=1}^n Y_i$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] = \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m \right] \pm o(1)$$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] = \sum_{k=0}^{\infty} \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = k \right] \Pr[Y = k]$$

choose

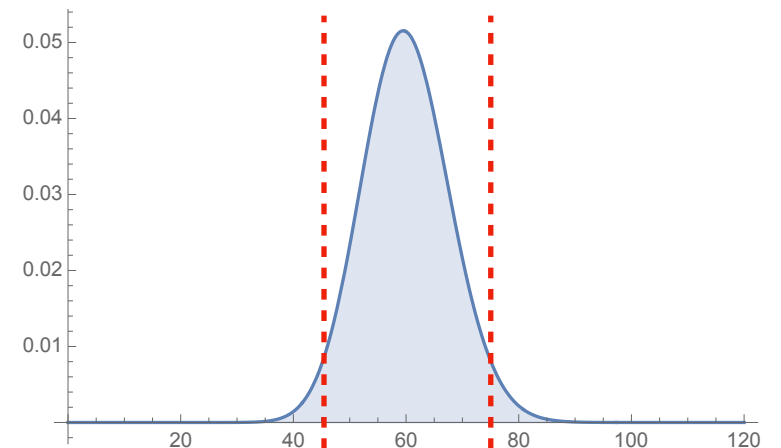
$$t = \sqrt{2m \ln m} \leq \sum_{k=m-t}^{m+t} \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = k \right] \Pr[Y = k]$$

$$+ \Pr[Y < m - t] + \Pr[Y > m + t] = o(1)$$

Lemma: $Y \sim \text{Pois}(m)$

$$k > m \Rightarrow \Pr[Y > k] < e^{-m} \left(\frac{em}{k} \right)^k$$

$$k < m \Rightarrow \Pr[Y < k] < e^{-m} \left(\frac{em}{k} \right)^k$$



$$m = n \ln n + cn$$

Thm: *i.i.d.* $Y_1, \dots, Y_n \sim \text{Pois}(\frac{m}{n})$ and $Y = \sum_{i=1}^n Y_i$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] = \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m \right] \pm o(1)$$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] = \sum_{k=m-t}^{m+t} \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = k \right] \Pr[Y = k] + o(1)$$

choose $t = \sqrt{2m \ln m}$ **is monotone in k**

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m - t \right] \leq \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m \right] \leq \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m + t \right]$$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m + t \right] - \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m - t \right] \leq ?$$

i.i.d. $Y_1, \dots, Y_n \sim \text{Pois}(\frac{m}{n})$ **conditioning on $Y = \sum_{i=1}^n Y_i = k$**
is identically distributed as loads $X_1, \dots, X_n$ for k balls into n bins

$$m = n \ln n + cn$$

Thm: *i.i.d.* $Y_1, \dots, Y_n \sim \text{Pois}(\frac{m}{n})$ and $Y = \sum_{i=1}^n Y_i$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] = \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m \right] \pm o(1)$$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] = \sum_{k=m-t}^{m+t} \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = k \right] \Pr[Y = k] + o(1)$$

choose $t = \sqrt{2m \ln m}$

is monotone in k

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m - t \right] \leq \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m \right] \leq \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m + t \right]$$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m + t \right] - \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m - t \right]$$

$$\leq \Pr[2t \text{ balls hit an empty bin}] \leq \frac{2t}{n} = o(1)$$

$$m = n \ln n + cn$$

Thm: *i.i.d.* $Y_1, \dots, Y_n \sim \text{Pois}(\frac{m}{n})$ and $Y = \sum_{i=1}^n Y_i$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] = \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m \right] \pm o(1)$$

$$\begin{aligned} \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] &= \sum_{k=m-t}^{m+t} \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = k \right] \Pr[Y = k] + o(1) \\ &= (1 - o(1)) \left(\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m \right] \pm o(1) \right) + o(1) \end{aligned}$$

$$m = n \ln n + cn$$

Thm: *i.i.d.* $Y_1, \dots, Y_n \sim \text{Pois}(\frac{m}{n})$ and $Y = \sum_{i=1}^n Y_i$

$$\Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] = \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m \right] \pm o(1)$$

- loads of n bins receiving m balls: $X_1, \dots, X_n$
- $(X_1, \dots, X_n)$ is identically distributed as $(Y_1, \dots, Y_n \mid Y = m)$

$$\Pr \left[\bigwedge_{i=1}^n X_i > 0 \right] = \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \mid Y = m \right]$$

$$= \Pr \left[\bigwedge_{i=1}^n Y_i > 0 \right] \pm o(1)$$

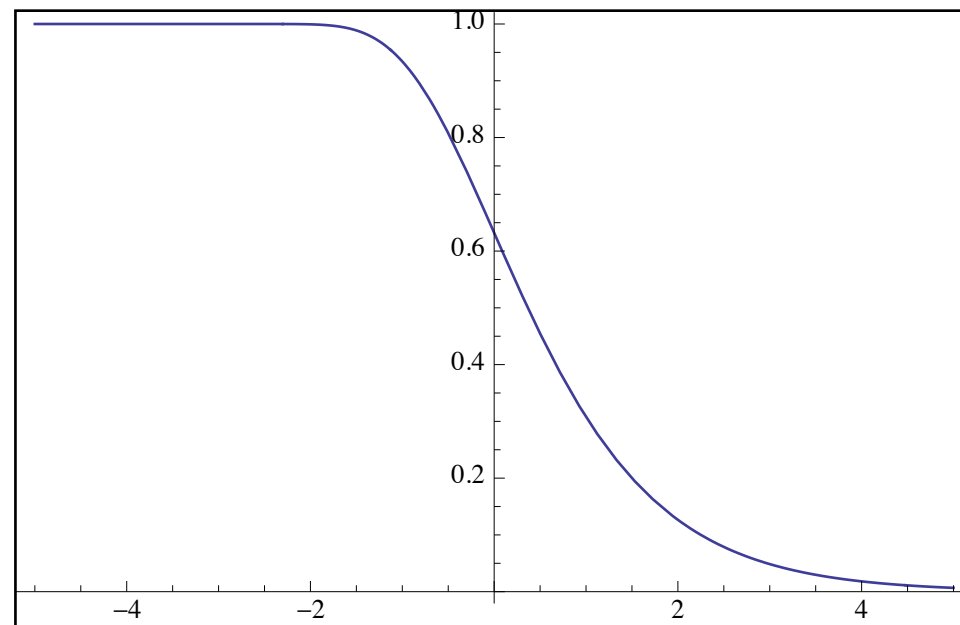
$$\rightarrow e^{-e^{-c}} \text{ as } n \rightarrow \infty$$

Coupon Collector

X : number of balls thrown to make
all the n bins nonempty

a sharp threshold:

$$\lim_{n \rightarrow \infty} \Pr[X \geq n \ln n + cn] = 1 - e^{-e^{-c}}$$



Poisson Approximation

- loads of n bins receiving m balls: $X_1, \dots, X_n$
- *i.i.d.* **Poisson** random variables $Y_1, \dots, Y_n \sim \text{Pois}(m/n)$

Theorem (Poisson Approximation): $\forall$ nonnegative function f

$$\mathbb{E} [f(X_1, \dots, X_n)] \leq e\sqrt{m} \cdot \mathbb{E} [f(Y_1, \dots, Y_n)]$$

$$\mathbb{E} [f(\vec{Y})] = \sum_{k=0}^{\infty} \mathbb{E} [f(\vec{Y}) \mid Y = k] \Pr[Y = k] \quad \text{where } Y = \sum_{i=1}^n Y_i \sim \text{Pois}(m)$$

$$\geq \mathbb{E} [f(\vec{Y}) \mid Y = m] \Pr[Y = m] = \mathbb{E} [f(\vec{X})] e^{-m} \frac{m^m}{m!} \geq \frac{1}{e\sqrt{m}} \mathbb{E} [f(\vec{X})]$$

$$\text{since } m! \leq e\sqrt{m} \left(\frac{m}{e}\right)^m$$

Poisson Approximation

- loads of n bins receiving m balls: $X_1, \dots, X_n$
- *i.i.d.* **Poisson** random variables $Y_1, \dots, Y_n \sim \text{Pois}(m/n)$

Theorem (Poisson Approximation): $\forall$ nonnegative function f

$$\mathbb{E} [f(X_1, \dots, X_n)] \leq e\sqrt{m} \cdot \mathbb{E} [f(Y_1, \dots, Y_n)]$$

Occupancy problem:

$$\Pr \left[\max_{1 \leq i \leq n} X_i < L \right] \leq e\sqrt{m} \Pr \left[\max_{1 \leq i \leq n} Y_i < L \right] = e\sqrt{m} (\Pr[Y_i < L])^n$$

$$(\text{when } m = n) \leq e\sqrt{n} (1 - \Pr[Y_i = L])^n = e\sqrt{n} \left(1 - \frac{1}{eL!} \right)^n$$

$$\leq \frac{1}{n} \quad \text{for } L = \frac{\ln n}{\ln \ln n}$$

Occupancy Problem

n balls are thrown into n bins.

X_i : number of balls in the i -th bin

Theorem:

With high probability, the maximum load is

$$\max_{1 \leq i \leq n} X_i = \Omega \left(\frac{\log n}{\log \log n} \right)$$