Probabilistic Inequalities and Examples

Lecture 3
January 22, 2019
## Outline

**Probabilistic Inequalities**

- Markov’s Inequality
- Chebyshev’s Inequality
- Bernstein-Chernoff-Hoeffding bounds
- Some examples
Part I

Inequalities
Massive randomness.. Is not that random.

Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k} \frac{1}{2^n}$.
Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k} \frac{1}{2^n}$. 

![Graph showing probability distribution for $n = 4$.]
Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k} \frac{1}{2^n}$.
Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k} \frac{1}{2^n}$. 
Consider flipping a fair coin \( n \) times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: \( k \) w.p. \( \binom{n}{k} \frac{1}{2^n} \).
Massive randomness.. Is not that random.

Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k} \frac{1}{2^n}$.
Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k}1/2^n$. 

![Graph showing binomial distribution for n=256]
Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k} \frac{1}{2^n}$.
Consider flipping a fair coin \( n \) times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: \( k \) w.p. \( \binom{n}{k} \frac{1}{2^n} \).
Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k} \frac{1}{2^n}$. 

![Graph showing the binomial distribution for $n = 2048$.]
Massive randomness.. Is not that random.

Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k}1^{1/2^n}$. 

![Probability distribution graph](#)
Massive randomness.. Is not that random.

Consider flipping a fair coin $n$ times independently, head gives 1, tail gives zero. How many 1s? Binomial distribution: $k$ w.p. $\binom{n}{k}1/2^n$. 

![Probability distribution graph](image-url)
Massive randomness.. Is not that random.

This is known as concentration of mass.
This is a very special case of the law of large numbers.
Informal statement of law of large numbers

For $n$ large enough, the middle portion of the binomial distribution looks like (converges to) the normal/Gaussian distribution.
Massive randomness.. Is not that random.

Intuitive conclusion

Randomized algorithm are unpredictable in the tactical level, but very predictable in the strategic level.
Intuitive conclusion
Randomized algorithm are unpredictable in the tactical level, but very predictable in the strategic level.

Use of well known inequalities in analysis.
Randomized **QuickSort**: A possible analysis

**Analysis**

- Random variable $Q = \#\text{comparisons}$ made by randomized **QuickSort** on an array of $n$ elements.
Randomized **QuickSort**: A possible analysis

**Analysis**

- Random variable $Q = \#\text{comparisons}$ made by randomized **QuickSort** on an array of $n$ elements.
- Suppose $\Pr[Q \geq 10n\log n] \leq c$. Also we know that $Q \leq n^2$. 
Randomized QuickSort: A possible analysis

Analysis

- Random variable $Q = \#\text{comparisons}$ made by randomized QuickSort on an array of $n$ elements.
- Suppose $\Pr[Q \geq 10n\log n] \leq c$. Also we know that $Q \leq n^2$.
- $E[Q] \leq 10n \log n + (n^2 - 10n \log n)c$. 
Randomized **QuickSort**: A possible analysis

### Analysis

- Random variable $Q = \#\textit{comparisons}$ made by randomized QuickSort on an array of $n$ elements.
- Suppose $\Pr\[Q \geq 10n\log n\] \leq c$. Also we know that $Q \leq n^2$.
- $\mathbb{E}[Q] \leq 10n \log n + (n^2 - 10n \log n)c$.

### Question:

How to find $c$, or in other words bound $\Pr\[Q \geq 10n \log n\]$?
Markov’s Inequality

Markov’s inequality

Let $X$ be a non-negative random variable over a probability space $(\Omega, \Pr)$. For any $a > 0$, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$. Equivalently, for any $t > 0$, $\Pr[X \geq t\mathbb{E}[X]] \leq 1/t$. 
Markov’s Inequality

Markov’s inequality

Let $X$ be a non-negative random variable over a probability space $(\Omega, \Pr)$. For any $a > 0$, $\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$. Equivalently, for any $t > 0$, $\Pr[X \geq t\mathbb{E}[X]] \leq 1/t$.

Proof:

\[
\begin{align*}
\mathbb{E}[X] &= \sum_{\omega \in \Omega} X(\omega) \Pr[\omega] \\
&= \sum_{\omega, \ 0 \leq X(\omega) < a} X(\omega) \Pr[\omega] + \sum_{\omega, \ X(\omega) \geq a} X(\omega) \Pr[\omega] \\
&\geq \sum_{\omega \in \Omega, \ X(\omega) \geq a} X(\omega) \Pr[\omega] \\
&\geq a \sum_{\omega \in \Omega, \ X(\omega) \geq a} \Pr[\omega] \\
&= a \Pr[X \geq a]
\end{align*}
\]
Markov’s Inequality

Let $X$ be a non-negative random variable over a probability space $(\Omega, \Pr)$. For any $a > 0$, $\Pr[X \geq a] \leq \frac{E[X]}{a}$. Equivalently, for any $t > 0$, $\Pr[X \geq tE[X]] \leq 1/t$. 
Markov’s Inequality

**Markov’s inequality**

Let $X$ be a non-negative random variable over a probability space $(Ω, Pr)$. For any $a > 0$, $Pr[X ≥ a] ≤ \frac{E[X]}{a}$. Equivalently, for any $t > 0$, $Pr[X ≥ tE[X]] ≤ 1/t$.

**Proof:**

\[
E[X] = \int_{0}^{\infty} zf_X(z)dz \\
\geq \int_{a}^{\infty} zf_X(z)dz \\
\geq a \int_{a}^{\infty} f_X(z)dz \\
= a Pr[X ≥ a]
\]
Markov’s Inequality: Proof by Picture

\[ \text{Area} = a \cdot p(x \geq a) \]

\[ \text{Area} = \sum x \cdot p(x = x) = E \left[ x \right] \]
Chebyshev’s Inequality: Variance

Variance

Given a random variable $X$ over probability space $(\Omega, \Pr)$, variance of $X$ is the measure of how much does it deviate from its mean value. Formally, $\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

Derivation

Define $Y = (X - \mathbb{E}[X])^2 = X^2 - 2X \mathbb{E}[X] + \mathbb{E}[X]^2$.

$$\text{Var}(X) = \mathbb{E}[Y]$$
$$= \mathbb{E}[X^2] - 2 \mathbb{E}[X] \mathbb{E}[X] + \mathbb{E}[X]^2$$
$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2$$
Chebyshev’s Inequality: Variance

**Independence**

Random variables $X$ and $Y$ are called mutually independent if

$$\forall x, y \in \mathbb{R}, \Pr[X = x \land Y = y] = \Pr[X = x] \Pr[Y = y]$$

**Lemma**

If $X$ and $Y$ are mutually independent, then

$$\operatorname{Var}(X + Y) = \operatorname{Var}(X) + \operatorname{Var}(Y).$$
### Independence

Random variables $X$ and $Y$ are called mutually independent if
\[
\forall x, y \in \mathbb{R}, \quad \Pr[X = x \land Y = y] = \Pr[X = x] \Pr[Y = y]
\]

### Lemma

If $X$ and $Y$ are independent random variables then
\[
\operatorname{Var}(X + Y) = \operatorname{Var}(X) + \operatorname{Var}(Y).
\]

### Lemma

If $X$ and $Y$ are mutually independent, then
\[
\operatorname{E}[XY] = \operatorname{E}[X] \operatorname{E}[Y].
\]
Chebyshev’s Inequality

If $\text{Var}X < \infty$, for any $a \geq 0$, $\text{Pr}[|X - E[X]| \geq a] \leq \frac{\text{Var}(X)}{a^2}$
Chebyshev’s Inequality

If $\text{Var}X < \infty$, for any $a \geq 0$, $\Pr[|X - \mathbb{E}[X]| \geq a] \leq \frac{\text{Var}(X)}{a^2}$

Proof.

$Y = (X - \mathbb{E}[X])^2$ is a non-negative random variable. Apply Markov’s Inequality to $Y$ for $a^2$.

$$
\Pr[Y \geq a^2] \leq \mathbb{E}[Y]/a^2 \iff \Pr[(X - \mathbb{E}[X])^2 \geq a^2] \leq \frac{\text{Var}(X)}{a^2}
$$

$$
\iff \Pr[|X - \mathbb{E}[X]| \geq a] \leq \frac{\text{Var}(X)}{a^2}
$$
Chebyshev’s Inequality

If \( \text{Var}X < \infty \), for any \( a \geq 0 \), \( \Pr[|X - \mathbb{E}[X]| \geq a] \leq \frac{\text{Var}(X)}{a^2} \)

Proof.

\( Y = (X - \mathbb{E}[X])^2 \) is a non-negative random variable. Apply Markov’s Inequality to \( Y \) for \( a^2 \).

\[
\Pr[Y \geq a^2] \leq \frac{\mathbb{E}[Y]}{a^2} \iff \Pr[(X - \mathbb{E}[X])^2 \geq a^2] \leq \frac{\text{Var}(X)}{a^2} \\
\iff \Pr[|X - \mathbb{E}[X]| \geq a] \leq \frac{\text{Var}(X)}{a^2}
\]

\( \Pr[X \leq \mathbb{E}[X] - a] \leq \frac{\text{Var}(X)}{a^2} \) AND \( \Pr[X \geq \mathbb{E}[X] + a] \leq \frac{\text{Var}(X)}{a^2} \)
Chebyshev’s Inequality

Given $a \geq 0$, $\Pr[|X - \mathbb{E}[X]| \geq a] \leq \frac{\text{Var}(X)}{a^2}$ equivalently for any $t > 0$, $\Pr[|X - \mathbb{E}[X]| \geq t\sigma_X] \leq \frac{1}{t^2}$ where $\sigma_X = \sqrt{\text{Var}(X)}$ is the standard deviation of $X$. 
Example: Random walk on the line

- Start at origin 0. At each step move left one unit with probability $1/2$ and move right with probability $1/2$.
- After $n$ steps how far from the origin?
Example: Random walk on the line

- Start at origin 0. At each step move left one unit with probability \(1/2\) and move right with probability \(1/2\).
- After \(n\) steps how far from the origin?

At time \(i\) let \(X_i\) be \(-1\) if move to left and \(1\) if move to right.

\[Y_n = \sum_{i=1}^{n} X_i\]

\[E[Y_n] = 0\] and \[\text{Var}(Y_n) = \sum_{i=1}^{n} \text{Var}(X_i) = n\]

By Chebyshev:

\[\text{Pr}[|Y_n| \geq t \sqrt{n}] \leq \frac{1}{t^2}\]
Example: Random walk on the line

- Start at origin 0. At each step move left one unit with probability $1/2$ and move right with probability $1/2$.
- After $n$ steps how far from the origin?

At time $i$ let $X_i$ be $-1$ if move to left and $1$ if move to right.
$Y_n$ position at time $n$
$Y_n = \sum_{i=1}^{n} X_i$

$E[Y_n] = 0$ and $Var(Y_n) = \sum_{i=1}^{n} Var(X_i) = n$
Example: Random walk on the line

- Start at origin $0$. At each step move left one unit with probability $1/2$ and move right with probability $1/2$.
- After $n$ steps how far from the origin?

At time $i$ let $X_i$ be $-1$ if move to left and $1$ if move to right.

$Y_n$ position at time $n$

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

$E[Y_n] = 0$ and $Var(Y_n) = \sum_{i=1}^{n} Var(X_i) = n$

By Chebyshev: $Pr[|Y_n| \geq t\sqrt{n}] \leq 1/t^2$
Chernoff Bound: Motivation

In many applications we are interested in $X$ which is sum of independent bounded random variables.

$$X = \sum_{i=1}^{k} X_i \text{ where } X_i \in [0, 1] \text{ or } [-1, 1] \text{ (normalizing)}$$

Chebyshev not strong enough. For random walk on line one can prove

$$\Pr[|Y_n| \geq t\sqrt{n}] \leq 2\exp(-t^2/2)$$
Chernoff Bound: Non-negative case

Lemma

Let $X_1, \ldots, X_k$ be $k$ independent binary random variables such that, for each $i \in [1, k]$, $E[X_i] = \Pr[X_i = 1] = p_i$. Let $X = \sum_{i=1}^{k} X_i$. Then $E[X] = \sum_i p_i$.

- **Upper tail bound:** For any $\mu \geq E[X]$ and any $\delta > 0$,

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

- **Lower tail bound:** For any $0 < \mu < E[X]$ and any $0 < \delta < 1$,

  $$\Pr[X \leq (1 - \delta)\mu] \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{1-\delta}}\right)^\mu$$
When $0 < \delta < 1$ an important regime of interest we can simplify.

**Lemma**

Let $X_1, \ldots, X_k$ be $k$ independent random variables such that, for each $i \in [1, k]$, $X_i$ equals 1 with probability $p_i$, and 0 with probability $(1 - p_i)$. Let $X = \sum_{i=1}^{k} X_i$ and $\mu = \mathbb{E}[X] = \sum_i p_i$. For any $0 < \delta < 1$, it holds that:

$$\Pr[|X - \mu| \geq \delta \mu] \leq 2e^{-\frac{\delta^2 \mu}{3}}$$

$$\Pr[X \geq (1 + \delta)\mu] \leq e^{-\frac{\delta^2 \mu}{3}} \quad \text{and} \quad \Pr[X \leq (1 - \delta)\mu] \leq e^{-\frac{\delta^2 \mu}{2}}$$
Lemma

Let $X_1, \ldots, X_k$ be $k$ independent random variables such that, for each $i \in [1, k]$, $X_i \in [-1, 1]$. 
Lemma

Let $X_1, \ldots, X_k$ be $k$ independent random variables such that, for each $i \in [1, k]$, $X_i \in [-1, 1]$. Let $X = \sum_{i=1}^{k} X_i$. For any $a > 0$,

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

When variables are not positive the bound depends on $n$ while in the non-negative case there is no dependence on $n$ (dimension-free)
Lemma

Let $X_1, \ldots, X_k$ be $k$ independent random variables such that, for each $i \in [1, k]$, $X_i \in [-1, 1]$. Let $X = \sum_{i=1}^{k} X_i$. For any $a > 0$,

$$
\Pr[|X - E[X]| \geq a] \leq 2\exp\left(\frac{-a^2}{2n}\right).
$$

When variables are not positive the bound depends on $n$ while in the non-negative case there is no dependence on $n$ (dimension-free).

Applying to random walk:

$$
\Pr[|Y_n| \geq t\sqrt{n}] \leq 2\exp(-t^2/2).
$$
Chernoff Bounds

Many variations and generalization that are useful in specific situations. See pointers on course webpage.
Part II

Ball and Bins
Balls and Bins

- \( m \) balls and \( n \) bins
- Each ball thrown independently and uniformly in a bin
- Want to understand properties of bin loads
- Fundamental problem with many applications
Balls and Bins

- $m$ balls and $n$ bins
- Each ball thrown independently and uniformly in a bin
- Want to understand properties of bin loads
- Fundamental problem with many applications

$Z_{ij}$ indicator for ball $i$ falling into bin $j$

$X_j = \sum_{i=1}^{m} Z_{ij}$ is number of balls in bin $j$

$\sum_{j=1}^{n} Z_{ij} = 1$ deterministically

$\mathbb{E}[Z_{ij}] = 1/n$ for all $i, j$, and hence $\mathbb{E}[X_j] = m/n$ for each bin $j$
**Maximum load**

**Question:** Suppose we throw $n$ balls into $n$ bins. What is the expectation of the *maximum* load?

\[
Y = \max_{j=1}^{n} X_j
\]

Theorem

Let $Y = \max_{j=1}^{n} X_j$ be the maximum load. Then

\[
\Pr[Y > \frac{10 \ln n}{\ln \ln n}] < \frac{1}{n^2} \quad \text{(high probability)}
\]

and hence

\[
E[Y] = O\left(\frac{\ln n}{\ln \ln n}\right)
\]

One can also show that

\[
E[Y] = \Theta\left(\frac{\ln n}{\ln \ln n}\right)
\]

Proof technique: combine Chernoff bound and union bound which is powerful and general template
Maximum load

**Question**: Suppose we throw \( n \) balls into \( n \) bins. What is the expectation of the *maximum* load?

**Theorem**

Let \( Y = \max_{j=1}^{n} X_j \) be the maximum load. Then

\[
\Pr[Y > 10 \ln n / \ln \ln n] < 1/n^2 \quad \text{(high probability)} \quad \text{and hence}
\]

\[
E[Y] = O(\ln n / \ln \ln n).
\]

One can also show that \( E[Y] = \Theta(\ln n / \ln \ln n) \).
**Maximum load**

**Question**: Suppose we throw $n$ balls into $n$ bins. What is the expectation of the *maximum* load?

**Theorem**

Let $Y = \max_{j=1}^n X_j$ be the maximum load. Then

\[
\Pr[Y > 10 \ln n / \ln \ln n] < 1/n^2 \quad \text{(high probability) and hence}
\]

\[
E[Y] = O(\ln n / \ln \ln n).
\]

One can also show that $E[Y] = \Theta(\ln n / \ln \ln n)$.

Proof technique: combine Chernoff bound and union bound which is powerful and general template
Focus on bin 1 without loss of generality since bins are symmetric. Simplifying notation \( X = \sum_i Z_i \) where \( X \) is load of bin 1 and \( Z_i \) is indicator of ball \( i \) falling in bin.

- Want to know \( \Pr[X \geq 10 \ln n / \ln \ln n] \)
- \( \mu = E[X] = 1 \)
- \( (1 + \delta) = 10 \ln n / \ln \ln n \). We are in large \( \delta \) setting
- Apply the Chernoff upper tail bound:

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

- Calculate/simplify and see that \( \Pr[X \geq 10 \ln n / \ln \ln n] \leq 1/n^3 \)
For each bin $j$, $\Pr[X_j \geq 10 \ln n / \ln \ln n] \leq 1/n^3$

Let $A_j$ be event that $X_j \geq 10 \ln n / \ln \ln n$

By union bound

$$\Pr[\bigcup_j A_j] \leq \sum_j \Pr[A_j] \leq n \cdot 1/n^3 \leq 1/n^2.$$ 

Hence, with probability at least $(1 - 1/n^2)$ no bin has load more than $10 \ln n / \ln \ln n$. 

Let $Y = \max_j X_j$. Hence

$$\mathbb{E}[Y] \leq (1 - 1/n^2)(10 \ln n / \ln \ln n) + (1/n^2)n.$$
For each bin $j$, $\Pr[X_j \geq 10 \ln n \ln \ln n] \leq 1/n^3$

Let $A_j$ be event that $X_j \geq 10 \ln n / \ln \ln n$

By union bound

$$\Pr[\bigcup_j A_j] \leq \sum_j \Pr[A_j] \leq n \cdot 1/n^3 \leq 1/n^2.$$ 

Hence, with probability at least $(1 - 1/n^2)$ no bin has load more than $10 \ln n / \ln \ln n$.

Let $Y = \max_j X_j$. $Y \leq n$. Hence

$$E[Y] \leq (1 - 1/n^2)(10 \ln n / \ln \ln n) + (1/n^2)n.$$
From a ball’s perspective

Consider a ball $i$. How many other balls fall into the same bin as $i$?
From a ball’s perspective

Consider a ball $i$. How many other balls fall into the same bin as $i$?

- Ball $i$ is thrown first wlog. And lands in some bin $j$.
- Then the other $n - 1$ balls are thrown.
- Now bin $j$ is fixed. Hence expected load on bin $j$ is $(1 - 1/n)$.
- What is variance? What is a high probability bound?
Part III

Approximate Median
Approximate median

- **Input:** \( n \) distinct numbers \( a_1, a_2, \ldots, a_n \) and \( 0 < \epsilon < 1/2 \)
- **Output:** A number \( x \) from input such that
\[
(1 - \epsilon)n/2 \leq \text{rank}(x) \leq (1 + \epsilon)n/2
\]
Approximate median

- **Input:** $n$ distinct numbers $a_1, a_2, \ldots, a_n$ and $0 < \epsilon < 1/2$
- **Output:** A number $x$ from input such that $(1 - \epsilon)n/2 \leq \text{rank}(x) \leq (1 + \epsilon)n/2$

Algorithm:
- Sample with replacement $k$ numbers from $a_1, a_2, \ldots, a_n$
- Output median of the sampled numbers

Theorem
For any $0 < \epsilon < 1/2$ and $0 < \delta < 1$, if $k = O\left(\frac{1}{\epsilon^2 \log(1/\delta)}\right)$, the algorithm outputs an $\epsilon$-approximate median with probability at least $1 - \delta$. 
Approximate median

- **Input:** \( n \) distinct numbers \( a_1, a_2, \ldots, a_n \) and \( 0 < \epsilon < 1/2 \)
- **Output:** A number \( x \) from input such that \((1 - \epsilon)n/2 \leq \text{rank}(x) \leq (1 + \epsilon)n/2\)

**Algorithm:**
- Sample with replacement \( k \) numbers from \( a_1, a_2, \ldots, a_n \)
- Output median of the sampled numbers

**Theorem**

For any \( 0 < \epsilon < 1/2 \) and \( 0 < \delta < 1 \), if \( k = O\left(\frac{1}{\epsilon^2 \log(1/\delta)}\right) \), the algorithm outputs an \( \epsilon \)-approximate median with probability at least \((1 - \delta)\).
Approximate median

- Let $S$ be random sample chosen by algorithm
- Imagine sorting the numbers
- Split numbers into $L$ (left), $M$ (middle), and $R$ (right)
- $M = \{y \mid (1 - \epsilon)n/2 \leq \text{rank}(y) \leq (1 + \epsilon)n/2\}$
- Algorithm makes a mistake only if $|S \cap L| \geq k/2$ or $|S \cap R| \geq k/2$. Otherwise it will output a number from $M$. 

Analysis:

Let $Y = |S \cap L|$. What is $E[Y]$?

$Y = \sum_{i=1}^{k} X_i$ where $X_i$ is indicator of sample $i$ falling in $L$.

Hence $E[Y] = k(1 - \epsilon)/2$.

Use Chernoff bound to argue that $\Pr[Y \geq k/2] \leq \delta/2$ if $k = 10\epsilon^2 \log(1/\delta)$.
Approximate median

- Let $S$ be random sample chosen by algorithm
- Imagine sorting the numbers
- Split numbers into $L$ (left), $M$ (middle), and $R$ (right)
- $M = \{y \mid (1 - \epsilon)n/2 \leq \text{rank}(y) \leq (1 + \epsilon)n/2\}$
- Algorithm makes a mistake only if $|S \cap L| \geq k/2$ or $|S \cap R| \geq k/2$. Otherwise it will output a number from $M$.

Analysis:

- Let $Y = |S \cap L|$? What is $E[Y]$?
- $Y = \sum_{i=1}^{k} X_i$ where $X_i$ is indicator of sample $i$ falling in $L$. Hence $E[Y] = k(1 - \epsilon)/2$
- Use Chernoff bound to argue that $\Pr[Y \geq k/2] \leq \delta/2$ if $k = \frac{10}{\epsilon^2} \log(1/\delta)$. 
Approximate median

Analysis:

- Let $Y = |S \cap L|$? What is $E[Y]$?

- $Y = \sum_{i=1}^{k} X_i$ where $X_i$ is indicator of sample $i$ falling in $L$. Hence $E[Y] = k(1 - \epsilon)/2$

- Use Chernoff bound to argue that $Pr[Y \geq k/2] \leq \delta/2$ if $k = \frac{10}{\epsilon^2} \log(1/\delta)$.

- By union bound at most $\delta$ probability that $|S \cap L| \geq k/2$ or $|S \cap R| \geq k/2$.

- Hence with $(1 - \delta)$ probability median of $S$ is an $\epsilon$-approximate median.
Part IV

Randomized QuickSort (Contd.)
Randomized QuickSort: Recall

**Input:** Array $A$ of $n$ numbers. **Output:** Numbers in sorted order.

**Randomized QuickSort**

1. Pick a pivot element *uniformly at random* from $A$.
2. Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and the pivot itself.
3. Recursively sort the subarrays, and concatenate them.
Randomized **QuickSort**: Recall

**Input**: Array $A$ of $n$ numbers. **Output**: Numbers in sorted order.

**Randomized QuickSort**

1. Pick a pivot element *uniformly at random* from $A$.
2. Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and the pivot itself.
3. Recursively sort the subarrays, and concatenate them.

**Note**: On every input randomized QuickSort takes $O(n \log n)$ time in expectation. On every input it may take $\Omega(n^2)$ time with some small probability.
Randomized **QuickSort**: Recall

**Input:** Array $A$ of $n$ numbers. **Output:** Numbers in sorted order.

### Randomized **QuickSort**

1. Pick a pivot element *uniformly at random* from $A$.
2. Split array into 3 subarrays: those smaller than pivot, those larger than pivot, and the pivot itself.
3. Recursively sort the subarrays, and concatenate them.

**Note:** On every input randomized **QuickSort** takes $O(n \log n)$ time in expectation. On every input it may take $\Omega(n^2)$ time with some small probability.

**Question:** With what probability it takes $O(n \log n)$ time?
Informal Statement

Random variable $Q(A) = \#$ comparisons done by the algorithm.

We will show that $\Pr[Q(A) \leq 32n \ln n] \geq 1 - \frac{1}{n^3}$.
**Informal Statement**

Random variable $Q(A) = \#$ comparisons done by the algorithm.

We will show that $\Pr[Q(A) \leq 32n \ln n] \geq 1 - \frac{1}{n^3}$.

If $n = 100$ then this gives $\Pr[Q(A) \leq 32n \ln n] \geq 0.99999$. 
Randomized QuickSort: High Probability Analysis

Informal Statement
Random variable $Q(A) = \#$ comparisons done by the algorithm.
We will show that $\Pr[Q(A) \leq 32n \ln n] \geq 1 - \frac{1}{n^3}$.

Outline of the proof
- If depth of recursion is $k$ then $Q(A) \leq kn$.
- Prove that depth of recursion $\leq 32 \ln n$ with high probability. Which will imply the result.
Randomized QuickSort: High Probability Analysis

**Informal Statement**

Random variable $Q(A) = \#$ comparisons done by the algorithm.

We will show that $\Pr[Q(A) \leq 32n \ln n] \geq 1 - \frac{1}{n^3}$.

**Outline of the proof**

- If depth of recursion is $k$ then $Q(A) \leq kn$.
- Prove that depth of recursion $\leq 32 \ln n$ with high probability. Which will imply the result.
  1. Focus on a single element. Prove that it “participates” in $> 32 \ln n$ levels with probability at most $\frac{1}{n^4}$.
  2. By union bound, any of the $n$ elements participates in $> 32 \ln n$ levels with probability at most
Informal Statement

Random variable $Q(A) = \#$ comparisons done by the algorithm.

We will show that $\Pr[Q(A) \leq 32n \ln n] \geq 1 - \frac{1}{n^3}$.

Outline of the proof

- If depth of recursion is $k$ then $Q(A) \leq kn$.
- Prove that depth of recursion $\leq 32 \ln n$ with high probability. Which will imply the result.

1. Focus on a single element. Prove that it “participates” in $> 32 \ln n$ levels with probability at most $\frac{1}{n^4}$.

2. By union bound, any of the $n$ elements participates in $> 32 \ln n$ levels with probability at most $\frac{1}{n^3}$. 
Randomized **QuickSort**: High Probability Analysis

**Informal Statement**

Random variable \( Q(A) = \# \) comparisons done by the algorithm.

We will show that \( \Pr[Q(A) \leq 32n \ln n] \geq 1 - \frac{1}{n^3} \).

**Outline of the proof**

- If depth of recursion is \( k \) then \( Q(A) \leq kn \).
- Prove that depth of recursion \( \leq 32 \ln n \) with high probability. Which will imply the result.
  1. Focus on a single element. Prove that it “participates” in \( > 32 \ln n \) levels with probability at most \( \frac{1}{n^4} \).
  2. By union bound, any of the \( n \) elements participates in \( > 32 \ln n \) levels with probability at most \( \frac{1}{n^3} \).
  3. Therefore, all elements participate in \( \leq 32 \ln n \) w.p. \( (1 - \frac{1}{n^3}) \).
Randomized **QuickSort**: High Probability Analysis

- If $k$ levels of recursion then $kn$ comparisons.
Randomized **QuickSort**: High Probability Analysis

- If \( k \) levels of recursion then \( kn \) comparisons.
- Fix an element \( s \in A \). We will track it at each level.
- Let \( S_i \) be the partition containing \( s \) at \( i^{th} \) level.
- \( S_1 = A \) and \( S_k = \{s\} \).
If $k$ levels of recursion then $kn$ comparisons.

Fix an element $s \in A$. We will track it at each level.

Let $S_i$ be the partition containing $s$ at $i^{th}$ level.

$S_1 = A$ and $S_k = \{s\}$.

We call $s$ lucky in $i^{th}$ iteration, if balanced split:

$|S_{i+1}| \leq (3/4)|S_i|$ and $|S_i \setminus S_{i+1}| \leq (3/4)|S_i|$. 

Randomized QuickSort: High Probability Analysis

- If $k$ levels of recursion then $kn$ comparisons.
- Fix an element $s \in A$. We will track it at each level.
- Let $S_i$ be the partition containing $s$ at $i^{th}$ level.
- $S_1 = A$ and $S_k = \{s\}$.
- We call $s$ lucky in $i^{th}$ iteration, if balanced split:
  $|S_{i+1}| \leq (3/4)|S_i|$ and $|S_i \setminus S_{i+1}| \leq (3/4)|S_i|$.
- If $\rho = \#$ lucky rounds in first $k$ rounds, then $|S_k| \leq (3/4)^\rho n$. 
Randomized **QuickSort**: High Probability Analysis

- If $k$ levels of recursion then $kn$ comparisons.
- Fix an element $s \in A$. We will track it at each level.
- Let $S_i$ be the partition containing $s$ at $i^{th}$ level.
- $S_1 = A$ and $S_k = \{s\}$.
- We call $s$ lucky in $i^{th}$ iteration, if balanced split:
  \[ |S_{i+1}| \leq \frac{3}{4}|S_i| \text{ and } |S_i \setminus S_{i+1}| \leq \frac{3}{4}|S_i|. \]
- If $\rho = \#\text{lucky rounds in first } k \text{ rounds}$, then
  \[ |S_k| \leq (3/4)^\rho n. \]
- For $|S_k| = 1$, $\rho = 4 \ln n \geq \log_{4/3} n$ suffices.
How may rounds before $4 \ln n$ lucky rounds?

- $X_i = 1$ if $s$ is lucky in $i^{th}$ iteration.
How may rounds before $4 \ln n$ lucky rounds?

- $X_i = 1$ if $s$ is lucky in $i^{th}$ iteration.
- Observation: $X_1, \ldots, X_k$ are independent variables.
- $\Pr[X_i = 1] = \frac{1}{2}$ Why?
How may rounds before $4 \ln n$ lucky rounds?

- $X_i = 1$ if $s$ is lucky in $i^{th}$ iteration.
- **Observation:** $X_1, \ldots, X_k$ are independent variables.
- $\Pr[X_i = 1] = \frac{1}{2}$ Why?
- Clearly, $\rho = \sum_{i=1}^{k} X_i$. Let $\mu = \mathbb{E}[\rho] = \frac{k}{2}$.
How may rounds before $4 \ln n$ lucky rounds?

- $X_i = 1$ if $s$ is lucky in $i^{th}$ iteration.
- **Observation:** $X_1, \ldots, X_k$ are independent variables.
- $\Pr[X_i = 1] = \frac{1}{2}$  Why?

Clearly, $\rho = \sum_{i=1}^{k} X_i$. Let $\mu = E[\rho] = \frac{k}{2}$.

Set $k = 32 \ln n$ and $\delta = \frac{3}{4}$. $(1 - \delta) = \frac{1}{4}$. 

How may rounds before $4 \ln n$ lucky rounds?

- $X_i = 1$ if $s$ is lucky in $i^{th}$ iteration.
- **Observation:** $X_1, \ldots, X_k$ are independent variables.
- $\Pr[X_i = 1] = \frac{1}{2}$ Why?

Clearly, $\rho = \sum_{i=1}^{k} X_i$. Let $\mu = \mathbb{E}[\rho] = \frac{k}{2}$.

Set $k = 32 \ln n$ and $\delta = \frac{3}{4}$. $(1 - \delta) = \frac{1}{4}$.

Probability of NOT getting $4 \ln n$ lucky rounds out of $32 \ln n$ rounds is,
How may rounds before $4 \ln n$ lucky rounds?

- $X_i = 1$ if $s$ is lucky in $i^{th}$ iteration.
- **Observation:** $X_1, \ldots, X_k$ are independent variables.
- $\Pr[X_i = 1] = \frac{1}{2}$  
  Why?

Clearly, $\rho = \sum_{i=1}^{k} X_i$. Let $\mu = \mathbb{E}[\rho] = \frac{k}{2}$.

Set $k = 32 \ln n$ and $\delta = \frac{3}{4}$. $(1 - \delta) = \frac{1}{4}$.

Probability of NOT getting $4 \ln n$ lucky rounds out of $32 \ln n$ rounds is,

$$\Pr[\rho \leq 4 \ln n] = \Pr[\rho \leq k/8] = \Pr[\rho \leq (1 - \delta)\mu]$$
How may rounds before $4 \ln n$ lucky rounds?

- $X_i = 1$ if $s$ is lucky in $i^{th}$ iteration.
- **Observation:** $X_1, \ldots, X_k$ are independent variables.
- $\Pr[X_i = 1] = \frac{1}{2}$ Why?
- Clearly, $\rho = \sum_{i=1}^{k} X_i$. Let $\mu = \mathbb{E}[\rho] = \frac{k}{2}$.
- Set $k = 32 \ln n$ and $\delta = \frac{3}{4}$, $(1 - \delta) = \frac{1}{4}$.

Probability of NOT getting $4 \ln n$ lucky rounds out of $32 \ln n$ rounds is,

$$\Pr[\rho \leq 4 \ln n] = \Pr[\rho \leq k/8] = \Pr[\rho \leq (1 - \delta)\mu]$$

$$(\text{Chernoff}) \leq e^{-\frac{\delta^2 \mu}{2}} = e^{-\frac{9k}{64}} = e^{-4.5 \ln n} \leq \frac{1}{n^4}$$
n input elements. Probability that depth of recursion in QuickSort is at most $32 \ln n$ is at most $\frac{1}{n^4} \times n = \frac{1}{n^3}$. 
Randomized **QuickSort** w.h.p. Analysis

- n input elements. Probability that depth of recursion in **QuickSort** \( \geq 32 \ln n \) is at most \( \frac{1}{n^4} \times n = \frac{1}{n^3} \).

**Theorem**

*With high probability (i.e., \( 1 - \frac{1}{n^3} \)) the depth of the recursion of **QuickSort** is \( \leq 32 \ln n \). Due to \( n \) comparisons in each level, with high probability, the running time of **QuickSort** is \( O(n \ln n) \).*
Randomized **QuickSort** w.h.p. Analysis

- $n$ input elements. Probability that depth of recursion in **QuickSort** > $32 \ln n$ is at most $\frac{1}{n^4} \times n = \frac{1}{n^3}$.

**Theorem**

*With high probability (i.e., $1 - \frac{1}{n^3}$) the depth of the recursion of **QuickSort** is $\leq 32 \ln n$. Due to $n$ comparisons in each level, with high probability, the running time of **QuickSort** is $O(n \ln n)$.***