Probabilistic Inequalities and Examples

Lecture 3

January 22, 2019

Outline

Probabilistic Inequalities

Markov's Inequality

Chebyshev's Inequality

Bernstein-Chernoff-Hoeffding bounds

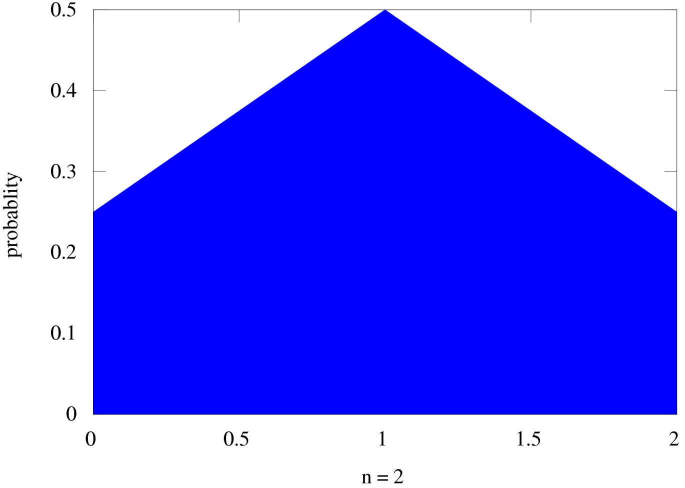
Some examples

Part I

Inequalities

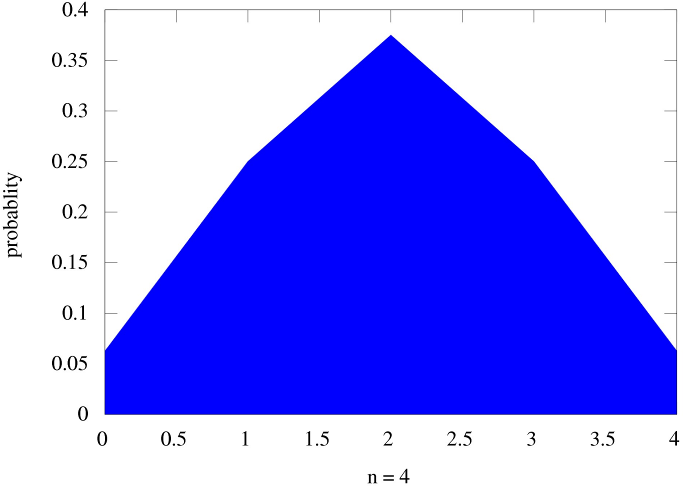
Massive randomness.. Is not that random.

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



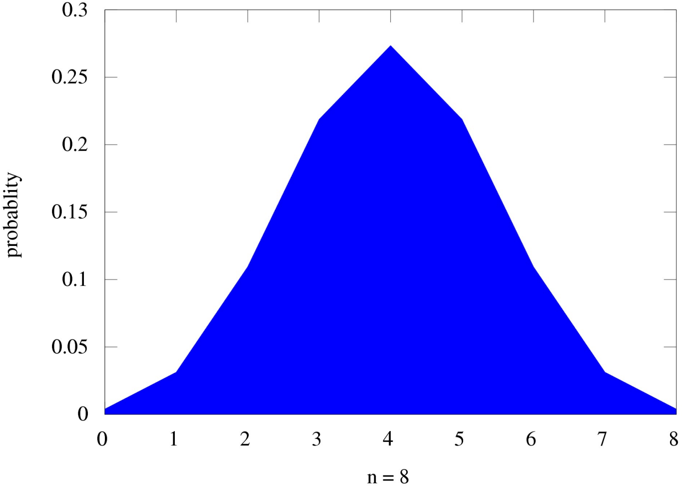
Massive randomness.. Is not that random.

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



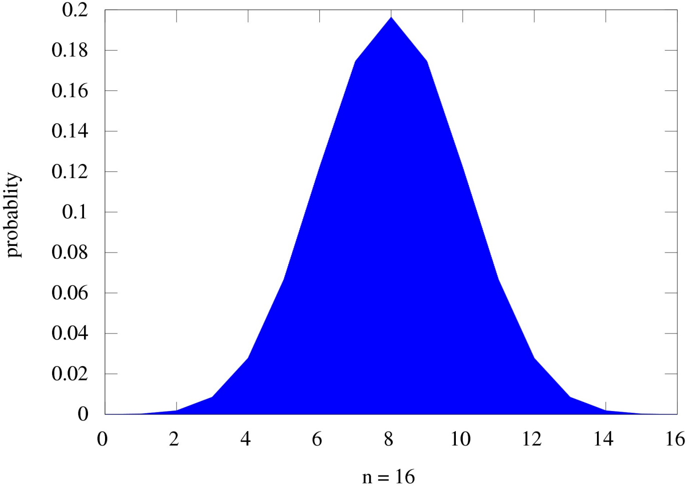
Massive randomness.. Is not that random.

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



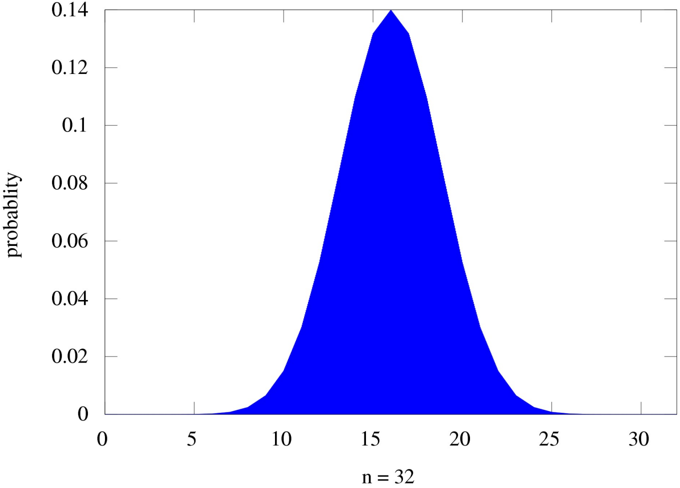
Massive randomness.. Is not that random.

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



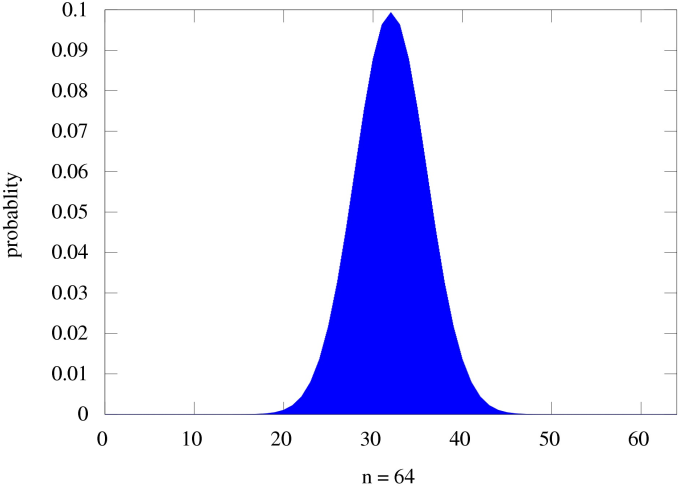
Massive randomness.. Is not that random.

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



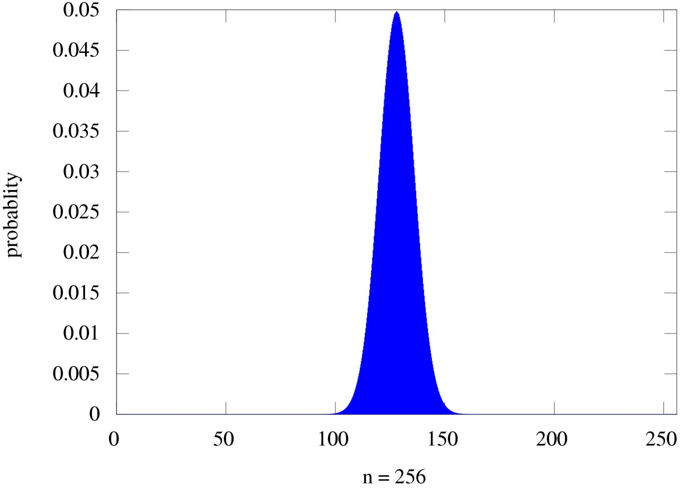
Massive randomness.. Is not that random.

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



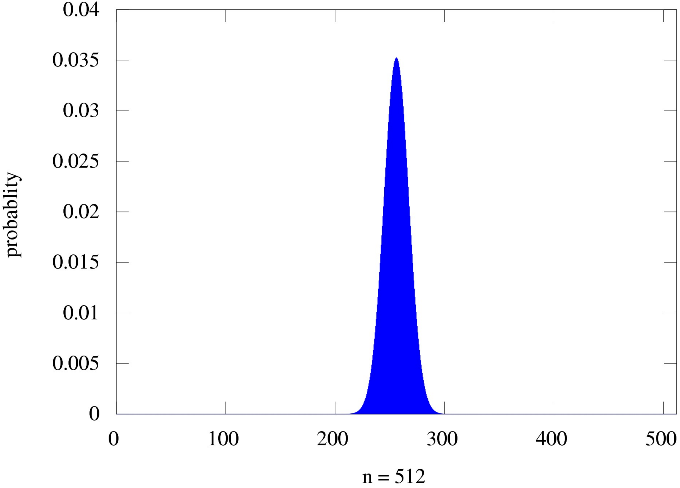
Massive randomness.. Is not that random.

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



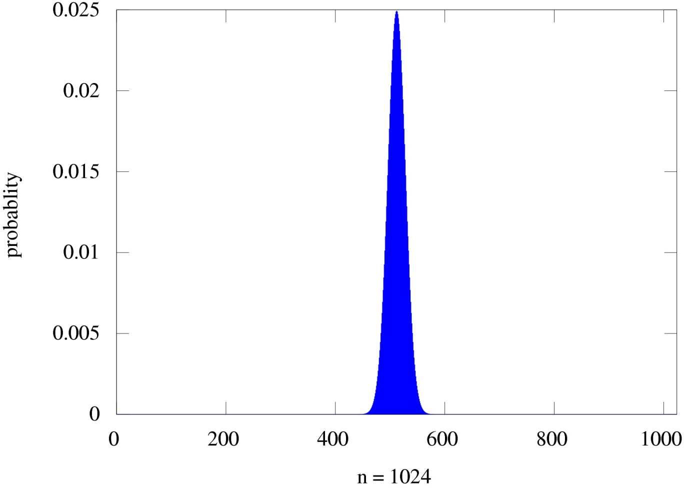
Massive randomness.. Is not that random.

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



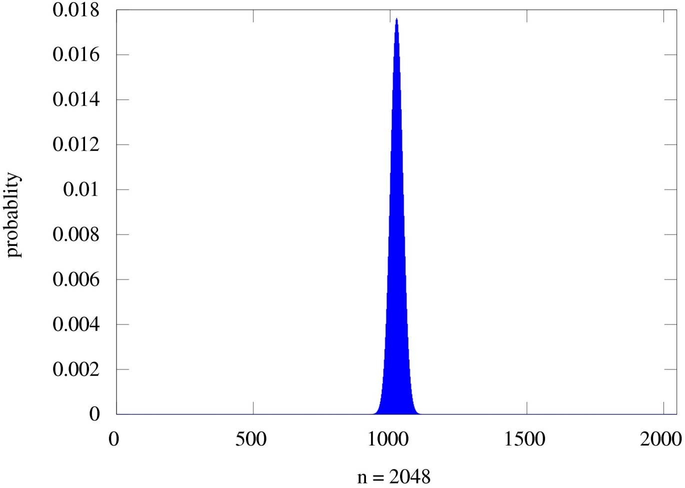
Massive randomness.. Is not that random.

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



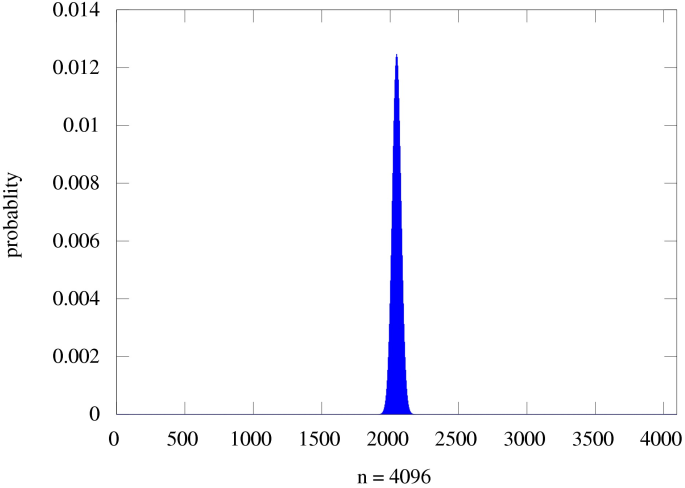
Massive randomness.. Is not that random.

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



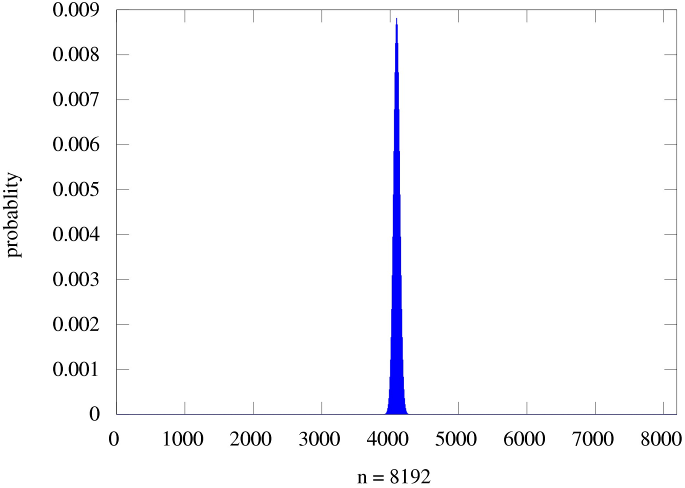
Massive randomness.. Is not that random.

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

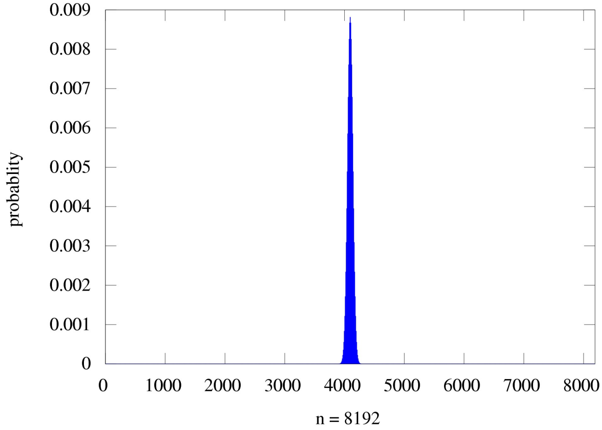


Massive randomness.. Is not that random.

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



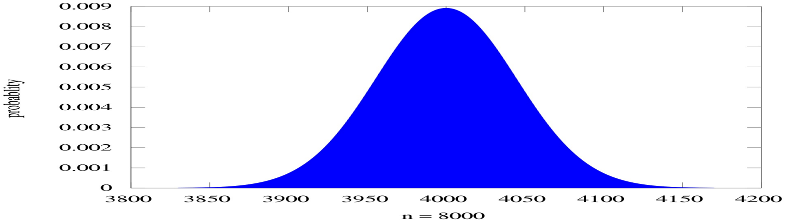
Massive randomness.. Is not that random.



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.

Massive randomness.. Is not that random.

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 $Q = \#comparisons$ made by randomized **QuickSort** on an array of n elements.

Randomized **QuickSort**: A possible analysis

Analysis

- Random variable $Q = \#comparisons$ made by randomized **QuickSort** on an array of n elements.
- Suppose $\Pr[Q \geq 10 n lg n] \leq c$. Also we know that $Q \leq n^2$.

Randomized **QuickSort**: A possible analysis

Analysis

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

Randomized **QuickSort**: A possible analysis

Analysis

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

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 $a > 0$, $\Pr[X \geq a] \leq \frac{E[X]}{a}$. 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 $(\Omega, \Pr)$. For any $a > 0$, $\Pr[X \geq a] \leq \frac{E[X]}{a}$. Equivalently, for any $t > 0$, $\Pr[X \geq tE[X]] \leq 1/t$.

Proof:

$$
\begin{aligned}
\mathbf{E}[X] &= \sum_{\omega \in \Omega} X(\omega) \Pr[\omega] \\
&= \sum_{\omega,\ 0 \leq X(\omega) < a} X(\omega) \Pr[\omega] + \sum_{\omega,\ X(\omega) \geq a} X(\omega) \Pr[\omega] \\
&\geq \sum_{\omega \in \Omega,\ X(\omega) \geq a} X(\omega) \Pr[\omega] \\
&\geq a \sum_{\omega \in \Omega,\ X(\omega) \geq a} \Pr[\omega] \\
&= a \Pr[X \geq a]
\end{aligned}
$$

Markov's Inequality

Markov's inequality

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

Markov's Inequality

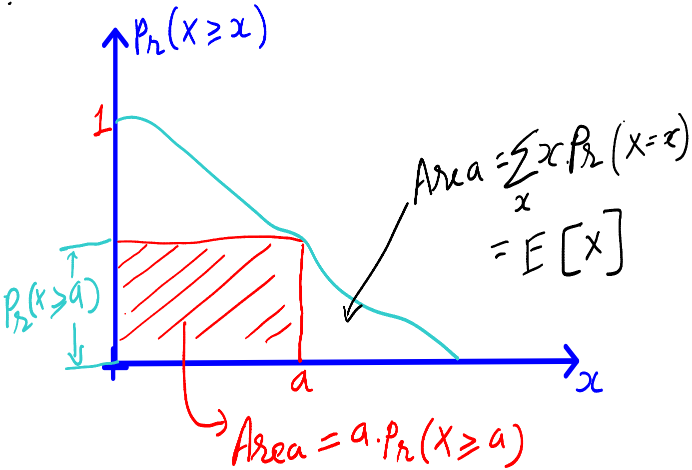
Markov's inequality

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

Proof:

$$\begin{aligned} \mathbf{E}[X] &= \int_0^\infty z f_X(z) dz \\ &\geq \int_a^\infty z f_X(z) dz \\ &\geq a \int_a^\infty f_X(z) dz \\ &= a \,\mathbf{Pr}[X \geq a] \end{aligned}$$

Markov's Inequality: Proof by Picture



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, $Var(X) = \mathbf{E}\left[(X - \mathbf{E}[X])^2\right] = \mathbf{E}\left[X^2\right] - \mathbf{E}[X]^2$

Derivation

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

$$\begin{aligned} Var(X) &= \mathbf{E}[Y] \\ &= \mathbf{E}\left[X^2\right] - 2\,\mathbf{E}[X]\,\mathbf{E}[X] + \mathbf{E}[X]^2 \\ &= \mathbf{E}\left[X^2\right] - \mathbf{E}[X]^2 \end{aligned}$$

Chebyshev's Inequality: Variance

Independence

Random variables X and Y are called mutually independent if

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

Lemma

If X and Y are independent random variables then $Var(X + Y) = Var(X) + Var(Y)$.

Chebyshev's Inequality: Variance

Independence

Random variables X and Y are called mutually independent if

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

Lemma

If X and Y are independent random variables then $Var(X + Y) = Var(X) + Var(Y)$.

Lemma

If X and Y are mutually independent, then $\mathbf{E}[XY] = \mathbf{E}[X]\,\mathbf{E}[Y]$.

Chebyshev's Inequality

Chebyshev's Inequality

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

Chebyshev's Inequality

Chebyshev's Inequality

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

Proof.

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

$$\Pr\left[Y \geq a^2\right] \leq {}^{\mathbf{E}[Y]}/_{a^2} \Leftrightarrow \Pr\left[(X - \mathbf{E}[X])^2 \geq a^2\right] \leq {}^{\mathit{Var}(X)}/_{a^2}$$
$$\Leftrightarrow \Pr[|X - \mathbf{E}[X]| \geq a] \leq {}^{\mathit{Var}(X)}/_{a^2}$$

$\square$

Chebyshev's Inequality

Chebyshev's Inequality

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

Proof.

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

$$\begin{aligned}\Pr\left[Y \geq a^2\right] \leq \mathbf{E}[Y]/a^2 &\Leftrightarrow \Pr\left[(X - \mathbf{E}[X])^2 \geq a^2\right] \leq Var(X)/a^2 \\ &\Leftrightarrow \Pr[|X - \mathbf{E}[X]| \geq a] \leq Var(X)/a^2\end{aligned}$$

$\square$

$\Pr[X \leq \mathbf{E}[X] - a] \leq Var(X)/a^2$ AND $\Pr[X \geq \mathbf{E}[X] + a] \leq Var(X)/a^2$

Chebyshev's Inequality

Chebyshev's Inequality

Given $a \geq 0$, $\mathbf{Pr}[|X - \mathbf{E}[X]| \geq a] \leq \frac{Var(X)}{a^2}$ equivalently for any $t > 0$, $\mathbf{Pr}[|X - \mathbf{E}[X]| \geq t\sigma_X] \leq \frac{1}{t^2}$ where $\sigma_X = \sqrt{Var(X)}$ 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 position at time n
$Y_n = \sum_{i=1}^{n} X_i$

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

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

Example: Random walk on the line

- Start at origin $\mathbf{0}$. At each step move left one unit with probability $\mathbf{1/2}$ and move right with probability $\mathbf{1/2}$.
- 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
$Y_n = \sum_{i=1}^{n} X_i$

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

By Chebyshev: $\mathbf{Pr}\left[|Y_n| \geq t\sqrt{n}\right] \leq 1/t^2$

Chernoff Bound: Motivation

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

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

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

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

Chernoff Bound: Non-negative case

Lemma

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

- *Upper tail bound: For any* $\mu \geq \mathbf{E}[X]$ *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* $0 < \mu < \mathbf{E}[X]$ *and any* $0 < \delta < 1$,

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

Chernoff Bound: Non-negative case, simplifying

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 $\mu = \mathbf{E}[X] = \sum_i p_i$. For any $0 < \delta < 1$, it holds that:

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

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

Chernoff Bound: general

Lemma

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

Chernoff Bound: general

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* $a > 0$,

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

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

Chernoff Bound: general

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 $a > 0$,

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

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:

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

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

Maximum load

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

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* $\Pr[Y > 10 \ln n / \ln \ln n] < 1/n^2$ *(high probability) and hence* $\mathbf{E}[Y] = O(\ln n / \ln \ln n)$.

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

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* $\Pr[Y > 10 \ln n / \ln \ln n] < 1/n^2$ *(high probability) and hence* $\mathbf{E}[Y] = O(\ln n / \ln \ln n)$.

One can also show that $\mathbf{E}[Y] = \Theta(\ln n / \ln \ln n)$.
Proof technique: combine Chernoff bound and union bound which is powerful and general template

Maximum Load

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 $\Pr[X \geq 10 \ln n / \ln\ln n]$
- $\mu = \mathrm{E}[X] = 1$
- $(1+\delta) = 10 \ln n / \ln\ln n$. We are in large δ 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 $\Pr[X \geq 10 \ln n / \ln\ln n] \leq 1/n^3$

Maximum load

- 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

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

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

Maximum load

- 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

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

- 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

$$\mathbf{E}[Y] \leq (1 - 1/n^2)(10 \ln n / \ln \ln n) + (1/n^2) n.$$

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 $(1-\epsilon)n/2 \leq rank(x) \leq (1+\epsilon)n/2$

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 $(1 - \epsilon)n/2 \leq rank(x) \leq (1 + \epsilon)n/2$

Algorithm:

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

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 $(1 - \epsilon)n/2 \leq rank(x) \leq (1 + \epsilon)n/2$

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 $k = O(\frac{1}{\epsilon^2} \log(1/\delta))$, the algorithm outputs an ϵ-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)
- $M = \{y \mid (1-\epsilon)n/2 \leq rank(y) \leq (1+\epsilon)n/2\}$
- 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.

Approximate median

- Let S be random sample chosen by algorithm
- Imagine sorting the numbers
- Split numbers into L (left), M (middle), and R (right)
- $M = \{y \mid (1-\epsilon)n/2 \leq rank(y) \leq (1+\epsilon)n/2\}$
- 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 $\mathbf{E}[Y]$?
- $Y = \sum_{i=1}^{k} X_i$ where X_i is indicator of sample i falling in L. Hence $\mathbf{E}[Y] = k(1-\epsilon)/2$
- Use Chernoff bound to argue that $\mathbf{Pr}[Y \geq k/2] \leq \delta/2$ if $k = \frac{10}{\epsilon^2}\log(1/\delta)$.

Approximate median

Analysis:

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

Randomized **QuickSort**: High Probability Analysis

Informal Statement

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

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

Randomized **QuickSort**: High Probability Analysis

Informal Statement

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

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

If $n = 100$ then this gives $\Pr[Q(A) \leq 32 n \ln n] \geq 0.99999$.

Randomized **QuickSort**: High Probability Analysis

Informal Statement

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

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

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 $\Pr[Q(A) \le 32n \ln n] \ge 1 - 1/n^3$.

Outline of the proof

- If depth of recursion is k then $Q(A) \le kn$.
- Prove that depth of recursion $\le 32 \ln n$ with high probability. Which will imply the result.
  1. Gocus on a single element. Prove that it "participates" in $> 32 \ln n$ levels with probability at most $1/n^4$.
  2. By union bound, any of the n elements participates in $> 32 \ln n$ levels with probability at most

Randomized QuickSort: High Probability Analysis

Informal Statement

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

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

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. Gocus on a single element. Prove that it "participates" in $> 32 \ln n$ levels with probability at most $1/n^4$.
  2. By union bound, any of the n elements participates in $> 32 \ln n$ levels with probability at most $1/n^3$.

Randomized **QuickSort**: High Probability Analysis

Informal Statement

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

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

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. Gocus on a single element. Prove that it "participates" in $> 32 \ln n$ levels with probability at most $1/n^4$.
  2. By union bound, any of the n elements participates in $> 32 \ln n$ levels with probability at most $1/n^3$.
  3. Therefore, all elements participate in $\leq 32 \ln n$ w.p. $(1 - 1/n^3)$.

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 $S_k = \{s\}$.

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

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: $|S_{i+1}| \leq (3/4)|S_i|$ and $|S_i \setminus S_{i+1}| \leq (3/4)|S_i|$.
- If $\rho =$#lucky rounds in first k rounds, then $|S_k| \leq (3/4)^{\rho} n$.
- 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 $\mu = \mathbf{E}[\rho] = \frac{k}{2}$.

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 $\mu = \mathbf{E}[\rho] = \frac{k}{2}$.
- 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 $\mu = \mathbf{E}[\rho] = \frac{k}{2}$.
- 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 $\mu = \mathbf{E}[\rho] = \frac{k}{2}$.
- 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,

$$\begin{aligned} \Pr[\rho \leq 4 \ln n] &= \Pr[\rho \leq k/8] \\ &= \Pr[\rho \leq (1 - \delta)\mu] \end{aligned}$$

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 $\mu = \mathbf{E}[\rho] = \frac{k}{2}$.
- 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,

$$\begin{aligned}
\Pr[\rho \leq 4 \ln n] &= \Pr[\rho \leq k/8] \\
&= \Pr[\rho \leq (1-\delta)\mu] \\
(\textit{Chernoff}) \quad &\leq e^{\frac{-\delta^2 \mu}{2}} \\
&= e^{-\frac{9k}{64}} \\
&= e^{-4.5 \ln n} \leq \frac{1}{n^4}
\end{aligned}$$

Randomized **QuickSort** w.h.p. Analysis

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

Randomized **QuickSort** w.h.p. Analysis

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

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 $\frac{1}{n^4} * n = \frac{1}{n^3}$.

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* $O(n \ln n)$.