**Randomized Algorithms**

- What are they?
  - Quicksort
  - Chernoff bounds

**What are they?**

- Algorithms with access to random source.
  - E.g., can flip coins/roll dice/etc.

- Funky: on same input on different executions, randomized algorithm may:
  - Run for different # steps
  - Produce different results

  [Depending on outcomes of coin flips]

- Flavors:
  - Monte Carlo
    - Always runs in polynomial time
    - \( \text{prob}[\text{output correct}] > \text{high} \)
  - Las Vegas
    - \( \text{Prob}[\text{output correct}] = 1 \)
    - Runs in expected polynomial time
- True Randomness?

- **Assumption for this class:**

  algorithm has access to subroutine that, given \( R \), outputs a uniformly random number \( r \in \{1, \ldots, R\} \)

- **In practice:** pseudo-random number generator

- **Fundamentally?** philosophical/religious belief

  Einstein: “... I, at any rate, am convinced that He does not throw dice.”
  letter to Max Born 1926

Q: Are randomized better than deterministic algs? 

- A1: not a fair comparison, as randomized algorithms are allowed to make errors, take longer

- A2: in practice, they are faster and simpler

- A3: in theory, only polynomial gain* in running time under complexity-theoretic assumptions
QuickSort  [Hoare 1962]

- comparison sort  (like Mergesort, unlike Radix Sort)
- divide & conquer
  - work on divide step
  - no work on combine step
- in place: \( O(1) \) extra space
- practical (with tuning)
- flavors:
  - basic: worst-case \( O(n^2) \)
    expected \( O(n \log n) \) for random input
  - randomized: expected \( O(n \log n) \) for all inputs
  - deterministic: \( O(n \log n) \) worst-case
    in practice, slower than randomized

Algorithm

INPUT: Array \( A \) of \( n \) elements

1. If \( n=1 \), stop; \( A \) is sorted
2. divide:
   - pick some \( i \in \{1, 2, \ldots, n\} \)
     - basic: \( i=1 \)
     - randomized: \( i \sim \text{uniform}\{1, \ldots, n\} \)
     - deterministic: \( i \) such that \( A[i] \) is median of \( A \)

\( x=A[i] \) is the pivot element
- partition $A$ into $\text{elts} \leq x$, $x$, $\text{elts} > x$

\[ \begin{array}{c}
\geq x \leq x \leq x \leq x \leq x \leq x \\
i \end{array} \Rightarrow \begin{array}{c}
\leq x \leq x \leq x \leq x \leq x \\
i \end{array} \]

② conquer: recursively sort $\text{elts} \leq x$ ($A[1..i-1]$)
recursively sort $\text{elts} > x$ ($A[i+1..n]$)

③ combine: do nothing :

E.g., execution of basic Quicksort

\[ 3 \ 1 \ 8 \ 2 \ 6 \ 7 \ 5 \]

\[ \begin{array}{c}
1 \ 2 \\
2 \ 6 \ 7 \ 5 \\
5 \ 7 \\
\end{array} \]

sorted array: $1 \ 2 \ 3 \ 5 \ 6 \ 7 \ 8$

Exercise: basic Quicksort may take $O(n^2)$

Exercise 2: no matter what strategy for picking pivot is used: output of Quicksort is sorted array and runtime is $O(n \log n)$

Exercise 3: if median is used as pivot Quicksort runs in $O(n \log n)$ time.
Analysis of Randomized Quicksort

- Want to show expected running time is $O(n \log n)$
- Will show something stronger, namely:

$$\Pr \left[ \text{Quicksort takes longer than } c \cdot n \log n \right] \leq \frac{1}{n}$$

the Bad Event

$$\Rightarrow \mathbb{E}[\text{running time}] \leq \left(1 - \frac{1}{n}\right) c \cdot n \log n + \frac{1}{n} \cdot c \cdot n^2 = O(n \log n)$$

Even under bad event runtime is bounded by $c \cdot n^2$

- Intuition: (c) Suppose we’re extremely lucky and every choice of pivot results in a 50-50 split of the elements

$$\begin{array}{c}
\text{A} \\
\downarrow n \\
\text{els \leq x} \quad \text{x} \quad \text{els > x}
\end{array} \quad \Rightarrow \\
\Rightarrow \frac{n}{2} \rightarrow \text{pivot} \quad \leftarrow \frac{n}{2}$$

Then $T(n) \leq 2 \cdot T\left(\frac{n}{2}\right) + O(n) = O(n \log n)$

Probability [single split 50%-50%] = $O\left(\frac{1}{n}\right)$ (unlikely)
(ii) why exact split? what if we always get a 10% - 90% (or more balanced) split?

\[
\begin{array}{c}
\text{then worst case:} \\
\text{\emph{O}(logn) levels} \\
\end{array}
\]

\[
\begin{array}{c}
\frac{n}{10} \\
\frac{g_n}{10^2} \\
\frac{g_n}{10^2} \\
\vdots \\
\frac{g^2_n}{10^2} \\
\end{array}
\]

Total work: \( O(\log n) \times O(n) = O(n \cdot \log n) \)

\[
Pr \left[ \text{single split is 10\% - 90\% or better} \right] = \frac{8}{10} \quad :)\]

problem? \[
Pr \left[ \text{all splits are 10\% - 90\% or better} \right] = \left( \frac{8}{10} \right)^0 = 0
\]

\text{Idea: It is not crucial that all splits are 10\% - 90\%. Let's formalize this idea...}
For every element $\alpha$ of the array and recursive depth $t$ of quicksort, define random variable $X_{\alpha, t} = \begin{cases} 0, & \text{if in the subarray where } \alpha \text{ belongs in depth } t, \text{ quicksort chose a good pivot} \\ 1, & \text{o.w.} \end{cases}$

Note: if quicksort finished before reaching rec. depth $t$, or if $\alpha$ does not appear in any sub-array at depth $t$ (because in some earlier depth the subarray of $\alpha$ had only $\alpha$ in it), we still define $X_{\alpha, t}$ and take it to be $X_{\alpha, t} = \begin{cases} 0, & \text{w.p. } 0.8 \\ 1, & \text{w.p. } 0.2 \end{cases}$

So for all $\alpha$, for all $t$: $\Pr[X_{\alpha, t} = 1] = 0.2$

Observe also that $\forall \alpha, X_{\alpha, 1}, X_{\alpha, 2}, \ldots$ are independent random variables

(in contrast the variables $\{X_{\alpha, t}\}$ for a fixed $t$ are not independent)
\( \diamond \) Let \( T = c \cdot \log n \) (with \( \log \) I mean \( \log_e \))
\[ c : \text{some constant to be set later} \]
\[ E \left[ \sum_{t=1}^{T} X_{\alpha,t} \right] = 0.2 \cdot T \]

\[ Q: \quad \Pr \left[ \sum_{t=1}^{T} X_{\alpha,t} \geq \gamma \cdot E \left[ \sum_{t=1}^{T} X_{\alpha,t} \right] \right] \quad \gamma: \text{constant to be set later} \]
\[ \Rightarrow \Pr \left[ \text{Bin}(T, 0.2) \geq \gamma \cdot 0.2 \cdot T \right] \quad ? \]

\[ A1: \quad \sum_{\lambda=0.2 \cdot 8^{T} \lambda}^{T} (\lambda) \cdot 0.2^{\lambda} \cdot 0.8^{T-\lambda} \]
\[ \Rightarrow \text{accurate but not usable} \]

\[ A2: \quad \text{Central Limit Theorem:} \]
\[ \text{as } T \to \infty: \text{Bin}(T, p) \to \text{Normal}(T \cdot p, T \cdot p(1-p)) \]
\[ \Rightarrow \Pr \left[ \text{Bin}(T, 0.2) \geq \gamma \cdot 0.2 \cdot T \right] \approx \Pr \left[ \text{Normal}(0.2T, 0.16T) \geq \gamma \cdot 0.2 \cdot T \right] \]

Problem: \( \times \) holds in the limit as \( T \to \infty \) so not usable directly either
A3: Chernoff Bounds

**Theorem (Chernoff):** Suppose $X_1, X_2, \ldots, X_n \in [0, 1]$ are independent random variables. Then $\forall \theta \in (0, 1]$:

$$
\Pr \left[ \sum_{i=1}^{n} X_i > (1+\theta) \mathbb{E}[X_i] \right] \leq e^{-\frac{\theta^2 \mathbb{E}[X_i]}{2}}
$$

and

$$
\Pr \left[ \sum_{i=1}^{n} X_i < (1-\theta) \mathbb{E}[X_i] \right] \leq e^{-\frac{\theta^2 \mathbb{E}[X_i]}{2}}
$$

**Proof:** later

Using Chernoff Bound:

$$
\Pr \left[ \sum_{t=1}^{T} X_{\alpha,t} \geq \gamma \cdot \mathbb{E} \left[ \sum_{t=1}^{T} X_{\alpha,t} \right] \right] \leq e^{-\frac{4}{3} \mathbb{E} \left[ \sum_{t=1}^{T} X_{\alpha,t} \right] \cdot (\gamma-1)}
$$

$$
\leq e^{-\frac{1}{15} \cdot T \cdot (\gamma-1)}
$$

$$
\leq e^{-\frac{1}{15} \cdot \frac{1}{n} \cdot c \cdot (\gamma-1)}
$$

if we chose $c \cdot (\gamma-1) \geq 30$

(e.g. $c=30, \gamma=2$ or $c=60, \gamma=1.5$ etc.)

$$
\leq \frac{1}{n^2}
$$

(\forall \gamma)
Union bound: If $E_1, E_2, \ldots, E_n$ are arbitrary events then
\[ \Pr\left[ E_1 \text{ or } E_2 \text{ or } \ldots \text{ or } E_n \right] \leq \sum_{i=1}^{n} \Pr\left[ E_i \right] \]

Since (*) holds for all $\alpha$, the union bound gives
\[ \Pr\left[ \exists \alpha \text{ s.t. } \sum_{t=1}^{T} X_{\alpha, t} \geq 0.2T \right] \leq n \cdot \frac{1}{n^2} = \frac{1}{n} \]

\[ \Rightarrow \Pr\left[ \forall \alpha : \sum_{t=1}^{T} X_{\alpha, t} < 0.2T \right] \geq 1 - \frac{1}{n} \]

This is a good event.

Under this event: $\forall \alpha$, if we trace the sub-arrays containing $\alpha$ inside the recursion tree of the quicksort execution, at least $T - 0.2T = (1 - \frac{4}{5})T$ of these sub-arrays shrunk by a 10%-30% or better split.

Hence, elements of subarray containing $\alpha$ at depth $t \leq \left( \frac{9}{10} \right)^{t} \cdot n = \left( \frac{9}{10} \right)^{t} \cdot n$

As long as $(1 - \frac{4}{5}) \cdot \log_{0.9} < 1$
e.g. $\approx 30$, $\approx 2$ works.
\[
\begin{align*}
\{c_1 \text{ means that there is no subarray as it can't contain any elements}\} \\
\Rightarrow \forall \lambda: \text{no subarray containing } \lambda \text{ at depth } T \\
(\text{i.e. } \lambda \text{ became lonely at an earlier depth}) \\
\Rightarrow \text{ i.e. quicksort finished earlier than depth } T = c \cdot \log n \\
\Rightarrow \text{ total runtime } O(n) \times T = O(n \cdot \log n)
\end{align*}
\]

Choice of constants \(c\) and \(\gamma\): for the above to go through the constants need to satisfy:
\[
\begin{align*}
&c \cdot (\gamma - 1) \geq 30 \\
&\left(1 - \frac{\alpha}{5}\right) \cdot c \cdot \log 0.9 < -1
\end{align*}
\]
Setting \(c = 30, \gamma = 2\) satisfies both with this choice of constants:

- the good event happens with probability \(\geq 1 - \frac{1}{n}\)
- under good event quicksort takes \(O(n \cdot \log n)\)

Final Remark: Is there anything special with the choice of a 10%-90% split?

A: No using 1%-99% or 49%-51% or more generally \(\Theta \%-\left(1-\Theta\right)\%\) for any constant \(\Theta\) would also work with exact same proof as long as we choose \(c\) to be a large enough constant
**Back To Chernoff Bound**

Chernoff: Suppose $X_1, X_2, \ldots, X_n \in [0, 1]$ are independent random variables. Then $\forall \theta \in (0, 1]$:

$$\Pr \left[ \sum X_i > (1+\theta) \mathbb{E}[X_i] \right] \leq e^{-\frac{\theta^2 \mathbb{E}[X_i]}{3}}$$

and

$$\Pr \left[ \sum X_i < (1-\theta) \mathbb{E}[X_i] \right] \leq e^{-\frac{\theta^2 \mathbb{E}[X_i]}{2}}$$

**Proof of Chernoff:**

\[ \text{\textbullet \hspace{1em} for convenience set: } p_i = \mathbb{E}[X_i], i = 1, \ldots, n \]
\[ \mu = \frac{\sum p_i}{n} \]
\[ P = \frac{\mu}{n} \]
\[ Y = \sum X_i \]

\[ \text{\textbullet \hspace{1em} we'll only bound } \Pr \left[ Y \geq (1+\theta)\mu \right] \text{ (upper tail) } \]
\[ \text{the lower tail is treated similarly} \]
Proof comprises basic idea + some calculus

\[ \Pr \left[ Y \geq (1 + \theta) \mu \right] = \Pr \left[ t \cdot Y \geq t \cdot (1 + \theta) \mu \right] \]

\[ = \Pr \left[ e^{t \cdot Y} \geq e^{t \cdot (1 + \theta) \mu} \right] \]

\[ \leq \frac{\mathbb{E} \left[ e^{t \cdot Y} \right]}{e^{t \cdot (1 + \theta) \mu}} \]

Markov's Inequality:

If \( X \) is positive random variable then for all \( \alpha > 0 \):

\[ \Pr \left[ X > \alpha \right] \leq \frac{\mathbb{E} \left[ X \right]}{\alpha} \]

in independence

\[ \frac{x + d_2 + \ldots + d_n}{n} \geq \sqrt[n]{x \cdot d_2 \cdot \ldots \cdot d_n} \]

\[ \leq \left( \frac{\sum_{i=1}^{n} \tfrac{t \cdot \pi_i}{1 - \pi_i}}{n} \right)^n \]

\[ = e^{-t(1+\theta)\mu} \cdot (e^{t \cdot p + 1 - p})^n \quad (**) \]
Bound \( (\rho) \) true for all \( t < 0 \). Minimized by setting
\[
 t = \log \frac{(1+\rho)(1-p)}{(1-p-\rho p)}
\]
and the bound becomes:
\[
\Pr \left[ Y \geq (1+\rho)\mu \right] \leq e^{-n \cdot H_p \left( \frac{(1+\rho)\mu}{n} \right)} \quad \text{(a.k.a. the entropic form of Chernoff bound)}
\]

where
\[
H_p(x) = x \cdot \log \frac{x}{p} + (1-x) \cdot \log \frac{1-x}{1-p}
\]
is the "relative entropy of \( x \) with respect to \( p \)
\[
\leq e^{-\frac{\rho^2 \mu^2}{3}}
\]
with a little calculus
use: \( \ln(1+x) < x \), \( \forall x > 0 \)
Example:

- Suppose $X_1, X_2, \ldots, X_n$ are independent outcomes of a fair coin, i.e. $E[X_1] = E[X_2] = \ldots = E[X_n] = \frac{1}{2}$

- $Y = \sum X_i$: total number of heads
  $E[Y] = \frac{n}{2}$, $\text{Var}(Y) = \frac{n}{4}$, $\text{std dev} = \sqrt{\frac{n}{2}}$

- Q1: Probability $Y$ is a constant factor away from its mean?

  A1: Directly from Chernoff
  
  $$\Pr\left[\left| Y - E[Y] \right| > 6 \cdot E[Y] \right] \leq 2 \cdot e^{-\frac{1}{3} \frac{6^2 \cdot n}{2}}$$

  exponentially small probability

  $\Rightarrow$ with probability $\geq 1 - 2 e^{-\frac{1}{3} \frac{6^2 \cdot n}{2}}$, $Y = E[Y] (1 \pm \theta)$

  often say w.v.h.p (with very high probability) for events that hold with probability 1 - exponentially small such as this one
Q2: Probability $Y$ is a few standard deviations away from its mean?

\[ \forall \delta: \Pr \left[ |Y - \mathbb{E}Y| > \delta \cdot \frac{\mathbb{n}}{2} \right] = \]

\[ \Pr \left[ |Y - \mathbb{E}Y| > \frac{\delta}{\sqrt{n}} \cdot \frac{n}{2} \right] \leq 2 \cdot e^{-\frac{1}{3} \left( \frac{\delta^2}{n} \right)} = 2 \cdot e^{-\frac{1}{6} \delta^2} \]

by Chernoff

Setting $\delta = 1$ gives $2 \cdot e^{-\frac{1}{6}}$

$\delta = \frac{1}{3}$ gives $2 \cdot e^{-\frac{1}{9}}$

$\delta = \frac{1}{5}$ gives $2 \cdot e^{-\frac{1}{5}}$

$\delta = \frac{1}{7}$ gives $2 \cdot e^{-\frac{1}{7}}$

$\delta = \sqrt{\log n}$ gives $2 \cdot e^{-\frac{1}{6} \log n} = 2 \cdot \frac{1}{n^{c/6}}$

$\Rightarrow$ with prob. $\geq 1 - \frac{2}{n^{c/6}}$, $Y = \mathbb{E}Y + \frac{\mathbb{n}}{2} \cdot \sqrt{\log n}$

often say w.h.p. (with high probability) for events that hold with probability $1 - (polynomially small)$
Conclusion: If \( n \) coins are tossed, the number of heads is:

- \( \frac{n}{2} \pm O(\ln \cdot \log n) \), with high probability
- \( \frac{n}{2} \pm 0.1 \cdot n \), with very high probability

\[ \uparrow \]

can replace this with any constant and will still have w.v.l.h.p.