Class 3: Randomized Algorithms - Expected Run Time
Online Class

- Attend synchronous class (if possible), participate, ask questions

On Zoom:
- Use the largest display you have. Your phone is not a good choice
- Join Zoom with video ON and audio Muted
- View options: Fit to Window
- To participate:
  - Use Raise Hand in Reactions
  - Unmute and talk
  - Chat to Everyone or Host.
Example: Finding the $k$-Smallest Element

Procedure Order($S, k$);

**Input:** A set $S$, an integer $k \leq |S| = n$.

**Output:** The $k$ smallest element in the set $S$.

1. If $|S| = k = 1$ return $S$.
2. Choose a random element $y$ uniformly from $S$.
3. Compare all elements of $S$ to $y$. Let $S_{\leq}(y) = \{x \in S \mid x \leq y\}$ and $S_{>}(y) = \{x \in S \mid x > y\}$.
4. If $k \leq |S_{\leq}(y)|$ return Order($S_{\leq}(y), k$) else return Order($S_{>}(y), k - |S_{\leq}(y)|$).

**Theorem**

1. The algorithm always returns the $k$-smallest element in $S$
2. The algorithm performs $O(n)$ comparisons in expectation.
Random Variable

**Definition**

A random variable \( X \) on a sample space \( \Omega \) is a real-valued function on \( \Omega \); that is, \( X : \Omega \rightarrow \mathbb{R} \). A discrete random variable is a random variable that takes on only a finite or countably infinite number of values.

Discrete random variable \( X \) and real value \( a \): the event "\( X = a \)" represents the set \( \{ s \in \Omega : X(s) = a \} \).

\[
\Pr(X = a) = \sum_{s \in \Omega : X(s) = a} \Pr(s)
\]

In the analysis of algorithm Order(\( S, k \)), the sample space is the set of possible choices of \( y \). The number of comparisons performed is a random variable \( T(S, k) \) that is a function of \( y \).
**Expectation**

**Definition**

The **expectation** of a discrete random variable $X$, denoted by $E[X]$, is given by

$$E[X] = \sum_i i \Pr(X = i),$$

where the summation is over all values in the range of $X$. The expectation is finite if $\sum_i |i| \Pr(X = i)$ converges; otherwise, the expectation is unbounded.

The expectation (or mean or average) is a weighted sum over all possible values of the random variable.
Linearity of Expectation

**Theorem**

For any two random variables $X$ and $Y$ and constants $a, b, c$,

Example: Finding the $k$-Smallest Element

Procedure Order($S, k$);

**Input:** A set $S$, an integer $k \leq |S| = n$.

**Output:** The $k$ smallest element in the set $S$.

1. If $|S| = k = 1$ return $S$.
2. Choose a random element $y$ uniformly from $S$.
3. Compare all elements of $S$ to $y$. Let $S_{\leq}(y) = \{x \in S \mid x \leq y\}$ and $S_{>}(y) = \{x \in S \mid x > y\}$.
4. If $k \leq |S_{\leq}(y)|$ return Order($S_{\leq}(y), k$) else return Order($S_{>}(y), k - |S_{\leq}(y)|$).

**Theorem**

1. The algorithm always returns the $k$-smallest element in $S$.
2. The algorithm performs $O(n)$ comparisons in expectation.
Proof

We say that a call to $\text{Order}(S, k)$ was successful if the random element was in the middle $1/3$ of the set $S$. A call is successful with probability $1/3$.

**Lemma**

*The algorithms terminate after no more than $\log_{3/2} n$ successful calls, where the $i$-th success call was to $\text{Order}(S_i, k)$, with $|S_i| \leq (2/3)^{i-1} n$.***

**Proof.**

After the $i$-th successful call the size of the set $S$ is bounded by $n(2/3)^i$. Thus, we need at most $\log_{3/2} n$ successful calls. □
Let $E = \text{call to Order}(S, k)$ was successful.
$t = \text{number of calls till } E$.

\[
\mathbb{E}[T(S, k)] \leq |S| + \mathbb{E}[T(S, k)]\Pr(\bar{E}) + \mathbb{E}[T(S', k)]\Pr(E)
\]

\[
\mathbb{E}[T(S, k)] \leq \mathbb{E}[t]|S| + \mathbb{E}[T(S', k)], \quad |S'| \leq \frac{2}{3}|S|
\]

For $|S| = n$, \[
\mathbb{E}[T(S, k)] \leq \sum_{j=0}^{\log_{3/2} n} \mathbb{E}[t]n \left(\frac{2}{3}\right)^j
\]

What is $\mathbb{E}[t]$?
The Geometric Distribution

Definition

A geometric random variable $X$ with parameter $p$ is given by the following probability distribution on $n = 1, 2, \ldots$.

$$\Pr(X = n) = (1 - p)^{n-1} p.$$ 

Example: repeatedly draw independent Bernoulli random variables with parameter $p > 0$ until we get a 1. Let $X$ be number of trials up to and including the first 1. Then $X$ is a geometric random variable with parameter $p$. 
Memoryless Distribution

Lemma

For a geometric random variable with parameter $p$ and $n > 0$,

$$\Pr(X = n + k \mid X > k) = \Pr(X = n).$$

Proof.

$$\Pr(X = n + k \mid X > k) = \frac{\Pr((X = n + k) \cap (X > k))}{\Pr(X > k)}$$

$$= \frac{\Pr(X = n + k)}{\Pr(X > k)} = \frac{(1 - p)^{n+k-1} p}{\sum_{i=k}^{\infty} (1 - p)^i p}$$

$$= \frac{(1 - p)^{n+k-1} p}{(1 - p)^k} = (1 - p)^{n-1} p = \Pr(X = n).$$
 Conditional Expectation

Definition

\[ E[Y \mid Z = z] = \sum_y y \Pr(Y = y \mid Z = z), \]

where the summation is over all \( y \) in the range of \( Y \).

Lemma

For any random variables \( X \) and \( Y \),

\[ E[X] = E_y\{E_x[X \mid Y]\} = \sum_y \Pr(Y = y)E[X \mid Y = y], \]

where the sum is over all values in the range of \( Y \).
Let $X$ be a geometric random variable with parameter $p$.
Let $Y = 1$ if the first trial is a success, $Y = 0$ otherwise.

$$E[X] = \Pr(Y = 0)E[X \mid Y = 0] + \Pr(Y = 1)E[X \mid Y = 1]$$
$$= (1 - p)E[X \mid Y = 0] + pE[X \mid Y = 1].$$

If $Y = 0$ let $Z$ be the number of trials after the first one.

$$E[X] = (1 - p)E[Z + 1] + p \cdot 1 = (1 - p)E[Z] + 1$$
Lemma

Let $X$ be a discrete random variable that takes on only non-negative integer values. Then

$$E[X] = \sum_{i=1}^{\infty} \Pr(X \geq i).$$

Proof.

$$\sum_{i=1}^{\infty} \Pr(X \geq i) = \sum_{i=1}^{\infty} \sum_{j=i}^{\infty} \Pr(X = j)$$

$$= \sum_{j=1}^{\infty} \sum_{i=1}^{j} \Pr(X = j)$$

$$= \sum_{j=1}^{\infty} j \Pr(X = j) = E[X].$$
For a geometric random variable $X$ with parameter $p$,

$$\Pr(X \geq i) = \sum_{n=i}^{\infty} (1 - p)^{n-1}p = (1 - p)^{i-1}.$$ 

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} \Pr(X \geq i)$$

$$= \sum_{i=1}^{\infty} (1 - p)^{i-1}$$

$$= \frac{1}{1 - (1 - p)}$$

$$= \frac{1}{p}$$
Let $E$ = call to Order($S, k$) was successful. 
$t = \text{number of calls till } E$.

$t$ has a Geometric distribution with $p = 1/3$, $t \sim G(\frac{1}{3})$ 
$E[t] = 1/p = 3$

For $|S| = n$, 
$E[T(S, k)] \leq \sum_{j=0}^{\log_{3}/2} n \ E[t] n (\frac{2}{3})^j = \sum_{j=0}^{\log_{3}/2} n 3n (\frac{2}{3})^j \leq 9n$.

Theorem

1. The algorithm always returns the $k$-smallest element in $S$
2. The algorithm performs $\leq 9n$ comparisons in expectation.
Finding the \( k \)-Smallest Element with no Randomization

Procedure \( \text{Det-Order}(S, k) \);

Input: An array \( S \), an integer \( k \leq |S| = n \).

Output: The \( k \) smallest element in the set \( S \).

1. If \( |S| = k = 1 \) return \( S \).
2. Let \( y \) be the first element is \( S \).
3. Compare all elements of \( S \) to \( y \). Let \( S_{\leq} = \{ x \in S \mid x \leq y \} \) and \( S_{>} = \{ x \in S \mid x > y \} \).
4. If \( k \leq |S_{\leq}| \) return \( \text{Det-Order}(S_{\leq}, k) \) else return \( \text{Det-Order}(S_{>}, k - |S_{\leq}|) \).

Theorem

The algorithm returns the \( k \)-smallest element in \( S \) and performs \( O(n) \) comparisons in expectation over all possible input permutations.
Randomized Algorithms:

- Analysis is true for any input.
- The sample space is the space of random choices made by the algorithm.
- Repeated runs are independent.

Probabilistic Analysis:

- The sample space is the space of all possible inputs.
- If the algorithm is deterministic repeated runs give the same output.
A **Monte Carlo Algorithm** is a randomized algorithm that may produce an incorrect solution.
For decision problems: A **one-side error** Monte Carlo algorithm errs only one one possible output, otherwise it is a **two-side error** algorithm.
A **Las Vegas** algorithm is a randomized algorithm that **always** produces the correct output.
In both types of algorithms the run-time is a random variable.
Finding the \textit{k}-Smallest Element

Procedure Order\((S, k)\);
\textbf{Input:} A set \(S\), an integer \(k \leq |S| = n\).
\textbf{Output:} The \(k\) smallest element in the set \(S\).

1. If \(|S| = k = 1\) return \(S\).
2. Choose a random element \(y\) uniformly from \(S\).
3. Compare all elements of \(S\) to \(y\). Let \(S_\leq = \{x \in S \mid x \leq y\}\) and \(S_\geq = \{x \in S \mid x > y\}\).
4. If \(k \leq |S_\leq|\) return Order\((S_\leq, k)\) else return Order\((S_\geq, k - |S_1|)\).

\textbf{Theorem}

1. \textit{The algorithm always returns the \(k\)-smallest element in \(S\)}
2. \textit{The algorithm performs \(\leq 9n\) comparisons in expectation.}

What is the probability that the algorithm performs much more than \(9n\) comparisons?
# Bounding Deviation from Expectation

**Theorem**

**[Markov Inequality]** For any non-negative random variable $X$, and for all $a > 0$,

$$\Pr(X \geq a) \leq \frac{E[X]}{a}.$$

**Proof.**

$$E[X] = \sum_i i \Pr(X = i) \geq a \sum_{i \geq a} \Pr(X = i) = a \Pr(X \geq a).$$

Example: The expected number of comparisons executed by the $k$-select algorithm is $\leq 9n$. The probability that it executes $18n$ comparisons or more $\leq \frac{9n}{18n} = \frac{1}{2}$. 
## Definition

The **variance** of a random variable $X$ is


## Definition

The **standard deviation** of a random variable $X$ is

$$\sigma(X) = \sqrt{Var[X]}.$$
Chebyshev’s Inequality

**Theorem**

For any random variable $X$, and any $a > 0$,

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

**Proof.**

$$\Pr(|X - E[X]| \geq a) = \Pr((X - E[X])^2 \geq a^2)$$

By Markov inequality

$$\Pr((X - E[X])^2 \geq a^2) \leq \frac{E[(X - E[X])^2]}{a^2}$$

$$= \frac{\text{Var}[X]}{a^2}$$
Theorem

For any random variable $X$ and any $a > 0$:

$$Pr(|X - E[X]| \geq a\sigma[X]) \leq \frac{1}{a^2}.$$  

Theorem

For any random variable $X$ and any $\varepsilon > 0$:

$$Pr(|X - E[X]| \geq \varepsilon E[X]) \leq \frac{Var[X]}{\varepsilon^2 (E[X])^2}.$$
Back to the $k$-select Algorithm

- Let $T(S, k)$ be the total number of comparisons.
- Let $t_i$ be the number of iterations between the $i$-th successful call (included) and the $i + 1$-th (excluded):
  - $T(S, k) \leq \sum_{i=0}^{\log_{3/2}n} n(2/3)^i t_i$.
- $T_i \sim G(1/3)$, $E[T_i] = 3$
- $E[T(S, k)] \leq \sum_{j=0}^{\log_{3/2}n} 3n \left(\frac{2}{3}\right)^j \leq 9n$.
- $Var[T(S, k)] = \sum_{i=0}^{\log_{3/2}n} 9n^2(2/3)^{2i} Var[t_i]$

What is the variance of $t_i$?
What is the variance of $T(S, k)$?
Variance of a Geometric Random Variable

- We use

\[ \text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2. \]

- To compute \( \mathbb{E}[X^2] \), let \( Y = 1 \) if the first trial is a success, \( Y = 0 \) otherwise.

\[
\mathbb{E}[X^2] = \Pr(Y = 0)\mathbb{E}[X^2 | Y = 0] + \Pr(Y = 1)\mathbb{E}[X^2 | Y = 1]
= (1 - p)\mathbb{E}[X^2 | Y = 0] + p\mathbb{E}[X^2 | Y = 1].
\]

- If \( Y = 0 \) let \( Z \) be the number of trials after the first one.

\[
\mathbb{E}[X^2] = (1 - p)\mathbb{E}[(Z + 1)^2] + p \cdot 1
= (1 - p)\mathbb{E}[Z^2] + 2(1 - p)\mathbb{E}[Z] + 1,
\]
\[
\begin{align*}
E[X^2] &= (1 - p)E[(Z + 1)^2] + p \cdot 1 \\
&= (1 - p)E[Z^2] + 2(1 - p)E[Z] + 1,
\end{align*}
\]

\[E[Z] = 1/p \text{ and } E[Z^2] = E[X^2].\]

\[
\begin{align*}
E[X^2] &= (1 - p)E[(Z + 1)^2] + p \cdot 1 \\
&= (1 - p)E[Z^2] + 2(1 - p)E[Z] + 1, \\
&= (1 - p)E[X^2] + 2(1 - p)/p + 1 \\
&= (1 - p)E[X^2] + (2 - p)/p,
\end{align*}
\]

\[E[X^2] = (2 - p)/p^2.\]

\[Var[X] = E[X^2] - E[X]^2 = \frac{2-p}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}.\]
Back to the $k$-select Algorithm

- Let $T(S,k)$ be the total number of comparisons. $|S| = n$
- Let $t_i$ be the number of iterations between the $i$-th successful call (included) and the $i + 1$-th (excluded):
  - $T(S,k) \leq \sum_{i=0}^{\log_3 2} n (2/3)^i t_i$
  - $t_i \sim G(1/3)$, therefore $E[t_i] = 3$, $Var[t_i] = 9/4$. 
  - $E[T(S,k)] \leq \sum_{j=0}^{\log_3 2} 3n (2/3)^j \leq 9n$
  - $Var[T(S,k)] = \sum_{i=0}^{\log_3 2} n^2 (2/3)^{2i} Var[t_i] \leq 5n^2$

$$Pr(|T(S,k) - E[T(S,k)]| \geq \delta E[T(S,k)]) \leq \frac{Var[T(S,k)]}{\delta^2 E[T(S,k)]^2} \leq \frac{5n^2}{\delta^2 81n^2}$$

$$Pr($$number of comparisons is $\geq 18n$$) \leq \frac{5}{81} \approx 0.07$$

With Markov inequality that probability was $\leq 1/2$. 