Solutions to selected HW problems:

Mid-sem policy:

\[
\begin{align*}
\text{Accuracy} & \quad \text{Confidence} \\
X_1 &= 1 \quad \text{if 1st sampled person votes Congress} \\
X_2 &= 1 \quad \text{if 2nd 1st} \\
\vdots & \quad \vdots \\
X_m &= 1 \quad \text{if mth sampled} \\
E[X_i] &= p, \quad E[X_2] = p, \quad \ldots \quad E[X_m] = p.
\end{align*}
\]

\[S = \sum X_i\]

\[E(S) = mp\]

Predict our estimate as \(\frac{S}{m}\)

Q1) How accurate is prediction? (2)

Q2) How confident are we? (1-8)

Q3) When are we 1-5 confident that accuracy is within \(\varepsilon\)?

When is \(Pr\left(\left|\frac{S}{m} - p\right| > \varepsilon\right) < \delta\)

\[\varepsilon \approx \delta \Rightarrow \Pr\left(\left|\frac{1}{m} - p\right| > \varepsilon\right) < \delta\]

\[\text{Chebyshev's Inequality:} \quad \text{Tail prob:} \quad \exp\left(-\frac{\varepsilon^2 m}{\text{Variance}(S) + \varepsilon^2}ight) \quad \text{Variance}(S) \leq mp \leq m\]
tail prob = \[ \exp\left( -\frac{2m^2}{m} \right) = \exp(-2m) \]

Want \[ \exp(-2m) \leq \delta \]

Confidence is easy to achieve (inside ln)

\[ m > \frac{1}{2} \ln \left( \frac{1}{\delta} \right) \]

\[ \delta = 0.01, \quad \varepsilon = 0.001 \]

\[ m > 10 \ln \frac{1}{\varepsilon} \]

Ex 2 (HW 10) ← Designing a tournament

There are \( m \) teams. (Unknown implicit rankings)

Team \( i \) vs Team \( j \): better team wins \( wp \frac{1}{2} + \delta \) (\( \approx 0.6 \))

Best of \( k \) tournament:

\[ \text{best of } k = \frac{1}{2} \left( 1 + \log_2 k \right) = 0.6 \]

Best of \( k \) tournament = quite small (\( \log_2 \))

Best team does not win first series

\[ = \text{When } k / 2 \text{ matches}
\]

\[ E[\text{matches best team wins in } j] = k \left( \frac{1}{2} + \delta \right) \]

\[ \implies E[\text{matches best team loses}] = k \left( \frac{1}{2} - \delta \right) \]

Chernoff:

\[ P_k(\text{bad event}) = \exp\left( -\frac{k^2 \delta^2}{k} \right) \]

Set \[ k = \frac{1}{\delta^2} \log \left( \frac{\log n}{\varepsilon} \right) \]

\[ P_k(\text{best team loses } 1^{st} \text{ series}) \leq \frac{\varepsilon}{\log n} \]

Can union bound over \( m \) to overall \( P \leq \varepsilon \).
Can \( \text{UNION bound} \) and different serves, to overall \( pr \leq \varepsilon \).

Overall \# matches \( = O\left( \frac{n \log \log n}{\varepsilon^2} \right) \)

\[
\text{better tournament}
\]
\[
\text{to have non-equal \# matches in each level: } \lambda \leq \varepsilon
\]
\[
\text{Overall \# failures: } \leq \left( \exp\left( -k_1 \varepsilon^2 \right) + \exp\left( -k_2 \varepsilon^2 \right) + \cdots + \exp\left( -k_m \varepsilon^2 \right) \right)
\]
\[
\text{Overall \# games: } = (k_1 \frac{n}{2} + k_2 \frac{n}{4} + k_3 \frac{n}{8} + \cdots + k_m \frac{n}{2^{m-1}})
\]
\[
k_1 \leq k_2 < k_3 \cdots < k_m \Rightarrow \Theta(n)
\]

JL exercise:

Given \( n \) vectors \( v_1, v_2, \ldots, v_n \in \mathbb{R}^D \) \( D \text{ is very high} \)

Can I sample a small \# co-ordinates to preserve the lengths of resulting vectors?

\( \frac{\log n}{\varepsilon^2} \) dimensions, say.

A: No!! (\( \square \))

\[
v_1 = (1, 0, 0, \ldots, 0)
\]
\[
v_2 = (0, 1, 0, 0, \ldots, 0)
\]
\[
v_3 = (0, 0, 1, 0, \ldots, 0)
\]
\[
v_n = (0, 0, 0, \ldots, 1)
\]

Even if I sample \( d = \frac{\log n}{\varepsilon^2} \) co-ordinates, almost all of these vectors collapse to 0.

Main issue: variance/co-ordinate values vary a lot!
Main Issue: Variances/co-ordinate values vary a lot!

Fix:
- Make sure all co-ordinates are bounded in a nice way.

Rotate each vector "randomly"

\[ v_i = \text{random rotation of every vector} \]

What was originally \((1, 0, 0, 0, \ldots)\)

now looks like \((u_1, u_2, \ldots, u_n)\)

where all \(u_i\) are \([-\frac{6}{\sqrt{n}}, \frac{6}{\sqrt{n}}]\)

- SAMPLING WORKS NOW!

We have like a uniformly random rotation of co-ordinate space.

Flows, Hamilton Path, & the Ambulance problem.

**OR directed**

if all \(e^L \leq T > 0\),

\(u^L \leq T > 0\)

Want a min-cost flow satisfying these bounds

FORMULATED AS LINEAR PROGRAM.

Specifically for flow problem, for \(OPT \geq 0\), the flow values on edges one

\[ \text{when } e^L \leq u^L \leq T > 0 \]

Can Hamilton Path be solved using the integrality of flow?

Is there a path from \(s\) to \(t\) which+
Is there a path from $s \to t$ which visits every other vertex exactly once?

undirected $\to$ directed is easy

there are now upper & lower capacities on edges also!

water capacities are easy to handle.

$V \to$ 

$l = 1$

$V_{in} \to$ $V_{out}$

$\Rightarrow Np = P$

does this solve Hamilton Path? 

flow can output some of this kind

AMBULANCE QUESTION

\[
\begin{align*}
\begin{array}{cccc}
\xi_1 & \xi_2 & \xi_3 & \xi_T \\
\end{array}
\end{align*}
\]

\[\Rightarrow \text{cast it as a flow problem.}\]
Find a min-cost flow from $s$ to $t$ satisfying all lower & upper capacity constraints, that will give optimal migration policy.