LEcTure 11

Last time
- Chernoff Bounds

Today
- Hoeffding Bounds
- Applications of Chernoff-Hoeffding Bounds
- Estimating a Parameter
- Set Balancing
Sums of independent RVs

Chernoff Bound (Upper Tail). Let $X_1, \ldots, X_n$ be independent Bernoulli RVs. Let $X = X_1 + \cdots + X_n$ and $\mu = \mathbb{E}[X]$. Then

- **(stronger)** for any $\delta > 0$,\[
\Pr[X \geq (1 + \delta)\mu] \leq \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}}\right)^\mu.
\]

- **(easier to use)** for any $\delta \in (0,1]$,\[
\Pr[X \geq (1 + \delta)\mu] \leq e^{-\mu \delta^2/3}.
\]
Sums of independent RVs

Chernoff Bound (Lower Tail). Let $X_1, \ldots, X_n$ be independent Bernoulli RVs. Let $X = X_1 + \cdots + X_n$ and $\mu = \mathbb{E}[X]$. Then

- (stronger) for any $\delta \in (0,1)$,

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

- (easier to use) for any $\delta \in (0,1)$,

$$\Pr[X \leq (1 - \delta)\mu] \leq e^{-\mu\delta^2/2}.$$
Sums of independent RVs

Chernoff Bound (Both Tails). Let $X_1, \ldots, X_n$ be independent Bernoulli RVs. Let $X = X_1 + \cdots + X_n$ and $\mu = \mathbb{E}[X]$. Then

- for any $\delta \in (0,1)$,

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

We throw $n$ balls uniformly and independently into $n$ bins. Let $Y_1$ be the number of balls that fell into bin 1. Determine $m$ such that $\Pr[Y_1 > m] \leq \frac{1}{n^2}$. 

10/7/2021 

Sofya Raskhodnikova; Randomness in Computing
Sums of independent RVs

Hoeffding Bound. Let $X_1, \ldots, X_n$ be independent RVs with $\mathbb{E}[X_i] = \mu_0$ and $\Pr[a \leq X_i \leq b] = 1$. Let $X = X_1 + \cdots + X_n$. Then

- (upper tail) $\Pr[X \geq \mu_0 n + \epsilon n] \leq e^{-2n\epsilon^2/(b-a)^2}$
- (lower tail) $\Pr[X \leq \mu_0 n - \epsilon n] \leq e^{-2n\epsilon^2/(b-a)^2}$
Application: Estimating a parameter

- **Unknown**: probability $p$ that a feature occurs in the population.
- Obtain an estimate by taking $n$ samples
- $X \sim \text{Bin}(n, p)$
- Let $\hat{p} = X/n$.
- A $1 - \gamma$ confidence interval for parameter $p$ is an interval $[\hat{p} - \delta, \hat{p} + \delta]$ such that $\Pr[p \in [\hat{p} - \delta, \hat{p} + \delta]] \geq 1 - \gamma$.
- Find a tradeoff between $\gamma$, $\delta$, and $n$. 
Application: Estimating a parameter

- A $1 - \gamma$ confidence interval for parameter $p$ is an interval $[\hat{p} - \delta, \hat{p} + \delta]$ such that $\Pr[p \in [\hat{p} - \delta, \hat{p} + \delta]] \geq 1 - \gamma$.
- Find a tradeoff between $\gamma$, $\delta$ and $n$.

**Solution:** $\mathbb{E}[X] = np$

- Suppose $p \notin [\hat{p} - \delta, \hat{p} + \delta]$
- Case 1: $p < \hat{p} - \delta$. Then $\hat{p} > p + \delta$
- Case 2: $p > \hat{p} + \delta$. Then $\hat{p} < p - \delta$

- $\gamma = 2 \cdot e^{-2\delta^2 n}$
Application: Set Balancing

- **Given:** an $n \times m$ matrix $A$ with 0-1 entries
- **Definition:** $\|(x_1, \ldots, x_n)\|_\infty = \max_{i \in [n]} |x_i|$
- **Find:** $b \in \{-1,1\}^m$ minimizing $\|Ab\|_\infty$

Partition subjects into two groups, so that each feature is balanced.
Algorithm: Choose each $b_i$ independently from $\{-1,1\}$.

Theorem. $\Pr\left[ \|Ab\|_\infty \geq \sqrt{4m \ln n} \right] \leq 2/n$.

Proof: Let $\bar{a}_i$ be the vector in row $i$ of $A$.

- Let $k$ be the number of $1$s in $\bar{a}_i$.
- For $i \in [n]$, let $Z_i$ be the random variable $\sum_{j \in [m]} a_{ij} b_j$.
- $\mathbb{E}[Z_i] = k$ nonzero independent RVs each $+1$ or $-1$ w.p. $1/2$
- By Hoeffding, $\Pr[|Z_i| \geq \epsilon k] \leq 2e^{-\frac{2\epsilon^2 k}{4}} = 2e^{-\frac{2m \ln n}{2m}} = 2n^{-2}$

By Union Bound,

$\Pr\left[ |Z_i| \geq \epsilon k \text{ for some } i \in [n] \right] \leq \frac{2m}{k} \leq m$
An $n$-dimensional hypercube is a directed graph with

- $N = 2^n$ nodes, each indexed by $n$-bit integer
- containing the directed edge $(x, y)$ iff $x$ and $y$ differ in exactly one bit

How many edges?

- **Routing.** Each node is a routing switch.
  Edges = communication channels
  An edge can carry one packet in one step.

- A *routing algorithm* specifies a path from $s$ to $t$ for each pair of nodes $(s, t)$ and a *queuing policy* for ordering packets that are waiting for the same link (e.g., FIFO – First In First Out or FTG – Furthest to Go)
Permutation Routing Problem

- Each node is the source of one packet
- Each node is the designation of one packet
  - E.g., on a complete graph, it can be solved in one time step.
  - On sparse graphs?

Hypercube: $N$ nodes $N\log N$ edges.

## Bit-Fixing Routing Algorithm for the Hypercube

1. Let $x$ be the current node and $y$ be the destination of a packet
2. Find smallest $i \in [n]$ such that $x_i \neq y_i$
3. Traverse the edge $(x, x_1 \ldots x_{i-1} \bar{x_i} x_{i+1} \ldots x_n)$
Bad example for congestion

Transpose permutation \((n \text{ is even})\)

From each \(x\), send a packet to \((x_{n/2+1}, \ldots, x_n, x_1, \ldots, x_{n/2})\)