cs473 Algorithms Out: *Fri., 2022-03-25 17:00*

Problem Set #7

Prof. Michael A. Forbes
Dr. Bhaskar Chaudhury Due: *Fri., 2022-04-01 17:00*

All problems are of equal value.

1. Load Balancing. Kleinberg-Tardos Chapter 13, Problem #11.

2. (Multiplicative Chernoff Bound). Let $\mathsf{X}_1, \ldots, \mathsf{X}_n$ be independent random variables taking values over the continuous interval $[0,1]$. Let $\mathsf{X} = \sum_i \mathsf{X}_i$. Show the following.

   (a) For $r \in (-\infty, \ln 2]$, prove that $\mathbb{E}[e^{r\mathsf{X}}] \leq e^{r\mathbb{E}[\mathsf{X}]+r^2\mathbb{E}[\mathsf{X}]}$, where you may use (without proof) that $1+x \leq e^x$ for all $x \in \mathbb{R}$, and $e^x \leq 1+x+x^2$ for $x \leq \ln 2$.

   (b) Explain how the above used the independence of the X_i.

   (c) Apply Markov's inequality ($\Pr[\mathsf{Y} \geq a] \leq \mathbb{E}[\mathsf{Y}]/a$) to $e^{r\mathsf{X}}$, and optimize over r, to conclude that:

      i. For $0 \leq \epsilon \leq \ln 4$, $\Pr[\mathsf{X} \geq (1+\epsilon)\mathbb{E}[\mathsf{X}]] \leq e^{-\epsilon^2\mathbb{E}[\mathsf{X}]/4}$.
      ii. For $\epsilon \geq \ln 4$, $\Pr[\mathsf{X} \geq (1+\epsilon)\mathbb{E}[\mathsf{X}]] \leq 2^{-\epsilon\mathbb{E}[\mathsf{X}]/2}$.
      iii. For $0 \leq \epsilon \leq 1$, $\Pr[\mathsf{X} \leq (1-\epsilon)\mathbb{E}[\mathsf{X}]] \leq e^{-\epsilon^2\mathbb{E}[\mathsf{X}]/4}$.
      iv. (Additive Chernoff Bound) For $\epsilon \geq 0$, $\Pr[|\mathsf{X} - \mathbb{E}[\mathsf{X}]| \geq \epsilon \cdot n] \leq 2e^{-\epsilon^2 n/4}$.

   Note: The additive Chernoff bound suffices for applications such as estimating the errors in polling, but the multiplicative bound is in general stronger and often needed (e.g. consider $\mathbb{E}[\mathsf{X}] = \lg n$ and the resulting bound for $\Pr[\mathsf{X} \geq 2\mathbb{E}[\mathsf{X}]]$). Note also that the above omits one range of parameters, where one can show that $\Pr[\mathsf{X} \geq (1+\epsilon)\mathbb{E}[\mathsf{X}]] \leq e^{-(1+\epsilon)\ln(1+\epsilon)\mathbb{E}[\mathsf{X}]/4}$ if $\epsilon \geq 1$.

3. Let G be an undirected graph $G = (V, E)$ with n vertices and m edges. We wish to partition the vertex set V into k disjoint sets $V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_k$, while minimizing the total number of *conflicting vertices*, where two vertices are in conflict if they are adjacent in G and belong to the same set in the partition.

   (a) Give an $O(n)$ time *randomized* algorithm that outputs a partition, where the expected number of conflicts is at most $\frac{m}{k}$.

   (b) Give an $O(n+m)$ time *deterministic* algorithm that outputs a partition, where the number of conflicts is at most $\frac{m}{k}$.
   Hint: Use a greedy approach.