cs473: Algorithms	Problem Set #4	Assigned: *Thu., Sep. 24, 2020* Updated: *Tue., Sep. 29, 2020*
Prof. Michael A. Forbes		Due: *Thu., Oct. 1, 2020 (5:00pm)*

All problems are of equal value.

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

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

   (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 $\mathrm{e}^{r\mathsf{X}}$, and optimize over r, to conclude that:

      i. For $0 \leq \epsilon \leq \ln 4$, $\Pr[\mathsf{X} \geq (1 + \epsilon)\mathbb{E}[\mathsf{X}]] \leq \mathrm{e}^{-\epsilon^2\mathbb{E}[\mathsf{X}]/4}$.

      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 $0 \leq \epsilon \leq 1$, $\Pr[\mathsf{X} \leq (1 - \epsilon)\mathbb{E}[\mathsf{X}]] \leq \mathrm{e}^{-\epsilon^2\mathbb{E}[\mathsf{X}]/4}$.

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

   Note: The additive Chernoff-Hoeffding 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 $\Pr[\mathsf{X} \geq (1 + \epsilon)\mathbb{E}[\mathsf{X}]] \leq \mathrm{e}^{-(1+\epsilon)\ln(1+\epsilon)\mathbb{E}[\mathsf{X}]/4}$ if $\epsilon \geq 1$.

2. For $0 < p \leq 1$, let $\mathsf{Geom}(p)$ denote the *geometric distribution with parameter* p, so that if $\mathsf{X} \sim \mathsf{Geom}(p)$ then for $i \geq 1$ we have $\Pr[\mathsf{X} = i] = p(1 - p)^{i-1}$. Let $\mathsf{X}_1, \ldots, \mathsf{X}_n$ be independent random variables distributed according to $\mathsf{Geom}(p)$, and let $\mathsf{X} = \sum_{i=1}^n \mathsf{X}_i$.

   (a) Compute $\mathbb{E}[\mathsf{X}]$ and $\mathrm{Var}(\mathsf{X})$.

   (b) For $c \geq 2$, use Chebyschev's inequality to give a bound for $\Pr[\mathsf{X} \geq c\mathbb{E}[\mathsf{X}]]$.

   (c) Let $\mathsf{Bern}(p)$ denote the *Bernoulli distribution with parameter* p, so that if $\mathsf{Y} \sim \mathsf{Bern}(p)$ then $\Pr[\mathsf{Y} = 1] = p = 1 - \Pr[\mathsf{Y} = 0]$. Let $\mathsf{Y}_1, \ldots, \mathsf{Y}_t$ be independent random variables distributed according to $\mathsf{Bern}(p)$, and let $\mathsf{Y} = \sum_{j=1}^t \mathsf{Y}_j$. Show that $\Pr[\mathsf{X} \geq t] \leq \Pr[\mathsf{Y} \leq n]$.
   Hint: Consider the following experiment. You are given a biased coin that yields heads with probability p, and tails with probability $1 - p$. Flip the coin until you both (a) have seen n heads, and (b) have flipped at least t times in total. Using only this experiment as a probability space, construct random variables identically distributed to the X_i and Y_j. First try proving that $\Pr[\mathsf{X} > t] = \Pr[\mathsf{Y} < n]$.

   (d) For $c \geq 2$, use Problem 1 to give a bound for $\Pr[\mathsf{X} \geq c\mathbb{E}[\mathsf{X}]]$.

3. Consider the following variant of quicksort. Given an array A of n distinct numbers the algorithm picks a pivot x uniformly at random from A and computes the rank of x. If the rank of x is between $n/4$ and $3n/4$ (a *balanced* pivot), the algorithm proceeds as with standard quicksort. However, if the pivot is not balanced, the algorithm selects a new random pivot.

(a) Write a formal description of the algorithm.

(b) Directly prove that the expected runtime of this algorithm is $O(n \log n)$.

(c) For $c \geq 2$, bound the probability that the algorithm takes more than $c \cdot \Theta(n \log n)$ steps.
Hint: Use Problem 2.

(d) For a random variable Z over non-negative integers $\mathbb{N}$, prove that $\mathbb{E}[\mathsf{Z}] = \sum_{i \geq 1} \Pr[\mathsf{Z} \geq i]$.
Hint: Use indicator random variables, and linearity of expectation over an *infinite* sum. (There are non-trivial issues of convergence here; ignore them.)

(e) Use the previous two parts to reprove that the expected runtime is $O(n \log n)$.