Section 4

1. **(Chernoff) (Book 4.5)** We plan to conduct an opinion poll to find out the percentage of people in a community who want its president impeached. Assume that every person answers yes or no. If the actual fraction of people who want the president impeached is p, we want to find an estimate X of p such that

$$\Pr(|X - p| \leq \varepsilon p) > 1 - \delta \tag{1}$$

   for a given ε and δ, with $0 < \varepsilon,\ \delta < 1$.

   We query N people chosen independently and uniformly at random from the community and output the fraction of them who want the president impeached. How large should N be for our result to be a suitable estimator of p? Use Chernoff bounds, and express N in terms of p, ε, and δ. Calculate the value of N from your bound if $\varepsilon = 0.1$ and $\delta = 0.05$ and if you know that p is between 0.2 and 0.8.

2. **(Tail Inequalities) (Book Exercise 4.2)** We have a standard six-sided die. Let X be the number of times that a 6 occurs over n throws of the die. Let p be the probability of the event $X \geq n/4$. Compare the best upper bounds on p that you can obtain using Markov's inequality, Chebyshev's inequality, and Chernoff bounds.

3. **(Permutations and Chernoff) (Book Exercise 4.8)** We show how to construct a random permutation π on $[1, n]$, given a black box that outputs numbers independently and uniformly at random from $[1, k]$ where $k \geq n$. If we compute a function $f[1, n] \to [1, k]$ with $f(i) \neq f(j)$ for $i \neq j$, this yields a permutation; simply output the numbers $[1, n]$ according to the order of the $f(i)$ values. To construct such a function f, do the following for $j = 1, ..., n$: choose $f(j)$ by repeatedly obtaining numbers from the black box and setting $f(j)$ to the first number found such that $f(j) \neq f(i)$ for $i < j$. Prove that this approach gives a permutation chosen uniformly at random from all permutations. Find the expected number of calls to the black box that are needed when $k = n$ and $k = 2n$. For the case $k = 2n$, argue that the probability that each call to the black box assigns a value of $f(j)$ to some j is at least $1/2$. Based on this, use a Chernoff bound to bound the probability that the number of calls to the black box is at least $4n$.

4. **(Chernoff) (Book Exercise 4.12)** Consider a collection $X_1, ..., X_n$ of n independent geometrically distributed random variables with mean 2. Let $X = \sum_{i=1}^{n} X_i$ and $\delta > 0$.

   (a) Derive a bound on $\Pr(X \geq (1 + \delta)(2n))$ by applying a Chernoff bound to a sequence of $(1 + \delta)(2n)$ fair coin tosses.

   (b) Derive a bound on $\Pr(X \geq (1 + \delta)(2n))$ using the moment generating function for geometric random variables.

   (c) Which bound is better?