CS5314 Randomized Algorithms

Homework 1
Due: 3:20 pm, April 10, 2007 (before class)

1. During the class, we have learnt that, for any two events E_1 and E_2,

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2).$$

In fact, there is a general equality for any finite number of events. This equality is called *principle of inclusion-exclusion*, which is expressed as follows.

Let $E_1, E_2, \ldots, E_n$ be any n events. Then

$$\begin{aligned} Pr\left(\bigcup_{i=1}^{n} E_i\right) &= \sum_{i=1}^{n} \Pr(E_i) - \sum_{i<j} \Pr(E_i \cap E_j) + \sum_{i<j<k} \Pr(E_i \cap E_j \cap E_k) \\ &\quad - \cdots + (-1)^{\ell+1} \sum_{i_1<i_2<\cdots<i_\ell} \Pr\left(\bigcap_{r=1}^{\ell} E_{i_r}\right) \\ &\quad + \cdots + (-1)^n \Pr\left(\bigcap_{i=1}^{n} E_i\right). \end{aligned}$$

   (a) Prove that the principle of inclusion-exclusion is correct.

   (b) Suppose I choose a number, uniformly at random, from the range [1,100000]. Using the principle of inclusion-exclusion, determine the probability that the number chosen is divisible by one ore more of 4, 6, and 9.

2. We have a function $F : \{0, \ldots, n-1\} \to \{0, \ldots, m-1\}$. We know that, for any x and y with $0 \leq x, y \leq n-1$, $F((x+y) \bmod n) = (F(x) + F(y)) \bmod m$.

Currently, the values of F are stored in a lookup table so that we can evaluate any $F(x)$ easily and correctly. However, our naughty classmate, Jimmy, has told us that he has changed exactly $1/5$ of the lookup table entries (assume n is a multiple of 5) while we were having lunch, and you know that Jimmy is not going to tell us which entries are now wrong.

   (a) Describe a simple randomized algorithm that, given any input z, outputs the correct value $F(z)$ with probability at least $1/2$. Note that your algorithm needs to work for every value of z, regardless which entries Jimmy has changed.

   (b) Suppose you can repeat your initial algorithm three times. Can you improve the probability of returning the correct value?

3. During the class, we have studied a simple randomized algorithm so that given any graph, we can find its min-cut with probability at least $2/(n(n-1))$. Now, we define an r-cut of a graph G to be a set of edges in G whose removal will break G into r or more connected components. (That is, the normal definition of a cut is equivalent to a 2-cut here.)

Describe a randomized algorithm for finding an r-cut with minimum number of edges. Also, analyze the probability that the algorithm succeeds in one iteration.

4. A permutation on the numbers $[1, n]$ can be represented by a function $\pi : [1, n] \to [1, n]$, where $\pi(i)$ is the position of i in the ordering given by the permutation. A fixed point of the permutation π is a value such that $\pi(x) = x$. Suppose that the permutation π is chosen, uniformly at random, from the $n!$ possible permutations. Find the expected number of fixed points in π.

5. You need a new staff assistant, and you have n people to interview. You want to hire the best candidate for this position. When you interview the candidates, you can give each of them a score, so that the one with the highest score will be the best, and there will be no ties.

   You interview the candidates one by one. Because of your company's hiring policy, after you interview the kth candidate, you either offer the candidate the job immediately, or you will forever lose the chance to hire that candidate.

   We suppose that the candidates are interviewed in a random order, chosen uniformly at random from all $n!$ possible orderings.

   Consider the following strategy: First interview m candidates but reject them all. Then, from the $(m+1)$th candidate, hire the first candidate who is better than *all* of the previous candidates you have interviewed.[1]

   (a) Let E be the event that we hire the best candidate, and let E_i be the event that the ith candidate is the best and we hire him. Determine $\Pr(E_i)$, and show that

   $$\Pr(E) = \frac{m}{n} \sum_{j=m+1}^{n} \frac{1}{j-1}.$$

   (b) Bound $\sum_{j=m+1}^{n} \frac{1}{j-1}$ to obtain

   $$\frac{m}{n}(\log_e n - \log_e m) \leq \Pr(E) \leq \frac{m}{n}(\log_e(n-1) - \log_e(m-1)).$$

   (c) Show that $m(\log_e n - \log_e m)/n$ is maximized when $m = n/e$. Explain why this means $\Pr(E) \geq 1/e$ for this choice of m.

6. (Further studies: No marks) Prove that $E[X^k] \geq E[X]^k$ for any positive even integer k.

7. (Further studies: No marks) Prove that if $E_1, E_2, \ldots, E_n$ are mutually independent, then the events $\overline{E_1}, \overline{E_2}, \ldots, \overline{E_n}$ are also mutually independent.

8. (Further studies: No marks) If X is a binomial random variable $Bin(n, 1/2)$ with $n \geq 1$, show that the probability that X is even is $1/2$.

[1] That is, you will hire the kth candidate if $k > m$ and this candidate is better than all of the $k-1$ candidates you have already interviewed.