Assignment II

Topics in Algorithms

- 1. Suppose you keep throwing a fair dice till two consecutive 6 is obtained, what is the expected number of throws required? (It is not 36).
- 2. Let X be a random variable taking values $\{1, 2, 3, ...\}$ with finite expectation. Show that $E(X) = \sum_{i=1}^{\infty} Pr(X \ge i)$
- 3. Given an Array A[1..n], consider the following algorithm:

```
select = A[1];
for i = 2 to n {
    set select = A[i] with probability p(i)
}
```

If $p(i) = \frac{1}{i}$ for $i \in \{1, 2, ..., n\}$, find Pr(select = i). This algorithm is useful to achieve uniform sampling from a large (on-line) data stream without having to store the entire stream. What if $p(i) = \frac{1}{2}$?

- 4. Here is another on-line streaming algorithm. n candidates are appearing before you for a job interview from which you have to select exactly one. Your company requires that at the end of the interview for each candidate, you have to tell her whether she is selected or not. Hence, after each interview, there is decision whether to accept the candidate or call the next. Here is a strategy: reject the first m candidates and then accept the first candidate after m who is better than the first m candidates. (Assume that no two candidates are "equal" and there is a proper ordering of the merit of the candidates).
 - 1. What is the probability that the $(m+1)^{th}$ candidate is the best and you select the candidate?
 - 2. show that the probability for you to select the best candidate is $\frac{m}{n}(\frac{1}{m+1} + \frac{1}{m+2} + ... + \frac{1}{n})$.
 - 3. Show that the probability of selecting the best is maximized when $m \approx \frac{n}{e}$, and that for this value of m, the probability of selecting the best candidate is nearly $\frac{1}{e}$.
- 5. This question presents another algorithm for finding a large independent set in a graph. First select each vertex with probability p and create a set S. Now, look at the subgraph induced by S. If this subgraph contains say r edges, remove each edge along with one of the end points to output the resultant set of at least |S| r vertices.
 - 1. Let Y be the random variable denoting the value of r. Show that $E(Y) = p^2 m$, where m is the number of edges.
 - 2. Show that the algorithm indeed returns an independent set.
 - 3. Find the expected size of the independent set returned by the algorithm.
 - 4. Find the value of p that maximizes the expected size of the independent set returned by the algorithm. Show that for this choice of p, the expected size of independent set returned is $\frac{n^2}{4m}$.
- 6. Let $a = (a_1, a_2, ..., a_n)$ be a permutation of $\{1, 2, ..., n\}$. A position *i* is said to be a fix point of the permutation if $a_i = i$. Suppose *a* is a random permutation of $\{1, 2, ..., n\}$. Find the expected number of fix points in *a*.
- 7. Suppose $cn \log n + cn$ balls are thrown into n bins uniformly at random:
 - 1. Show that the probability that the first bin is empty is at most $\frac{1}{ne^c}$. (Use the approximation $(1-\frac{1}{n})^n \leq \frac{1}{e}$).
 - 2. Show that the probability that there exists at least one empty bin is at most $\frac{1}{e^c}$.
 - 3. Given $\epsilon > 0$, what must be the choice of c for which the probability for an empty bin to exist is less than ϵ ?
 - 4. What is the probability bound given by Chebyshev inequality? (You will see that it is weaker than the above derived union bound).
- 8. Let p be a prime number. Let a, b be numbers chosen uniformly at random from $\{0, 1, 2, .., p-1\}$. Define the sequence of numbers $(z_0, z_1, .., z_{p-1})$ as $z_i = a + ib \mod p$. Show that $z_0, .., z_{p-1}$ are pairwise independent. Where do you use the fact that p must be prime? (This question has been corrected subsequent to the initial upload).
- 9. Recall that the random vertex colouring approximation algorithm for MAXCUT returns a cut of expected size $\frac{m}{2}$, where *m* is the number of edges. We now find out the probability for the algorithm to return a cut of size at least $\frac{m}{2}$. Let *X* be a random variable denoting the size of the cut returned by the algorithm and let *E* be the event that the algorithm returns a cut of size at least $\frac{m}{2}$. Let Pr(E) = p.
 - 1. Convince yourself about the obvious facts: $E(X|E) \leq m, E(X|\overline{E}) \leq \frac{m-1}{2}$.
 - 2. Now Write E(X) in terms of the above and solve for p to show that $p \ge \frac{1}{(m+1)}$.
 - 3. Show that repeating the algorithm (m + 1) times, a cut of size at least $\frac{m}{2}$ is obtained with probability at least $1 \frac{1}{e}$