CSE 632: Analysis of Algorithms II: Randomized Algorithms Fall 2019

Homework 3

Lecturer: Shi Li **Deadline: Nov 4, 2019**

Problem 1 (30 points). Consider the procedure of throwing m balls into n bins, where each ball is thrown into 1 of the n bins, uniformly at random independently of the other balls. For any integer f, let E_f be the event that some bin contains at least f balls. We are interested in the following question: how big should f be so that $\Pr[E_f] \leq 0.1$? Prove the following statements for three different values of m:

- If $m = n$, then for some $f = O\left(\frac{\log n}{\log \log n}\right)$, we have $\Pr[E_f] \leq 0.1$.
- If $m = n \log_2 n$ (assuming $\log_2 n$ is an integer), then for some $f = O(\log n)$, we have $\Pr[E_f] \leq 0.1$.
- If $m = n^2$, then for some $f = n + O(\sqrt{n} \log_2 n)$, we have $\Pr[E_f] \leq 0.1$.

Problem 2 (20 points). Suppose we have an array A of n distinct integers. We say an integer in A is an approximate median, if its rank in A is between $0.4n$ and $0.6n$ (the rank of x is the number of integers in A that are smaller than or equal to x). Consider the following algorithm for finding the approximate median of A:

```
1: for i ← 1 to m do
2:     let B[i] be a number chosen uniformly at random from the array A
3: output the median of array B
```

Show that for the algorithm to succeed with probability at least $1 - 1/n^2$, it suffices to choose $m = O(\log n)$. (Thus, the running time to find an approximate median is much smaller than the running time to find the (exact) median.)

Problem 3 (10 points). Consider the following game.

```
1: put a white balls and b black balls into a bin
2: repeat n times:
3:     randomly pick a ball from the bin
4:     put the picked ball back and additionally add another ball of the same color to the bin
```

So at the end of the game, we have in total $a + b + n$ balls.

For every $i = 1, \cdots, n$, let Y_i be the fraction of black balls in the bin after the i-th iteration of the loop. Show that $Y_1, Y_2, \cdots, Y_i$ is a martingale sequence.