CS 174: Combinatorics and Discrete Probability Fall 2012

Homework 5

Due: Thursday, October 4, 2012 by **9:30am**

Instructions: *You should upload your homework solutions on bspace. You are strongly encouraged to type out your solutions using LaTeX. You may also want to consider using mathematical mode typing in some office suite if you are not familiar with LaTeX. If you must handwrite your homeworks, please write clearly and legibly. We will not grade homeworks that are unreadable. You are encouraged to work in groups of 2-4, but you* **must** *write solutions on your own. Please review the homework policy carefully on the class homepage.*

Note: You *must* justify all your answers. In particular, you will get no credit if you simply write the final answer without any explanation.

Problem 1. *(Exercise 5.9 from MU – 5 points)* Consider the probability that every bin receives exactly one ball when n balls are thrown randomly into n bins.

(a) Give an upper bound on this probability using the Poisson approximation.

(b) Determine the *exact* probability of this event.

Problem 2. *(Exercise 5.13 from MU – 5 points)* Let Z be a Poisson random variable with mean μ, where $\mu \geq 1$ is an integer. First, show that $\Pr[Z = \mu + h] \geq \Pr[Z = \mu - h - 1]$ for $0 \leq h \leq \mu - 1$, and use this to conclude that $\Pr[Z \geq \mu] \geq 1/2$.

Problem 3 *(Exercise 5.14 from MU – 5 points)* Let $Y_1, \ldots, Y_n$ be Poisson random variables with mean $\mu(= m/n)$. Let $X_1, X_2, \ldots, X_n$ be the random variables denoting the number of balls in each bin when m balls are thrown in n bins. In class, we showed that for any non-negative function, f,

$$\mathbb{E}[f(Y_1, \ldots, Y_n)] \geq \mathbb{E}[f(X_1, \ldots, X_n)] \Pr[\sum_{i=1}^{n} Y_i = m]$$

When f is monotonically increasing, show that

$$\mathbb{E}[f(Y_1, \ldots, Y_n)] \geq \mathbb{E}[f(X_1, \ldots, X_n)] \Pr[\sum_{i=1}^{n} Y_i \geq m]$$

Use this and problem 2 to conclude that $\mathbb{E}[f(X_1, \ldots, X_n)] \leq 2\mathbb{E}[f(Y_1, \ldots, Y_n)]$ (see Theorem 5.10).

Problem 4 *(5 points)* Let $X_1, \ldots, X_n$ be geometric 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 squence of $(1+\delta)(2n)$ independent coin tosses.

(b) Consider the quantity $\mathbb{E}[e^{tX}]$ and derive a Chernoff bound for $\Pr[X \geq (1+\delta)(2n)]$ using Markov's inequality for the random variable e^{tX}.

(c) Which bound is better?

Problem 5. *(Exercise 4.25 from MU - 10 points)* In this exercise, we design a randomized algorithm for the following packet routing problem. We are given a network that is an undirected connected graph, G, where nodes represent processors and the edges between the nodes represent wires. We are also given a set of N packets to route. For each packet we are given a source node, a destination node, and the exact route (path in the graph) that the packet should take from the source to the destination. (We may assume that there are no loops in the path.) In each time step, at most one packet can traverse an edge. A packet can wait at any node during any time step, and we assume unbounded queue sizes at each node.

A *schedule* for a set of packets specifies the timing for the movement of packets along their respective routes. That is, it specifies which packet should move and which should wait at each time step. Our goal is to produce a schedule for the packets that tries to minimize the total time and the maximum queue size needed to route all the packets to their destinations.

(a) The dilation, d, is the maximum distance travelled by any packet. The congestion, c, is the maximum number of packets that must traverse a single edge during the entire course of the routing. Argue that the time required for any schedule should be at least $\Omega(c+d)$. (Hint: Show that the time should be at least $\max\{c,d\}$ which is $\Omega(c+d)$.)

(b) Consider the following *unconstrained* schedule, where many packets may traverse an edge during a single time step. Assign each packet an integral delay chosen randomly, independently, and uniformly from the interval $[1, \lceil \alpha c/\log(Nd) \rceil]$, where α is a *sufficiently large* constant. A packet that is assigned a delay of x waits in its source node for x time steps; then it moves on to its final destination through its specified route without ever stopping. Give an upper bound on the probability that more than $O(\log(Nd))$ packets use a particular edge e at a particular time step t.

(c) Again using the unconstrained schedule of part (b), show that the probability that more than $O(\log(Nd))$ packets pass through any edge at any time step is at most $1/(Nd)$ for a sufficiently large α.

(d) Use the unconstrained schedule to devise a simple randomized algorithm that, with high probability, produces a schedule of length $O(c + d\log(Nd))$ using queues of size $O(\log(Nd))$ and following the constraint that at most one packet crosses an edge per time step. (By high probability, we mean $1 - O(1/N)$.)