CMSC 858S, Spring 2007: HW 4, due on April 2nd (Mon.) at the start of class

Notes: Please work on this with your group-mate(s); just submit *one* writeup per group. Consulting other sources (including the Web) is not allowed. Write your solutions neatly; if you are able to make partial progress by making some additional assumptions, then state these assumptions clearly and submit your partial solution.

Technical Note: For a random variable X and constant $t > 0$, $\mathbf{E}[e^{tX}]$ is called an *exponential moment-generating function* for X, since this exponential function generates all the higher moments of X (when expanded as a Taylor series). Recall that the Chernoff-Hoeffding bounds use such functions crucially.

1. We start with a graph-theoretic result that will be useful. Recall that given a graph G with maximum degree Δ, G can be properly colored using $(\Delta + 1)$ colors. In fact, a much stronger result is known. Given a proper coloring of G, define the *color class* of color c to be the set of vertices that are colored c in this coloring. It is known that G can be colored using $(\Delta + 1)$ colors in such a manner that *any two different color classes have cardinalities that differ by at most one.*

Now, given binary random variables $X_1, X_2, \ldots, X_n$ with $\Pr[X_i = 1] = p$ for all i, suppose we can define an undirected graph G on the set $\{1, 2, \ldots, n\}$ of vertices, with the following property:

> for any set $S = \{i_1, i_2, \ldots, i_k\}$ that is an independent set in G, the random variables $X_{i_1}, X_{i_2}, \ldots, X_{i_k}$ are mutually independent.

Suppose further that the maximum degree of G is d. Prove that for $X = \sum_i X_i$ and any $\delta > 0$,

$$\Pr[X \geq np(1 + \delta)] \leq (d + 1) \cdot F^+(\lfloor n/(d + 1)\rfloor \cdot p,\ \delta);$$

as usual, F^+ is the Chernoff upper-tail function. **(5 points)**

2. Recall that a random variable X has a Poisson distribution with mean μ if X takes on only non-negative integer values, with $\Pr[X = i] = e^{-\mu}\mu^i/i!$ for $i \geq 0$. (Verify for yourself that $\mathbf{E}[X] = \mu$; you don't need to prove this as part of this problem.) Use exponential moment-generating functions to get an upper bound on $\Pr[X \geq \mu(1 + \delta)]$ for $\delta > 0$. Do not express the result as an infinite sum etc.: express it in closed form. Do you see a similarity with the Chernoff bound we derived in class? **(5 points)**

3. A random variable X takes values in a given interval $[a, b]$, and its mean is some unknown quantity μ, which we want to estimate as follows. Given error paremeters ϵ and α such that $0 < \epsilon, \alpha \leq 1/2$, we want a randomized algorithm A that will output a value μ' such that $\Pr[|\mu' - \mu| \leq \epsilon] \geq 1 - \alpha$; the only randomness in this probabilistic statement is in the random choices made by the algorithm A. The only way to access X is that we have the ability to generate it any number of times as desired, independently.

Suggest a simple sampling-based algorithm for this task, and give a number of samples that is sufficient for this estimation task. (This number should be a function of a, b, ϵ and α only.) **(5 points)**