CSC 2420S 2012, Assignment 3
Due: April 25

1. Let $G = (V, E)$ be an n node regular graph of degree 3.

   (a) Suppose we independently remove nodes (and their adjacent edges) with probability 4/5. By looking at the expected number of nodes and edges that remain and using part (b) argue that G has an independent set of size at least $\frac{7}{50}n$.
   Note: as has been pointed out to me, there is a simple argument that there must be an independent set of size at least n/4 for graphs of degree at most 3. The goal of this question is to give another example of the probabilistic method noting again how it can lead immediately to an algorithm.

   (b) Suppose we follow this random removal process by additonally removing an (arbitrary) end point from each remaining edge. For a fixed $\epsilon > 0$, argue (perhaps using both Markov and Chernoff bounds) why (for sufficiently large n) this then leads to a randomized polynomial time algorithm that always produces an independent set S and with high probability $|S| \geq (1-\epsilon)\frac{7}{50}n$. Your argument should show that the failure probability is at most c^{-n} for some $c > 1$.

   (c) Using the Lovasz Local Lemma prove that there is a non zero probability that the nodes remaining after the random removal process (but without removing any further nodes as in part (b) above)will form an independent set. (This is not to say that the remaining nodes are a large independent set.) Hint: The remaining set is an independent set iff all edges were removed during the process.
   NOTE: One can establish easily this fact without use of the local lemma. The point of the exercise (from which this was taken) was to show an input that does not have a big independent set and yet the removal process will almost surely (except with some exponentially small probability) end with a large set of vertices. This in turn then implies that the LLL probability must be exponentially small. I am only asking you this question to see how to formulate the problem in terms of the LLL.

2. Let N be prime. For a, b chosen randomly in Z_N, let $r_i = (ai+b)\ modN$.

   The following results were used in showing how to improve the error bound by repeated trials using pairwise independent randomness of a 1-sided randomized algorithm.

   (a) Show that the $\{r_i\}$ are uniformly distributed and pairwise independent; that is for all $x \in Z_N, Prob[r_i = x] = 1/N$ and for all $x, y \in Z_N$ and $i \neq j\ modN \wedge Prob[r_i = x, r_j = y] = 1/N^2$.

   (b) Let f be a $\{0,1\}$ function and let $Y = \sum_{i=1}^{t} f(r_i)$, for $t \leq N$. Show that the variance $Var[Y] \leq t/4$.

3. Consider a random walk on a connected d-regular non bipartite n node graph $G = (V, E)$ defined by a doubly stochastic matrix matrix $P = (p_{ij})$ whose non zero entries correspond to the edges in G; that is, $p_{ij} \geq 0$ for all i, j, $\sum_j p_{ij} = 1$ for all i, $\sum_i p_{ij} = 1$ for all j, and $(i, j) \in E$ implies $p_{ij} > 0$. Hence this defines an ergodic Markov chain.

   (a) Show that the stationary distribution π (satisfying $\pi P = \pi$ and π is a probability row vector) is the uniform distribution. ; i.e. $\pi_i = 1/n$ for all i.

(b) Assume $1 = \lambda_1 > \lambda_2 \geq \ldots \geq \lambda_n \geq 0$ are the (left) eigenvalues of P with $(e_1, e_2, \ldots, e_n)$ an orthonormal basis of corresponding eigenvectors. That is, $e_i P = \lambda_i P$. Suppose that the initial state of the random walk is defined by a probability vector p_0 and let p_t be the probability distribtion after t steps of the random walk defined by P. Show that $||p_t - \pi|| \leq (\lambda_2)^t ||p_0 - \pi||$ where $||v||$ is the Euclidean norm $\sqrt{\sum_i v_i^2}$ and π is the stationary distribution.