CSIS8601: Probabilistic Method & Randomized Algorithms
Homework 1 **Due Date:** 16 Sept 2009

Rules: Discussion of the problems is permitted, but writing the assignment together is not (i.e. you are not allowed to see the actual pages of another student).

Please justify your arguments carefully. Unless otherwise stated, you may assume that there is an underlying probability space $(\Omega, \mathcal{F}, Pr)$, and all random variables take discrete values.

1. (5 points) Prove the union bound: $Pr(\cup_i A_i) \leq \sum_i Pr(A_i)$.
(Hint: Use induction and the fact $Pr(A) + Pr(B) = Pr(A \cup B) + Pr(A \cap B)$.)

2. (5 points) Prove the linearity of expectation: $E[\sum_i a_i X_i] = \sum_i a_i E[X_i]$.

3. (10 points) Suppose $X, Y : \Omega \to \mathbb{Z}$ are independent random variables taking integer values. In this question, we shall prove the basic result $E[XY] = E[X]E[Y]$.

   Prove from first principle (without using any results not proved in class) that
   $\sum_{\omega \in \Omega} X(\omega)Y(\omega)Pr(\omega) = \sum_{\omega_1 \in \Omega} X(\omega_1)Pr(\omega_1) \sum_{\omega_2 \in \Omega} Y(\omega_2)Pr(\omega_2)$.

4. (30 points) **Max Cut.** Let $G = (V, E)$ be a graph. Recall the randomized algorithm mentioned in class for finding a cut $C \subset V$ for the graph G, namely, a point $v \in V$ is included in C independently with probability $\frac{1}{2}$. Let $E(C) := \{\{u, v\} \in E : u \in C, v \in V \setminus C\}$ be the edges in the cut. It is shown that $E[|E(C)|] = \frac{|E|}{2}$. The point of this question is to show that the algorithm finds a large cut with at least a constant probability.

   (a) Define the random variable $Y := |E| - |E(C)|$. Compute $E[Y]$.

   (b) Find a suitable upper bound for $Pr[Y > \frac{3|E|}{4}]$, and conclude that the randomized algorithm finds a cut of size at least $\frac{|E|}{4}$ with probability at least $\frac{1}{3}$.

5. (50 points) **Max 3-SAT.** Let ϕ be a 3-CNF formula with m clauses and n variables. Recall the basic randomized procedure mentioned in class for maximizing the number of satisfied clauses in the formula ϕ. In that procedure, each variable takes the value TRUE independently with probability $\frac{1}{2}$. Assume that generating a random TRUE/FALSE value takes 1 random bit. Hence, the basic random procedure needs n independent random bits. We showed in class that the expected number of satisfied clauses is $\frac{7m}{8}$.

   The goal of this question is to design another randomized algorithm with better guarantees. Let $0 < \epsilon < 1$ and $0 < \delta < 1$. We shall design a randomized algorithm that, with failure probability at most δ, returns an assignment such that the number of satisfied clauses is at least $\frac{7-\epsilon}{8}m$.

   (a) Give an upper bound on the failure probability that the basic randomized procedure returns an assignment that satisfies less than $\frac{7-\epsilon}{8}m$ clauses.

(b) Show that by repeating the basic randomized procedure, it is possible to obtain a better randomized algorithm with failure probability at most δ. Compute the number of independent random bits used by your algorithm.

(Hint: You might find the following inequality useful: for $0 < \epsilon < 1$, $1 + \epsilon \geq e^{\frac{\epsilon}{2}}$.)