CS174 Sp2001 **Homework 3** **due: Feb 8, 2001**

This homework is due by 5pm on Thursday Feb 8th. Please hand it to the CS174 homework box on the second floor of Soda Hall.

1. So far we have used probabilities to study occupancy problems. We can still use probabilities to count outcomes by multiplying by the total number of outcomes. If m balls are placed in n bins randomly:

   (a) How many distributions (sample points) are there? i.e. how many ways are there to distribute m balls into n bins? You should assume that balls and bins are distinguishable.

   (b) How many distributions give k balls in bin 1?

   (c) What is the number of distributions where exactly two bins are empty?

2. Suppose you have n items to store in a hash table and that the hash function behaves like a random function. That is, hashing an item is like assigning it to a random location in the hash table. How large should the hash table be to have probability at least a half for a collision (two items hashing to the same location)?

3. Use tail bounds to give an upper limit on the probability that after 1000 tosses of a fair coin, there are more than 800 heads. Give an answer for both Markov and Chebyshev bounds.

4. Suppose 48 balls are independently and randomly allocated to 12 bins, and let X_i be the number of balls in bin i. Assume the Poisson distribution is a good approximation to the distribution of X_i, and let $\lambda = m/n$.

   (a) Use the bound derived in class for the tail of a Poisson distribution (lecture 5) to compute an upper bound on the probability $\Pr[X_i \geq 20]$. Compare with Markov and Chebyshev bounds for the same probability.

   (b) Now consider the probability $\Pr[X_i \geq k]$ as a function $f(k)$. Give a big-O bound for $f(k)$ as a function of k using the Poisson formula derived in class, and then for Markov and Chebyshev.