Stats 318 **Spring 2020**

Homework 5

Due Thursday, June 4

The problem of counting the number of solutions to a knapsack instance can be defined as follows: given items with sizes $a_1, a_2, \ldots, a_n > 0$ and an integer $b > 0$, find the number of vectors $x = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$ such that $\sum_{i=1}^n a_i x_i \leq b$. The number b can be thought as the size of the knapsack, and the x_i's denote whether or not each item is put into the knapsack. Counting solutions corresponds to counting the number of different sets of items that can be placed in the knapsack without exceeding its capacity.

1. A naive way of counting the number of solutions is to repeatedly choose $x = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$ uniformly at random and return 2^n times the fraction of samples that yield valid solutions. Explain why this is not a good solution; in particular, that it will work poorly when each a_i is 1 and b is about $\sqrt{n}$.

2. Consider a Markov chain $X_0, X_1, \ldots$ on $\{0,1\}^n$. Suppose that $X_t = x$. At each step, we choose $i \in \{1, \ldots n\}$ uniformly at random. If $x_i = 1$, then X_{t+1} is obtained from X_t by setting x_i to zero. If $x_i = 0$, then X_{t+1} is obtained from X_t by setting x_i to one if doing so maintains the condition $\sum_{i=1}^n a_i x_i \leq b$. Otherwise $X_{t+1} = X_t$.

   Argue that this Markov chain has a uniform stationary distribution whenever $\sum_{i=1}^n a_i > b$. Be sure to argue that the chain is irreducible and aperiodic.

3. Suppose that the distribution $\mu^{(t)}$ of the Markov chain obeys $d_{TV}(\mu^{(t)}, \pi) \leq \rho$ for all $\rho > 0$ in time t which is polynomial in n and in $\log 1/\rho$. Then show that one can get an (ϵ, δ) approximation to our problem that is polynomial in $1/\epsilon$, $\log 1/\delta$ and n.

   Hint: This is very much like the problem we studied in class (counting the number of q-colorings). To set up the problem. Assume without loss of generality that $a_1 \leq a_2 \leq \ldots \leq a_n$. Let $b_0 = 0$ and $b_i = \sum_{j=1}^i a_j$. Let N be the number we wish to estimate and N_j be the number of x obeying $\sum_{i=1}^n a_i x_i \leq b_j$. Put k to be the smallest integer such that $b_k \geq b$ and consider

$$N = \frac{N}{N_{k-1}} \times \frac{N_{k-1}}{N_{k-2}} \times \ldots \times \frac{N_1}{N_0} \times N_0.$$

   You will need to argue that N_{j-1}/N_j is not too small. Specifically, show that $N_j \leq (n+1)N_{j-1}$.