Final Exam.

June 12, 2017

1. Multinomial distribution (2 pts each)

   A binomial random variable handles experiments that have two types of outcomes. The number of Heads appeared in n coin flips is an example of a binomial random variable. Multinomial distribution is an extension of binomial distribution where experiments have more than two types of outcomes. Die rolling is a case of multinomial distribution. Assume a biased die that comes up with k (K=1, 2,..., 6) with p_k and $\sum_{i=1}^{6} p_i = 1$.

   A) Let X_i be the number of times that i (i=1,2,...,6) appears out of n die rolls. Prove that Pr$[X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n] = \frac{n!}{x_1!\ x_2!\ \cdots\ x_6!}\ p_1^{x_1} \cdots p_6^{x_6}$.

   B) Given that $X_6 = k\ \ (0 \leq k \leq= n)$, Derive the conditional probability Pr$[X_1 = x_1, X_2 = x_{2,}.., X_5 = x_5\ \mid X_6 = k]$.

2. Continuous random variables & Inequality (2, 2, 3 pts)

   A) X~Uni[0, 1], is a continuous uniform random variable over [0, 1]. Compute E[X] and Var[X].

   B) Y~Uni[0,1] is another uniform random variable independent of X. Let Z=X-Y. Find E[Z] and Var[Z]. Also find distribution and density functions of Z.

   C) X_i~Uni[0,1], i=1, 2,.., n are independent uniform random variables. Let X=X1+...+Xn. Find Markov and Chebyshev inequalities for Pr[X >= 0.7*n].

3. BB model (2, 2, 3 pts)

   Assume that n balls are thrown randomly into n bins.

   A) Compute the exact probability that each bin has exactly one ball.

   B) Compute the probability (both exact and Poisson cases) that a bin has three balls.

   C) Compute the expected number of bins that has three balls in Poisson approximation. Also compute the bound of the exact case. Your bound should be as tight as possible.

4. Probabilistic method. (3 pts)

   Let G be a graph with an even number of vertices. Show that the vertices of G can be partitioned into two parts A and B of equal size such that the number of edges between A and B is ***strictly larger*** than $|E|/2$. Describe a method to find a such partition.

5. Markov Chain & Random Walk (3 pts each)

   A) Assume that there are two biased coins and one fair coin. The probabilities of Heads of Coin1, 2 and 3 are 0.4, 0.5 and 0.6, respectively. At the beginning, a player picks Coin1 and flips the coin until Heads appears for the first time. Then the player picks coin2 or coin3 randomly and flips the selected coin until the first Heads. This process continues forever. Compute the stationary distribution of playing with Coin1.

B) An *n* x *n* matrix P with entries *Pi, j* is called stochastic if all entries are nonnegative and if the sum of the entries in each row is 1. It is called doubly stochastic if, additionally, the sum of the entries in each column is 1. Show that the uniform distribution is a stationary distribution for any Markov chain represented by a doubly stochastic matrix.

6. Monte Carlo Method (3, 4 pts)

The problem of counting the number of solutions to a knapsack problem can be defined as follows: Given n items with sizes $a_1, a_2, \ldots, a_n$ and an integer *b* > 0, find the number of subsets S $\subseteq$ {1,2,…,n} such that $\sum_{i \in S} a_i \ \leq b.$ The number *b* can be thought of as the size of a knapsack. For example, assume that there are five items whose sizes are 2, 5, 7, 3, and 4, respectively and b=10. Then a subset {1, 4, 5} satisfies the constrains because 2+3+4 <= 10 while a subset {2, 3} fails the restriction because 5+7 > 10. Counting solutions corresponds to counting the number of different sets of items that can be placed in the knapsack without exceeding its capacity.

A) A naïve way of counting the number of solutions to this problem is to repeatedly choose subsets uniformly at random, and return the 2^n times the fraction of samples that yield valid solutions. Argue why this is not a good strategy in general; in particular, argue that it will work poorly when each $a_i = 1$ and *b*= $\sqrt{n}$.

B) Consider a Markov chain Xo, X1, … on subsets $\subseteq$ {1,2,…,n}. Suppose *Xj* has k items and $X_j = \{j_1, j_{2,\ldots,} j_k\}$. At each step, the Markov chain chooses *i* ∈ $[1, \mathrm{n}]$ uniformly at random. If i ∈ X_j, then X_{j+1} is obtained from X_j by removing i from X_j. If *not*. then X_{j+1} is obtained from X_j by adding i to X_j if doing so maintains the restriction $\sum_{i \in X_{j+1}} a_i \ \leq b.$ Otherwise, X_{j+1} =X_j. 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.