UC Berkeley
Department of Electrical Engineering and Computer Sciences

EE126: Probability and Random Processes

Problem Set 3
Fall 2015

Issued: Thursday, September 10, 2015 **Due:** 9am, Thursday, September 17, 2015

Problem 1. This problem will explore an important probabilistic concept of clustering that is widely used in machine learning applications today. Consider n students. For each pair of students, say student i and student j, they are friends with probability p, independently of other pairs. Here we assume that friendship is mutual, then we can see that the friendship among the n students can be represented by an undirected graph G. Let $N(i)$ be the number of friends of student i and $T(i)$ be the number of triangles attached to student i. We define the *clustering coefficient* $C(i)$ for student i as follows:

$$C(i) = \frac{T(i)}{\binom{N(i)}{2}}.$$

Clustering coefficient is not defined for the students who have no friends. An

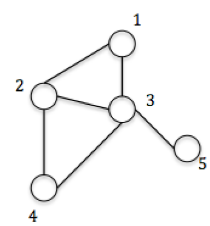

Figure 1: Friendship and clustering coefficient.

example is shown in Figure 1. Student 3 has 4 friends: 1,2,4,5, and there are two triangles attached to student 3, i.e., triangle 1-2-3 and triangle 2-3-4. Therefore $C(3) = \frac{2}{\binom{4}{2}} = \frac{1}{3}$. Find $\mathbb{E}[C(i)|N(i) \geq 2]$.

Problem 2. Suppose that you are applying to graduate school, and you want to improve your GRE writing score. Assume that on a given day, your score is distributed uniformly on the set $\{3, 3.5, 4, 4.5, 5, 5.5, 6\}$ independent of other days. You plan to take 3 exams and report the highest score: $X = \max\{X_1, X_2, X_3\}$, where X_i is the score in the i-th exam you take.

(a) Calculate the PMF of X.

(b) How much your expected reported score will change compared to taking only 1 exam?

Problem 3. n graduate students at a Berkeley Research Lab come to work by bicycles every day, parking their bikes in an unlocked bike rack in their lab. On a given day, after a hard day's work, the students begin leaving one by one, taking their bikes out of the rack on their way out. The first student has forgotten what his bike looks like, and picks up one of the n bikes from the rack uniformly at random. Subsequently, the other $(n-1)$ students who leave one by one, follow the "honest if possible" policy of picking up their own bike if it is there, else picking a bike uniformly at random from the remaining collection in the rack.

(a) What is the probability that the first student finds his/her own bike?

(b) List the bikes that the last student may possibly take. What is the probability that the last student finds his/her own bike?

(c) List the bikes that the ith student may possibly take (for $i = 2, \dots, n-1$). What is the probability that the this student finds his/her own bike?

(d) What is the expected number of students who go home with their own bikes? (What is this as n gets large)?

Problem 4. Two points are picked uniformly at random in the interval $[0, L]$.

(a) What is the expected distance between these points?

(b) Suppose that the selected points are X_1 and X_2 such that $0 \leq X_1 \leq X_2 \leq L$. What is the probability that a triangle can be formed from the lengths X_1, $X_2 - X_1$ and $L - X_2$?

Problem 5. Figure 2 shows the joint density $f_{X,Y}$ of random variables X and Y.

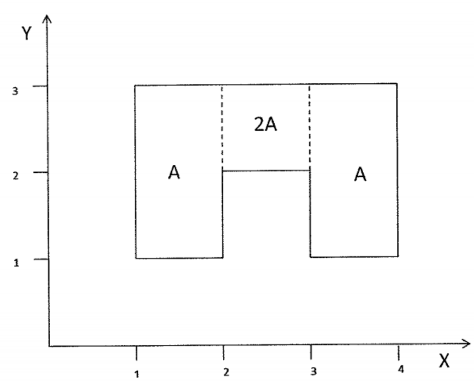

Figure 2: Joint pdf of X and Y.

(a) Find A and sketch f_X, f_Y and $f_{X|X+Y\leq 3}$.

(b) Find $\mathbb{E}[X|Y=y]$ for $1 \leq y \leq 3$ and $\mathbb{E}[Y|X=x]$ for $1 \leq x \leq 4$.

(c) Find $\text{cov}(X, Y)$.

Problem 6. In class, we discussed the Buffon's needle experiment. Let's explore a small extension here. A needle of length ℓ is dropped randomly on a plane surface that is partitioned in rectangles by horizontal lines that are a apart and by vertical lines that are b apart. Suppose that $\ell < a$ and $\ell < b$. What is the expected number of rectangle sides crossed by the needle? What is the probability that the needle crosses at least one side of a rectangle?