Homework 2 – Due Thursday, September 12, 2024

Page limit You can submit **at most** 2 pages per problem, even if the problem has multiple parts. If you submit a longer solution for some problem, only the first 2 pages will be graded.

Reminder Collaboration is permitted, but you must write the solutions *by yourself without assistance*, and be ready to explain them orally to the instructor if asked. You must also identify your collaborators and whether you gave help, received help, or worked something out together. Getting solutions from outside sources such as the Web or students not enrolled in the class is strictly forbidden.

Exercises Please practice on exercises in Chapter 1 of Mitzenmacher-Upfal.

Problems

1. (**Marble bag**) You are given a bag containing ten marbles and are told that the number of red marbles in the bag is equally likely to be any number between 0 and 10 inclusively; the other marbles are green. So, in particular, the probability that all ten marbles are red is $\frac{1}{11}$.

   (a) Suppose you pick a marble at random from the bag and it is red. Find the probability that all the marbles in the bag are red, *given* that the marble you picked is red. In other words, we want to compute the conditional probability $\Pr[A|R]$, where R is the event that the chosen marble is red, and A is the event that all marbles in the bag are red.

   (b) Is the probability $\Pr[A|R]$ larger or smaller than $\Pr[A]$? Explain why it is reasonable.

   (c) Now suppose you pick m marbles from the bag uniformly at random without replacement, where $m \in \{1, \ldots, 10\}$. Let R_m be the event that all of them are red. What is the probability of R_m?

   (d) You are given the second bag that satisfies the same condition as the first bag. Moreover, the number of red marbles in the first bag and the number of red marbles in the second bag are independent. Let m and R_m be defined as before (for the first bag). You pick m marbles from the second bag uniformly at random without replacement. Let R'_m be the event that all marbles picked from the second bag are red. Are R_m and R'_m independent? **Justify your answer.**

   (e) You roll a standard fair 6-sided die. Let random variable D be the number (from 1 to 6) that you rolled. Then you pick D marbles from the first bag and D marbles from the second bag independently and uniformly at random without replacement. Let E_1 be the event that all marbles selected from the first bag are red, and E_2 be the event that all marbles selected from the second bag are red. Are E_1 and E_2 independent? **Prove that your answer is correct.**

2. (**Mice with random genes**) There are n genes that your roommate is trying to investigate. For simplicity, let's call them $1, \ldots, n$. Whenever your roommate breeds a mouse, the mouse has mutations in a subset of the n genes. (As far as you know, all the experiments are happening in her lab, not in your apartment).

(a) First, a mouse Adam is born. Suppose his set A of mutations is generated as follows: a fair coin is flipped independently for each gene; if the coin lands heads then the gene is added to A, and otherwise it is not. Argue that the resulting set A is equally likely to be any one of the 2^n possible subsets of $\{1, \ldots, n\}$.

(b) Now, consider two mice, Adam and Eve. Suppose their sets of mutations, A and E, are chosen independently and uniformly at random from all the 2^n subsets of $\{1, \ldots, n\}$. Determine the probability that A and E don't have any elements in common. *Hint: Use part (a).*

(c) Three new mice, called Xavier, Yasmin, and Zeus, are born. Suppose that their sets X, Y and Z of mutations are chosen independently and uniformly at random from all the 2^n subsets of $\{1, \ldots, n\}$. Determine the probability that each of the n mutations appears in at least one of the three sets.

3. (**Verifying matrix multiplication**)

(a) Recall that in the problem of verifying matrix multiplication, we are given $n \times n$ matrices A, B, C, and we want to check whether $A \cdot B = C$. Your friend Daniel proposes the following algorithm for the problem: *Let $D = A \cdot B - C$. Pick i and j uniformly and independently from $[n]$ and compute D_{ij}. Accept if D_{ij} is 0 and reject otherwise.*

i. Give the best upper bound you can on the probability that this algorithm accepts incorrectly (for a worst-case input).

ii. Give the best upper bound you can on the number of independent runs of this algorithm you need to ensure that the error probability is at most $1/3$.
Hint: Use the inequality $1 - x \leq e^{-x}$.

iii. State the running time of the resulting algorithm (with the number of runs you specified) using asymptotic notation.

(b) In the algorithm from class for verifying matrix multiplication, we chose a uniformly random vector $\bar{r}$ from $\{0, 1\}^n$. Suppose we choose $\bar{r}$ from $\{0, 1, \ldots, v - 1\}^n$ instead. Prove a better bound (than $1/2$) on the probability of success of one check performed by the algorithm.

(c) Suppose your prior belief is that a given matrix multiplication identity is correct with probability $1/2$. The analysis that starts at the bottom of page 11 of the textbook demonstrates how your prior belief about the probability that the identity is correct changes after i runs of the algorithm that do not find any mistakes. Modify this analysis to work with your bound from part (b).