Problem Set for Unit 11

Prof. Amit Chakrabarti
Computer Science Department
Dartmouth College

Here are some problems on **basic probability**. Before working on these problems, it is **absolutely required** that you read Chapter 17 from the [LLM] book. Several of these problems are from the [LLM] book, but may be slightly modified, so read the wording carefully.

The symbols $\mathbb{N}$, $\mathbb{N}^+$, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ have their usual meanings.

PS11-1

Solve the following problems, each of which asks you to compute a probability. Be systematic and use the Four-Step Method (resist the urge to jump right into the calculations).

a. What is the probability that '0' does not appear among k digits chosen independently and uniformly at random?

b. A box contains 90 good and 10 defective screws. What is the probability that if we pick 10 screws from the box, none will be defective?

c. First one digit is chosen uniformly at random from $\{1,2,3,4,5\}$ and is removed from the set; then a second digit is chosen uniformly at random from the remaining digits. What is the probability that an odd digit is picked the second time?

d. Suppose that you *randomly* permute the symbols $1, 2, \ldots, n$ (that is, you select a permutation uniformly at random). What is the probability that the symbol k ends up in the ith position after the permutation?

e. A fair coin is flipped n times. What is the probability that all the heads occur at the end of the sequence? (If no heads occur, then "all the heads are at the end of the sequence" is vacuously true.)

f. You hold the following hand of five cards from a standard 52-card deck: $\{\spadesuit 2, \spadesuit 3, \spadesuit 9, \heartsuit \mathrm{K}, \diamondsuit 5\}$. If you discard the two non-spade cards from this hand and replace them with two uniformly random cards from the rest of the deck, what is the probability that your hand now consists of all spades?

PS11-2

Solve the following problems, each of which asks you to compute a probability. Be systematic and use the Four-Step Method (resist the urge to jump right into the calculations).

a. You are dealt a poker hand of 5 cards drawn at random from a well-shuffled standard deck of 52 cards. The hand is called a *full house* if it has three cards of one rank and two of another rank (e.g., three kings and two 7s). What is the probability that your hand is a full house?

b. A standard bag of Scrabble tiles has 100 tiles, exactly two (2) of which are *blank tiles*. Blanks are valuable assets in the game, since they can be turned into whatever letter you want.

You have just begun a game of Scrabble, drawing your first rack of seven (7) tiles from a full bag. What is the probability that your rack contains a blank?

c. How does the answer to the above question change if your opponent is to make the first move, so your opponent picks seven random tiles first and *then* you get to pick seven from the 93 remaining tiles?

PS11-3 *HW*

The New York Yankees and the Boston Red Sox are playing a two-out-of-three series. In other words, they play until one team has won two games. Then that team is declared the overall winner and the series ends. Assume that the Red Sox win each game with probability $3/5$, regardless of the outcomes of previous games.

Answer the questions below using the Four-Step Method. You can use the same tree diagram for all three problems.

a. What is the probability that a total of 3 games are played?

b. What is the probability that the winner of the series loses the first game?

c. What is the probability that the *correct* team wins the series? [6 points]

PS11-4

Which is more likely: rolling a total of 8 when two dice are rolled or rolling a total of 8 when three dice are rolled? Answer the same question with 8 changed to 9. Assume that all dice involved are standard, six-sided, fair dice.

PS11-5 *HW*

To determine which of two people gets a prize, a coin is flipped twice. If the flips are H (Head) followed by T (Tail), the first player wins. If the flips are T followed by H, the second player wins. However, if both flips land the same way, the flips don't count and the whole process starts over.

Assume that each flip results in H with probability p, regardless of what happened on other flips. Use the Four-Step Method to find a simple formula for the probability that the first player wins. What is the probability that neither player wins? [7 points]

Hint: The tree diagram and sample space are infinite, so you're not going to finish drawing the tree. Try drawing only enough to see a pattern. Summing all the winning outcome probabilities directly is cumbersome. However, a neat trick solves this problem—and many others. Let s be the sum of all winning outcome probabilities in the whole tree. Notice that *you can write the sum of all the winning probabilities in certain subtrees as a function of* s. Use this observation to write an equation in s and then solve it.

PS11-6 *HW*

We play a game with a deck of 52 regular playing cards, of which 26 are red and 26 are black. I randomly shuffle the cards and place the deck face down on a table. You have the option of "taking" or "skipping" the top card. If you skip the top card, then that card is revealed and we continue playing with the remaining deck. If you take the top card, then the game ends. If we get to a point where there is only one card left in the deck, you must take it.

You win if the card you took was revealed to be black, and you lose if it was red.

Prove that you have no better strategy than to take the top card—which means your probability of winning is exactly 1/2. [8 points]

Hint: Prove by induction the more general claim that for a randomly shuffled deck of n cards that are red or black—not necessarily with the same number of red cards and black cards—there is no better strategy than taking the top card. To precisely state this more general claim, first work out your probability of winning if you simply take the top card.

PS11-7

The Disjoint Sum Rule for probabilities says that if A and B are two disjoint events then $\Pr[A \cup B] = \Pr[A] + \Pr[B]$. As you know, this is a consequence of the Sum Principle from basic counting. There is also the extended version of the Disjoint Sum Rule, which applies to multiple events. This states that if $E_1, E_2, \ldots, E_n$ are pairwise disjoint events, then

$$\Pr[E_1 \cup E_2 \cup \cdots \cup E_n] = \Pr[E_1] + \Pr[E_2] + \cdots + \Pr[E_n],$$

or in shorthand,

$$\Pr\left[\bigcup_{i=1}^{n} E_i\right] = \sum_{i=1}^{n} \Pr[E_i].$$

Starting with the Disjoint Sum Rule, derive (i.e., prove) each of the following other useful rules for reasoning about probability.

a. Difference Rule: $\Pr[A - B] = \Pr[A] - \Pr[A \cap B]$.

b. Complement Rule: $\Pr[\overline{A}] = 1 - \Pr[A]$.

c. Inclusion-Exclusion: $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$.

d. Union Bound for Two Events: $\Pr[A \cup B] \le \Pr[A] + \Pr[B]$.

e. Monotonicity: If $A \subseteq B$, then $\Pr[A] \le \Pr[B]$.

State your reasoning for each step of each proof. Some of the proofs will be simple calculations with just a couple of steps. Venn diagrams could be useful.

PS11-8 ^{EC}

In the birthday problem discussed in class, there are n students with birthdays distributed *uniformly* across the d days in a year. We showed that the probability of some two students sharing a birthday is "paradoxically" high. Suppose we remove the uniformity assumption, so that not all days in $\{1, 2, \ldots, d\}$ are equally likely to be birthdays. Specifically, there are non-negative real numbers $p_1, p_2, \ldots, p_d$ with $p_1 + p_2 + \cdots + p_d = 1$ such that a student's birthday equals i with probability p_i. The birthdays of different students are still independent of one another.

Prove that in this more general situation, the probability of some two students sharing a birthday is at least as high as in the uniform case.