CS 70 Discrete Mathematics and Probability Theory
Summer 2011 Kamil HW 6

Due Tuesday, July 26, 4:59pm

1. (10 pts.) Pairwise vs mutual independence

In this problem we will see that pairwise independence of events does not imply mutual independence.

Two fair dice are thrown. Let A be the event that the number on the first die is odd, B the event that the number on the second die is odd, and C the event that the sum of the two numbers is odd.

(a) Show that the three events A, B, C are *pairwise independent* (i.e., each pair (A,B), (B,C) and (A,C) is independent).

(b) Show that the events A, B, C are *not* mutually independent.

2. (10 pts.) Discreet math

A non-profit wants to poll a sample of people to ask them whether they have ever had an extramarital affair. This being an extremely sensitive subject, one obvious problem is that if the surveyers ask this question straight-out, respondents may lie to avoid revealing personal information about their private lives.

The surveyers come up with the following clever scheme. They will ask the respondent to secretly roll a fair die. If the die comes up 1, 2, 3, or 4, the respondent is supposed to answer truthfully. If the die comes up 5 or 6, the respondent is supposed to answer the *opposite* of the truthful answer. The respondent is cautioned not to reveal what number came up on the die. Notice that if the respondent answers "Yes," this answer is not necessarily incriminating: for all the surveyer knows, this particular respondent might have rolled a 5 or 6 and might have never had an affair in his/her life.

Let p be the probability that, if we select a person at random, then they will have had an extramarital affair. (Of course, the surveyers do not know p; that is what they want to estimate.) Let q denote the probability that, if we select a person at random and have them follow the scheme above, then they will answer "Yes."

(a) Calculate a simple formula for q, as a function of p.

(b) Next, suppose the surveyers have estimated q. Now they want to solve for p. Find a simple formula for p, as a function of q.

3. (12 pts.) Practice with Bayes rule

In 1998, researchers at the Max Planck Institute conducted a survey of 24 doctors (doctors who had an average of 14 years of professional experience), to study how they understand risk. The doctors were presented with the following information:

> The probability that a woman has breast cancer is 0.008. If she has breast cancer, the probability that a mammogram will show a positive result is 0.90. If a woman does not have breast cancer the probability of a positive result is 0.07. Take, for example, a woman who has a positive result. What is the probability that she actually has breast cancer?

Each doctor was asked to estimate the probability that a woman whose mammogram comes back positive actually has breast cancer. One-third of the doctors said "0.90"; another one-third's estimates were in the range 0.50–0.80; one-sixth estimated something in the range 0.05–0.10; and another one-sixth estimated about 0.01.

(a) Calculate the correct probability that a woman has breast cancer, given that her mammogram shows a positive result. Show your work. Evaluate your answer to get a concrete number.

(b) What fraction of doctors were in the right ballpark? Which ones?

(c) For the doctors whose answers weren't close, how would you explain to them why the number you got in part (a) is correct, in terms they'd likely be able to understand? The doctors probably don't need to know how to get the exact answer, but they need to understand why your answer is approximately right.

(If your explanation involves mathematical formulas, references to Bayes theorem and conditional probability, or technical jargon, many doctors are probably going to have a hard time understanding what you're trying to say. How could you explain this in terms that an intelligent layperson would have a fighting chance of understanding?)

4. (12 pts.) Serve on the jury

In the OJ Simpson murder trial, OJ Simpson was accused of murdering his ex-wife, Nicole Simpson. The prosecution introduced evidence showing that OJ had previously abused Nicole. One of Simpson's defense lawyers, Alan Dershowitz, made the following argument in OJ Simpson's defense. Dershowitz stated that 1 in 1,000 women abused by their husbands are later killed by their abuser, so the fact that OJ Simpson had previously abused his wife is not relevant and should be disregarded. Assume for this problem that Dershowitz's 1 in 1,000 statistic is accurate.

(a) Are we entitled to conclude that there is only a $1/1000$ probability that OJ Simpson murdered Nicole? Why or why not?

(b) Suppose we select at random a woman who has been abused by her husband. Define the following events: M is the event that the woman is murdered at some point in her life; G is the event that the woman is murdered by her abuser at some point in her life. A plausible estimate is that 0.2% of abused women will be murdered by someone other than their abuser at some point in their life. Calculate the probability that the selected woman is murdered by her abuser, given that she is murdered.

(c) Based upon your answer to part (b), do you agree or disagree with Dershowitz's argument? Based upon your calculation, would you consider it relevant that OJ Simpson previously abused Nicole? Would you judge it more accurate to use the 1 in 1,000 number or the number you calculated in part (b)? Why?

5. (10 pts.) Probability models

Suppose a document has n letters and you are told the probability that each letter is typed wrongly is p.

(a) Explain why there is not enough information to calculate the probability that the entire document is typed correctly.

(b) Find the smallest possible value for this probability in terms of p. For what values of p is this zero? For the remaining values of p (i.e. when the minimum probability that the document is typed correctly is nonzero), specify the probability assignment to the underlying sample space for which the smallest value is achieved.

(c) Let $n = 2$. Suppose you are told additionally that q_1 is the conditional probability that the second letter is typed wrongly given that the first letter is typed wrongly, and q_2 is the conditional probability

that the second letter is typed wrongly given that the first letter is typed correctly. Do we now have a complete probability assignment on all the sample points of the experiment? If so, give it.

(d) What constraint(s) do q_1 and q_2 have to satisfy?

(e) Under what values of q_1 and q_2 is the probability that the document is typed correctly (i) equal to, (ii) larger than, and (iii) smaller than the probability when the events that the two letters are typed incorrectly are independent?

6. (14 pts.) A Very Small Example of Hashing

Suppose we hash three objects randomly into a table with three (labelled) entries. We are interested in the lengths of the linked lists at the three table entries.

(a) List all the outcomes in the sample space of the experiment. How many of them are there?

(b) Let X be the length of the linked list at entry 1 of the table. Write down X explicitly as a function on the sample space mapping to the real line (either in a figure as in class or as a list). Compute and plot the distribution and expectation of X.

(c) Let Y be the length of the *longest* linked list among all three. Write down Y explicitly as a function on the sample space mapping to the real line. Compute and plot the distribution and expectation of Y.

(d) Is the expectation of X larger than, equal to or smaller than that of Y?

(e) Compute the distribution of X for the general case when m objects are hashed randomly into a table of size n, i.e. give an expression for the probability that X takes on each value in its range. (Computing the distribution of Y for the general case is not so easy, so we won't ask you to do it!)

7. (14 pts.) Random Variables and Their Distributions

A biased coin with probability p landing with a head is flipped 4 times.

(a) Give the sample space and assign probabilities to the sample points.

(b) Let X be the total number of heads in the four flips. Draw a Venn diagram showing the five events $X = i, i = 0,1,2,3,4$ as well as the sample space and the sample points. Is X a random variable?

(c) Are the events $X = 1$ and $X = 2$ disjoint? Are they independent?

(d) Let Y be the first flip when a head appears. Draw a Venn diagram showing the four events $Y = i, i = 1,2,3,4$ as well as the sample space and the sample points. Is Y a random variable?

(e) Are the events $X = 3$ and $Y = 4$ disjoint? Are they independent? How about the events $X = 2$ and $Y = 2$? Disjoint? Independent?

(f) If X is a random variable, compute its distribution. If not, modify its definition appropriately to make it into a random variable and then compute its distribution.

(g) If Y is a random variable, compute its distribution. If not, modify its definition appropriately to make it into a random variable and then compute its distribution.

8. (10 pts.) Chopping up DNA

(a) In a certain biological experiment, a piece of DNA consisting of a linear sequence (or string) of 4001 nucleotides is subjected to bombardment by various enzymes. The effect of the bombardment is to randomly cut the string between pairs of adjacent nucleotides: each of the 4000 possible cuts occurs independently and with probability $\frac{1}{500}$. What is the expected number of pieces into which the string is cut? Justify your calculation.

[HINT: Use linearity of expectation! If you do it this way, you can avoid a huge amount of messy calculation. Remember to justify the steps in your argument; i.e., do not appeal to "common sense."]

(b) Suppose now that the cuts are no longer independent, but highly correlated, so that when a cut occurs in a particular place other cuts close by are much more likely. The probability of each individual cut remains $\frac{1}{500}$. Does the expected number of pieces increase, decrease, or stay the same? Justify your answer with a precise explanation.

9. (8 pts.) Bubblesort

In Bubblesort one exchanges adjacent elements if they are in the wrong order until the entire list is sorted. What is the expected number of exchanges for Bubblesort on a list that was initially ordered at random? That is, any of the $n!$ permutations is equally likely. (Assume that the elements are distinct.)

[HINT: any inverted pair requires one exchange to fix.]

Optional Problems

These problems are for extra practice, and they will not be graded, so don't turn them in! We will provide solutions, however.

10. (0 pts.) Grab 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 inclusive; the other marbles are green. So, in particular, the probability that all ten marbles are red is $\frac{1}{11}$.

Suppose you pick a marble at random from the bag and it is red. We want to calculate the probability that all the marbles in the bag are red, *given* that we have picked one red marble. 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. From Bayes' Rule (Note 11) we know that this probability can be written as

$$\Pr[A|R] = \frac{\Pr[R|A] \times \Pr[A]}{\Pr[R]}.$$

(a) What is $\Pr[R|A]$?

(b) Show that $\Pr[R]$ can be written as

$$\Pr[R] = \sum_{k=0}^{10} \Pr[R|W_k] \times \Pr[W_k],$$

where W_k is the event that the bag contains exactly k red marbles. [HINT: Note that the events W_k partition the probability space, i.e., these events are disjoint, and their probabilities sum to 1.]

(c) Use parts (a) and (b) to evaluate the probability $\Pr[A|R]$ that all marbles in the bag are red. Is this probability larger or smaller than it was before you picked the marble? Does this seem reasonable?

11. (0 pts.) Bonferroni's inequalities

(a) For events A, B in the same probability space, prove that

$$\Pr[A \cap B] \geq \Pr[A] + \Pr[B] - 1.$$

(b) Generalize part (a) to prove that, for events $A_1, \ldots, A_n$ in the same probability space (and any n),

$$\Pr[A_1 \cap \ldots \cap A_n] \geq \Pr[A_1] + \cdots + \Pr[A_n] - (n-1).$$

12. (0 pts.) Function of a Random Variable

Consider a random variable X defined on a sample space Ω. Let $Y = X^2$.

(a) Explain why Y is also a random variable defined on Ω. Draw a diagram like Figure 1 in Note 13 to visualize the relationship between Ω, X and Y.

(b) You only know the distribution of X but not how X is defined on Ω. Is that enough information for you to compute the distribution of Y? If so, express the distribution of Y in terms of the distribution of X. If not, specify what additional information you will need. You may want to draw a Venn Diagram relating various events.