CS 70 Spring 2006 | Discrete Mathematics for CS Vazirani | HW 10

Due Friday April 21

1. Chopping Up DNA

(a) In a certain biological experiment, a piece of DNA consisting of a linear sequence (or string) of 2001 nucleotides is subjected to bombardment by various enzymes. The effect of the bombardment is to randomly cut the string between pairs of nucleotides: each of the 2000 possible cuts occurs independently and with probability $\frac{1}{400}$. 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. Do not appeal to "common sense".

(b) Suppose 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 cut remains $\frac{1}{400}$. Does the expected number of pieces increase, decrease or stay the same? Justify your answer with a precise explanation.

2. Independent Random Variables

Two random variables X and Y on the same probability space are said to be *independent* if the events "$X = a$" and "$Y = b$" are independent for all pairs of values a,b. (i.e. the value taken on by X has no effect on the distribution of Y, and vice versa).

(a) Show that, for independent random variables X, Y, we have

$$\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$$

(Hint: Show first - carefully! - that, even if the r.v.'s are *not* independent, $\mathbf{E}[XY] = \sum_a \sum_b ab \cdot \Pr[X = a \wedge Y = b]$.)

(b) Give a simple example to show that the conclusion of part (a) is not necessarily true when X and Y are not independent.

3. Random Variables in GF_p

Let the random variables X and Y be distributed independently and uniformly at random in the set $\{0, 1, \ldots, p-1\}$, where $p > 2$ is a prime.

(a) What is the expectation $\mathbf{E}[X]$?

(b) Let $S = (X+Y) \bmod p$ and $T = XY \bmod p$. What are the distributions of S and T?

(c) What are the expectations $\mathbf{E}[S]$ and $\mathbf{E}[T]$?

(d) By linearity of expectation, we might expect that $\mathbf{E}[S] = (\mathbf{E}[X] + \mathbf{E}[Y]) \bmod p$. Explain why this does not hold in the present context; i.e. why does the value for $\mathbf{E}[S]$ obtained in part (c) not contradict linearity of expectation?

(e) Since X and Y are independent, we might expect that $\mathbf{E}[T] = \mathbf{E}[X]\mathbf{E}[Y] \bmod p$. Does this hold in this case? Explain why/why not.

4. (16 pts.) How to beat the heat

It's a hot summer day in the Central Valley. Three children Alice, Bob, and Carlos are engaged in a three-way duel with water balloons. They start by drawing lots to determine who throws first, second, and third, then take their places at the corners of an equilateral triangle. They agree to throw single water balloons in turn and continue in the same cyclic order until two of them have been soaked. Each player may throw at any other in his or her turn. You should assume the following: all the children have an essentially infinite supply of ammunition; a water balloon explodes on contact, drenching its target (who then leaves the game); when a water balloon misses its target, it explodes far enough away not to get anyone wet.

All three know that Alice always hits her target, Bob is 75% accurate, and Carlos is 50% accurate. Of course, if for some reason any of them deliberately decides to miss they can do so with certainty. Suppose that Bob has drawn the first shot, and Carlos second. What is Bob's best strategy, and what is the chance that he comes out the eventual winner? What about Alice?

5. Geometric Distribution

James Bond is imprisoned in a cell from which there are three possible ways to escape: an air-conditioning duct, a sewer pipe and the door (which is unlocked). The air-conditioning duct leads him on a two-hour trip whereupon he falls through a trap door onto his head, much to the amusement of his captors. The sewer pipe is similar but takes three hours to traverse. Each fall produces temporary amnesia and he is returned to the cell immediately after each fall. Assume that he always immediately chooses one of the three exits from the cell with probability $\frac{1}{3}$. On the average, how long does it take before he realizes that the door is unlocked and escapes?

Hint: If you are doing complicated calculations you're taking the wrong approach.

6. Extra Credit

I think of two distinct real numbers between 0 and 1 but do not reveal them to you. I now choose one of the two numbers at random and give it to you. Can you give a procedure for guessing whether you were shown the smaller or the larger of the two numbers, such that your guess is correct with probability *strictly* greater than 0.5 (anything larger than 0.5 is acceptable; exactly how much larger than 0.5 may depend on the actual values of the two numbers)?