Problem Set 9

1. *Marbles*
   Box A contains 2 black and 5 white marbles, and box B contains 1 black and 4 white marbles. A box is selected at random, and a marble is drawn at random from the selected box.
   - What is the probability that the marble is black?
   - Given that the marble is white, what is the probability that it came from the box A?

2. *Money bags*
   I have a bag containing either a \$1 or \$5 bill (with equal probability assigned to both possibilities). I then add a \$1 bill to the bag, so it now contains two bills. The bag is shaken, and you draw out a \$1 bill. If a second student draws the remaining bill from the bag, what is the chance that it is a \$1 bill? Show your work.

3. *Happy families*
   (a) Consider a collection of families, each of which has exactly two children. Each of the four possible combinations of boys and girls, bg, gb, bb, gg, occurs with the same frequency. A family is chosen uniformly at random, and we are told that it contains at least one boy. What is the (conditional) probability that the other child is a boy? Justify your answer with a precise calculation.
   (b) On the same probability space as in part (a), let A be the event that the chosen family has children of both sexes, and B the event that the family has at least one girl. Are the events A and B independent? Justify your answer carefully.
   (c) Consider now the probability space of families with *three* children, with each of the eight possible combinations of boys and girls equally likely. Define the events A and B as in part (b). Are these events independent? Again, justify your answer carefully.

4. *Triply-repeated ones*
   We say that a string of bits has k *triply-repeated ones* if there are k positions where three consecutive 1's appear in a row. For example, the string 011100111110 has four triply-repeated ones.

   What is the expected number of triply-repeated ones in a random n-bit string, when $n \geq 2$ and all n-bit strings are equally likely? Justify your answer.

5. *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 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.

6. *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 Carlos has drawn the first shot, and Bob second. What is Carlos's best strategy, and what is the chance that he comes out the eventual winner? What about Alice?

7. *Extra Credit*

   n points are uniformly and independently picked from the unit interval $[0, 1]$. This divides the interval into $n + 1$ subintervals. What is the expected length of the leftmost interval (i.e. the one containing 0)?

   Suppose we select one of the $n+1$ subintervals at random as follows: pick a point x in the unit interval $[0, 1]$ uniformly at random, and select the subinterval that contains x (think of this as pointing at the line at random and selecting the subinterval that you are pointing at). What is the expected length of the chosen subinterval?