CompSci 102 Discrete Mathematics for CS
Spring 2006 Forbes — HW 5

Due Wednesday, March 29

Please note the following on your assignment: (1) all people with whom you collaborated or contacted outside of the instructor and TA, (2) resources you used other than the book and class notes, and (3) how much time you spent working on the assignment.

1. (10 pts.) Book Problems

(a) 5.1.12

(b) 5.2.2

(c) 5.2.36

(d) 5.3.12

2. (10 pts.) Game

The Chevalier de Méré proposes two games to you.

I He bets that at least one 6 would appear during a total of four rolls of one die

II He bets that he would roll a double 6, once in 24 rolls of two dice

In each game, if he is successful, you pay him \$10, if he is unsuccessful, he pays you \$10. Should you take either bet? Which game is better for you?

3. (15 pts.) Multiple Choice Tests

Answer each of the following using the definition of expectation. Be sure to define all your random variables precisely.

(a) A teacher is setting a multiple choice test. Each question has five possible responses, exactly one of which is correct. One point is awarded for each correct answer, and $-b$ points for each incorrect answer. The teacher wants to ensure that a student who guesses a response at random on any given question will achieve an expected score of zero for that question. What value of b should the teacher choose?

(b) Suppose there are 100 questions in all, and some student randomly guesses the answer to every one of them. What is his expected score, using the same value of b as in part (a)?

(c) A less clueless student is able to identify two false responses on every question, but then has to guess at random among the remaining three. What is her expected score on the test, again with the same value of b?

4. (10 pts.) Chopping up DNA

In a certain biological experiment, a piece of DNA consisting of a linear sequence (or string) of 1000 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 999 possible cuts occurs independently and with probability $\frac{1}{250}$.

(a) 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 250 . Does the expected number of pieces increase, decrease, or stay the same? Justify your answer with a precise explanation.

5. (15 pts.) Martingales

Consider a fair game in a casino: on each play, you may stake any amount \$$s$; you win or lose with probability 1/2 each (all plays being independent); if you win you get your stake back plus \$$s$; if you lose you lose your stake.

(a) What is the expected number of plays before your first win (including the play on which you win)?

(b) The following gambling strategy, known as the "martingale," was popular in European casinos in the 18th century: on the first play, stake \$1; on the second play \$2; on the third play \$4; on the kth play \$$2^{k-1}$. Stop (and leave the casino!) when you first win.

Show that, if you follow the martingale strategy, and assuming you have unlimited funds available, you will leave the casino \$1 richer with probability 1. [Maybe this is why the strategy is banned in most modern casinos.]

(c) To discover the catch in this seemingly infallible strategy, let X be the random variable that measures your maximum loss before winning (i.e., the amount of money you have lost before the play on which you win). Show that $E(X) = \infty$. What does this imply about your ability to play the martingale strategy in practice?

(d) Colin and Diane enter the casino with \$10 and \$1,000,000 respectively. Both play the martingale strategy (leaving the casino either when they first win, or when they lack suficient funds to place the next bet as required by the strategy). What is the probability that Colin wins? What is the probability that Diane wins?

6. (20 pts.) Dice

Let E_k be the event "the outcome of rolling two fair dice is k" for $k = 2, 3, \ldots, 12$ (the "outcome" is the sum of the two numbers on the top faces of the two dice after rolling).

(a) What is the sample space S for the roll of two dice, and what is the cardinality of S?

(b) Find p_k, the probability of E_k, for $k = 2, 3, \ldots, 12$.

(c) Do you expect the values of p_k to add up to 1? If yes, state why and check that they do. If not, explain why not.

(d) The value of k associated with event E_k is a random variable with distribution p_k. Compute the expected value (or mean) m and variance σ^2 of this random variable.

(e) What is the conditional probability of the events E_k, defined earlier, given that the outcome of the first die is 1? Give the numerical value of the conditional probability for each k in $\{2, \ldots, 12\}$.

7. (20 pts.) Central Limit

You can calculate the probability that a random variable, X, with probability density function f is in a range (a, b) by integrating the density function over that range as in:

$$\Pr(a \leq X \leq b) = \int_a^b f(x)dx$$

For the following problems, you will need to calculate the area under a normal curve. You can use the definition of a normal distribution belwo and approximate the integral by writing a program.

The *Normal distribution* with mean μ and variance σ^2 is the distribution with density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-(x-\mu)^2/2\sigma^2}.$$

(a) Write a method *normalDistArea* that approximates the probability that a normally distributed variable is ina particular range,

```
public class ProbUtils {
  // returns the probability that a normal distributed with mean mu
  // and variance sigma^2 is between a and b
  public double normalDistArea(double mu, double sigma, double a, double b)
  {
  {
}
```

(b) What is the probability, that a normal distributed values is in the interval $(-1.5\sigma, 1.5\sigma)$?

(c) Let X be the number of times that a fair coin flipped 50 times, lands heads. Find the probability that $X = 25$ by using the normal approximation and then compare it to the exact solution.

(d) The lifetime of a hard drive is a random variable X with mean value 4 years with a standard deviation of 2 years. A hard drive is used until it fails at which point it is replaced by a new one. Assumming a stockpile of 10 such drives with independent lifetimes, approximate the probability that over 40 years of use can be obtained.