Problem Set 10

Due: December 9

Reading: Lecture notes for weeks 13 and 14.

Problem 1. MIT students sometimes delay laundry for a few days. Assume all random values described below are mutually independent.

(a) A ***busy*** student must complete 3 problem sets before doing laundry. Each problem set requires 1 day with probability $2/3$ and 2 days with probability $1/3$. Let B be the number of days a busy student delays laundry. What is $\mathrm{E}[B]$?

Example: If the first problem set requires 1 day and the second and third problem sets each require 2 days, then the student delays for $B = 5$ days.

(b) A ***relaxed*** student rolls a fair, 6-sided die in the morning. If he rolls a 1, then he does his laundry immediately (with zero days of delay). Otherwise, he delays for one day and repeats the experiment the following morning. Let R be the number of days a relaxed student delays laundry. What is $\mathrm{E}[R]$?

Example: If the student rolls a 2 the first morning, a 5 the second morning, and a 1 the third morning, then he delays for $R = 2$ days.

(c) Before doing laundry, an ***unlucky*** student must recover from illness for a number of days equal to the product of the numbers rolled on two fair, 6-sided dice. Let U be the expected number of days an unlucky student delays laundry. What is $\mathrm{E}[U]$?

Example: If the rolls are 5 and 3, then the student delays for $U = 15$ days.

(d) A student is ***busy*** with probability $1/2$, ***relaxed*** with probability $1/3$, and ***unlucky*** with probability $1/6$. Let D be the number of days the student delays laundry. What is $\mathrm{E}[D]$?

Problem 2. There are about 250,000,000 people in the United States who might use a phone. Assume that each person is on the phone during each minute mutually independently with probability $p = 0.01$.

(To keep the problem simple, we are putting aside the fact that people are on the phone more often at certain times of day and on certain days of the year.)

(a) What is the expected number of people on the phone at a given moment?

(b) Suppose that we construct a phone network whose capacity is a mere **one percent** above the expectation. Upper bound the probability that the network is overloaded in a given minute. (Use the approximation formula given in the notes. You may need to evaluate this expression in a clever way because of the size of numbers involved. For example, you could first evaluate the *logarithm* of the given expression.)

(c) What is the expected number of minutes (approximately) until the system is overloaded for the first time?

Problem 3. We are given a set of n distinct positive integers. We then determine the maximum of these numbers by the following procedure:

Randomly arrange the numbers in a sequence.

Let the "current maximum" initially be the first number in the sequence and the "current element" be the second element of the sequence. If the current element is greater than the current maximum, perform an "update": that is, change the current maximum to be the current element. Either way, change the current element to be the next element of the sequence. Repeat this process until there is no next element.

Prove that the expected number of updates is $\sim \ln n$.

Hint: Let M_i be the indicator variable for the event that the ith element of the sequence is bigger than all the previous elements in the sequence.

Problem 4. In a certain card game, each card has a point value.

- Numbered cards in the range 2 to 9 are worth five points each.
- The card numbered 10 and the face cards (jack, queen, king) are worth ten points each.
- Aces are worth fifteen points each.

(a) Suppose that you thoroughly shuffle a 52-card deck. What is the expected total point value of the three cards on the top of the deck after the shuffle?

(b) Suppose that you throw out all the red cards and shuffle the remaining 26-card, all-black deck. Now what is the expected total point value of the top three cards? (Note that drawing three aces, for example, is now impossible!)

Problem 5. A true story from World War II:

The army needs to identify soldiers with a disease called "klep". There is a way to test blood to determine whether it came from someone with klep. The straightforward approach is to test each soldier individually. This requires n tests, where n is the number of soldiers. A better approach is the following: group the soldiers into groups of k. Blend the blood samples of each group and apply the test once to each blended sample. If the group-blend doesn't have klep, we are done with that group after one test. If the group-blend fails the test, then someone in the group has klep, and we individually test all the soldiers in the group.

Assume each soldier has klep with probability, p, independently of all the other soldiers.

(a) What is the expected number of tests as a function of n, p, and k? (Assume for simplicity that n is divisible by k.)

(b) How should k be chosen to minimize the expected number of test performed, and what is the resulting expectation?

(c) What fraction of the work does the grouping method expect to save over the straightforward approach in a million-strong army where 1% have klep?

Problem 6. The hat-check staff has had a long day, and at the end of the party they decide to return people's hats at random. Suppose that n people have their hats returned at random. We previously showed that the expected number of people who get their own hat back is 1, irrespective of the total number of people. In this problem we will calculate the variance in the number of people who get their hat back.

Let $X_i = 1$ if the ith person gets his or her own hat back and 0 otherwise. Let $S_n ::= \sum_{i=1}^{n} X_i$, so S_n is the total number of people who get their own hat back. Show that

(a) $\mathrm{E}[X_i^2] = 1/n$.

(b) $\mathrm{E}[X_i X_j] = 1/n(n-1)$ for $i \neq j$.

(c) $\mathrm{E}\left[S_n^2\right] = 2$. *Hint:* Use (a) and (b).

(d) $\mathrm{Var}\left[S_n\right] = 1$.

(e) Explain why you cannot use the variance of sums formula to calculate $\mathrm{Var}\left[S_n\right]$.

(f) Using Chebyshev's Inequality, show that $\Pr\{S_n \geq 11\} \leq .01$ for any $n \geq 11$.

Problem 7. Let R_1 and R_2 be independent random variables, and f_1 and f_2 be any functions such that domain $(f_i) =$ codomain (R_i) for $i = 1, 2$. Prove that $f_1(R_1)$ and $f_2(R_2)$ are independent random variables.

Problem 8. Let A, B, C be events, and let I_A, I_B, I_C be the corresponding indicator variables. Prove that A, B, C are mutually independent iff the random variables I_A, I_B, I_C are mutually independent.

Student's Solutions to Problem Set 10

Your name:

Due date: December 9

Submission date:

Circle your TA: David Jelani Sayan Hanson

Collaboration statement: Circle one of the two choices and provide all pertinent info.

1. I worked alone and only with course materials.
2. I collaborated on this assignment with:

   got help from:[1]

   and referred to:[2]

DO NOT WRITE BELOW THIS LINE

Problem	Score
1	
2	
3	
4	
5	
6	
7	
8	
Total	

[1]People other than course staff.

[2]Give citations to texts and material other than the Fall '02 course materials.