CS 174, Fall 1998
Midterm 1
Professor A. Sinclair

Read these instructions carefully

1. This is a closed book exam. Caculators are permitted

2. This midterm consists of 10 questions. The first seven questions are multiple choice: the remaining three require written answers.

3. Answer the multiple choice questions by circling the correct answer. You should be able to answer all of these from memory, by inspection, or with a very small calculation. Incorrect answers may attract a negative score, so if you do not know the answer you should not guess.

4. Write you answers to the other questions in the spaces provided below them. None of these questions requires a long answer, so you should have enough space; if not, continue on the back of the page and state clearly that you have done so. Show all your working.

5. The questions vary in difficulty: if you get stuck on some part of a question, leave it and go on to the next one.

Problem 1. Two standard decks of 52 cards are randomly shuffled separately. The expected number of cards that are at the *same* position in both decks is

1/52 1/2 1
2 13 26

Problem 2. Three fair six-sided dice are rolled. *Given that* at least one of the dice comes up 6, the probability that *exactly* one of them comes up 6 is

25/216 75/216 125/216
25/91 75/91 1/3

Problem 3. Alice and Bob are two people in a group of size n. The group is ordered randomly in a line. The probability that there are exactly k people between Alice and Bob is

k/n (n-k)/(n*(n-1))
(n-k-1)/n! (n-k-1)/(n*(n-1)) 2(n-k-1)/(n*(n-1)

Problem 4. Each cereal box contains one coupon, chosen independently and uniformly at random from a set of n different coupons.

(a) The expected number of boxes that need to be bought before a copy of *some particular* coupon is obtained is

n^2 n*(ln (ln n))
n^(1/2) n n*(ln n) e^n

(b) The expected number of boxes that need to be bought before at least one copy of *all* n coupons is obtained is on the order of

n^2 n*(ln (ln n))
n^(1/2) n n*(ln n) e^n

Problem 5. 10*n* balls are thrown at random into *n* bins.

(a) The probability that the first bin contains exactly k balls i

(10n choose k)*((1/n)^k)*((n-1)/n)^(10n-k) (n choose k)*((1/n)^k)*((n-1)/n)^(10n-k)

10*(n choose k)*((1/n)^k)*((n-1)/n)^(10n-k)
((1/n)^k)*((n-1)/n)^(10n-k) (10n choose k)*((1/n)^10n)

(b) As n goes to infinity, for fixed k, this probability becomes very close to

1/10 (e^-1)*(1/k!)
(e^-10)*(1/k!) (e^-1)*(10^k)/k! (e^-10)*(10^k)/k!

Problem 6. X and Y are arbitrary independent random variables satisfying E[X] = E[Y] = 1 and Var[X] = Var[Y] = 2. Circle those three of the following statements that must be true:

Pr[X = 1] < 1 E[1/X] =
1 E[x^2] = 1

Pr[X >= 2] <= 1/2 E[XY] =
1 Var[2X+Y] = 10

Problem 7.

(a) A coin with heads probability *p* is tossed *n* times independently. The random variable X measures the number of heads observed *minus* the number of tails observed. The *expectation* of X is

0 np
n(p-1) n(2p-1) np(1-p)

(b) The *variance* of X is

np n(p^2)
(n^2)(p^2) np(1-p) 4np(1-p)

Problem 8. The Candy Bar Problem

A candy bar of total length n is made up of a sequency of n unit-length blocks. Assume that n is odd. Suppose you cut the bar at a randomly chosen boundary between two blocks. Let the r.v. X be the length of the <u>longer</u> of the two resulting pieces.

(a) Compute E[X] in the case n = 5.

(b) Compute E[X] exactly as a function of n. Check your answer against the value you obtained in part (a). [Note: Recall that the value of the sum (Σi=1 to m i) is (1/2)*m(m+1)]

Problem 9. Knowing All The Right People

Consider the following scenario. We have a large group of *n* people, some of whom know each other. We call a person *influential* if he/she knows at least n/100 other people (i.e., one person in 100). Since we would like to have access to all influential people, we'd like to find a small set *S* of people so that, for every influential person, there is someone in *S* who knows him/her. We'll call such a set S a *covering set*.

Here is a very simple randomized method for constructing a possible covering set S: for each of the n people independently, flip a coin with heads probability p; if the coin comes up heads, put that person in S.

(a) In the above method, what is E[|S|], the expected size of the set S?

(b) Let x be some particular influential person. For S constructed randomly as above, show that the probability that nobody in S knows X is at most (1-p)^(n/100).

np n(p^2) (n^2)(p^2) np(1-p) 4np(1-p)

(c) Show that. If we take p = (100*ln(3n))/n, then the probability in (b) is at most 1/(3n).

(d) Deduce that, with this value of p, the set S is a covering set with probability at least 2/3.

(e) Deduce from parts (a) and (d) that (again with this same value of p), with probability at least 1/6, the set S is a covering set of size at most 200ln(3n). [Hint: Let E1 be the event that S is not a covering set, and E2 the event that |S| > 200ln(3n). Part (d) bounds Pr[E1]. Use Markov's inequality and part (a) to bound Pr[E2]. Then combine your bounds on Pr[E1] and Pr[E2] to get a bound on Pr[E1 or E2].]

(f) Part (e) means that our simple randomized method, with the value of p given in part (c), will find a covering set of size at most 200ln(3n) with probability at least 1/6. How would you boost this probability to 1-ε for any desired ε>0.

Problem 10. To Hire or Not To Hire

A certain company has selected n candidates to interview for a vacant position. According to the local law, the candidates have to be interviewed in sequence, and each one has to be either offered the position or rejected immediately after the interview, before the next candidate is interviewed.

The company has decided to adopt the following strategy. Schedule the n candidates in a random order. For some value k (to be determined), interview and reject the first k candidates. After that, offer the position to the first candidate who is better qualified than all of the first k. (The position remains open if no such candidate exists.) The company assumes that no two candidates are equally qualified, so there are no ties. In this problem, we will investigate how good this strategy is in selecting the most qualified candidate.

(a) Suppose first that k = n-1, i.e., all candidates but one are interviewed and rejected and then the last one is hired if he/she is the best; otherwise the position remains open. Find the probability that the best candidate is hired in this special case.

(b) Now let's turn to the general case and consider the probability of hiring the best candidate as a function of n and k. Let us assume that the best candidate is scheduled as the mth one to be interviewed. Show that

Pr[best candidate is hired | best candidate is the mth interviewd] = { 0 if 1 <= m <= k; k/(m-1) if k+1 <= m <= n.

[Hint: Consider assigning the n candidates to the n interview slots, starting with the best candidate and ending with the worst. In this process, what event determines whether the best candidate is hired?]

(c) Use the result of the previous part to show that, as k, n approaches infinity,

Pr[best candidate is hired] ~ (k/n)*(ln n - ln k)

[Hint: Recall that (Σi=1 to n-1 (1/i)) ~ (ln n) as n approaches infinity]

(d) Using the fact that the function x*ln(1/x) is maximized when x=1/e, find the (approximate) value of k (as a function of n) that the company should use in order to maximize the probability of hiring the best candidate, and compute this probability for your value of k.