CSM 70 Discrete Mathematics and Probability Theory

Spring 2016

Final Review

I. Propositions

1. [True/False]: $\forall x\, (P(x) \wedge Q(x))$ is equivalent to $(\forall x, P(x)) \wedge (\forall x, Q(x))$. Explain

2. [True/False]: $(\forall n \in N)(P(n) \Rightarrow Q(n)) \Rightarrow (\neg \exists n \in N)(Q(n) \Rightarrow \neg P(n))$ for all P,Q? Explain.

3. Suppose $P(k) \Rightarrow P(k+2)$. For the following, label as always true, sometimes true, never true:

   a. $P(0) \Rightarrow \forall n\, P(n+1)$

   b. $\exists n\ \exists m>n$ s.t $[P(2n) \wedge \neg P(2m)]$

II. Stable Marriage

1. [True/False]: If w is the top choice of m and m is the top choice of w, then m and w must be paired with each other in any stable matching

2. [True/False]: Suppose that in an instance of the original Stable Marriage problem with n couples (so that every man ranks every woman and vice versa), there is a man M who is last on each woman's list and a woman W who is last on every man's list. If the Gale-Shapley algorithm is run on this instance, than M and W will be paired with each other.

3. [True/False]: Suppose we have an instance of the original Stable Marriage problem with M and W as above. In *any* stable solution to the instance, M and W will be paired with each other.

4. [True/False]: Consider an alternative to the Propose and Reject algorithm (with no rejections), where women take turns choosing the best available husband from the remaining unchosen men. On day 1, the oldest woman chooses her most preferred man, and marries him. On day k, the k-th eldest woman chooses her most preferred choice from the remaining unmarried men, and marries him. No matter what the preferences are, this process always results in a stable matching.

5. Given m men and m women, for any m≥2, what is the minimum number of stable pairings that must exist for any sets of preferences? Justify your answer with a specific example attaining the minimum

III. Mod Arithmetic

1. Find the inverse of 7 mod 48.

2. Let p,q be distinct prime numbers. Prove that $p^{q-1} + q^{p-1} = 1$ (mod pq)

IV. RSA

1. Fahad's Fault

Suppose, CSM is using RSA with modulus *n* and public exponent *e*. One day they are hacked, and their private key *d* becomes known to the attackers. Fahad, the overlord security consultant, suggests that instead of regenerating the new keys completely from the scratch, only the new exponents *e', d'* need to be re-computed, leaving the modulus *n* unchanged (after all, indeed modulus computation requires more work).
Is this safe? If yes, explain why. If not, show how the pirates can compromise the new system (i.e. compute new *d'* from *e, d, n, e'*).

Attacker HKN intercepted some packets $c_1, c_2, \ldots, c_k$, encrypted using RSA with public exponent *e* and modulus $n=pq$ (i.e. HKN knows only ciphertexts, *n, e,* but not *p, q*). His spies also learnt that one of the plaintext packets m_i (for some *i*, s.t. $c_i = m_i^e \bmod n$) is divisible by *p*.
Can HKN decipher all intercepted packets now? How or why not?

2. **We Will Make a Change**
   Nick and Nikhil use RSA cryptosystems with the same number n = pq, but different encryption exponents e_1 and e_2 , which are relatively prime. Alex sends the message m to both Nick and Nikhil using their individual encryption keys. Show that Erik can recover the message m from the two ciphertexts.

V. Polynomials and Error Correcting Codes

1. We would like to send a message of length n over a channel. You know that at most k packets can be dropped, and of the packets not dropped, j might be corrupted. How many packets do we need to send in order to guarantee a correct decoding?

2. We have a polynomial of degree 2 that goes through the points (1,0), (2,3), and (4,0) modulo 7. What is P(3)?

VI. Graphs

1. [True/False]: The hypercube graph always has an Eulerian tour

2. Let G be a non bipartite triangle-free simple graph with n vertices and minimum degree k. Let l be the minimum length of an odd cycle in G.
    a. Let C be a cycle of length l in G. Prove that every vertex not in V (C) has at most two neighbors in V (C).

    b. By counting the edges joining V (C) and V (G) − V (C) in two ways, prove that n $\geq$ kl/2 (and thus l $\leq$ 2n/k).

    c. When k is even, prove that the inequality of part (b) is best possible.

3. In a village there are three schools with n students in each of them. Every student from any of the schools is on speaking terms with at least n + 1 students from the other two schools. Show that we can find three students, no two from the same school, who are on speaking terms with each other

4. Let T1 and T2 be spanning trees of G with T1 != T2. Prove that there exists e that is in T1 and not T2 and f that is in T2 and not T1 so that both T1 −e+f and T2 −f +e are spanning trees.

VII. **Counting**

1. There are 9 representatives in the UC Berkeley Computer Science Conference: 2 from 61A, 3 from 61B, and 4 from 70. During the opening ceremony, 3 of the representatives fall asleep. In how many ways can exactly two of the sleepers be from the same class?

2. Corrina is choosing fruit from a tray. She only wants 6 pieces of fruit and she can choose from grapes, strawberries, and raspberries. There are at least 6 pieces of each of these fruits on the tray in front of her. How many different assortments of 6 fruits can be selected?

3. There are 2000 white balls in a box, and unlimited supply of white, green and red balls, initially outside the box. in each turn, we may replace: 2 white with 1 green, 2 reds with green, 2 greens with white and red, a white and green with a red, or a green and red with a white. if we end up with just three balls, prove at least one ball is green.

VIII. Halting

1. The function Neverloops(P) is 0 if program P does not halt on some input x, and 1 if P halts on every input x. Is there a program that computes Neverloops? Justify your answer.

2. [True/False]: The problem of determining whether a program halts in time 2^(n^2) hours on an input of size n is undecidable.

IX. Probability

Recall that in a secret sharing scheme the secret p(0) mod q can be reconstructed from the values of the polynomial p(x) of degree d at any d + 1 points. However, the values of the polynomial p(x) at any d points reveal absolutely no information about the secret p(0). This condition can be formally stated using conditional probability as follows:

Pr[p(0) = a | p(1), p(2), …, p(d)] = 1/q for every a mod q.

Now suppose Alice wishes to share a secret that consists of two numbers a and b, each mod q. She picks a random degree d polynomial p(x) mod q such that p(0) = a and p(1) = b. She distributes shares p(2), …, p(k) as with standard secret sharing (where k ≥ d + 2), and claims that any d + 1 people can reconstruct the secret, but any d people have absolutely no information about the secret.

a. Formally state (using conditional probability) Alice's claim that the values p(2), …, p(d + 1) reveal absolutely no information about the secret a, b.

b. Is Alice's claim correct? If so prove it, if not give a precise reason why not.

X. Countability

1. Give a bijection from the real number interval (1,oo) to the real number interval (0,1). What does it say about both of these sets?

2. Prove that the set of all programs if countably infinite.

XI. Variance/Expectation/Distributions

1. If a coin is biased, it will show up heads 80% of the time. There is a 20% chance that it is biased. What is the expected number of heads we should get when we flip this same coin 100 times.

2. You roll a die twice. What is the expectation of the minimum value between the two faces? The variance?

3. Know the Distribution
   a. Number of days until you begin studying for CS70 (you start studying with chance p)

   b. Number of days you'll waste procrastinating before your final (you procrastinate with chance p)

   c. Number of times that the teacher will give everyone a clarification during the exam (at any moment, can happen with chance p)

   d. Which room you take your final in, chance that you are taking your final in 1 Pimentel (out of 2 possible rooms) is p

XII. Markov and Chebyshev

1. Mark off Madoff

It’s 2008 and Bernie Madoff is selling pipe dreams. You’re smart some days and impulsive on others. 20% of the time, you’re smart and don’t invest that day. The rest of the time, you let temptation get the better of you and the amount you invest (lose) is distributed over a Poisson with variance $2 million.

After nearly a month (25 days), Madoff is arrested.

Bound the probability of losing more than a $100 million.

2. Fake it ‘til you make it

You’re pitching your latest app (Tinder for playdates for dogs) to a local VC.

As any CEO knows, the amount that the VC will invest depends on your level of confidence when pitching, which depends on how prepared you are. Say you prepare a week ahead with probability 20% and the night before the rest of the time, resulting in high and low confidence, respectively.

After surveying the data, you learn that high-confidence pitches average $4 million while high-confidence pitches net $1.5 million on average, both with standard deviation of $0.25 million, distributed normally.

Use Chebyshev’s to find a 98% confidence interval (2 sig figs).

XIII. Covariance

1. You flip coin c_1 twice and it lands heads with probability p_1= .5. Every time it lands heads, you flip another coin c_2 that lands heads with probability p_2=.5. What is the covariance between the number of times c_1 lands heads the number of times c_2 lands heads?

2. X = N(0,1), Y = N(0,1), Z = N(0,1)
   A = X + Y + Z
   B = fX +gY + hZ for some constants f, g, h
   Find Cov(A, B).

3. Continue 2: C = sA + tB, find Cov(A, C)

XIV. LLSE

1. **X**, **Y** independent

2. **Y** = 2**X** w.p. ½, 1 w.p. ½; **var(X)** = 8

3. **X** = die roll
   Y = payout: 3 * that die roll + 2 * a coin toss

4. Jean is practicing his prediction skills. The percentage of clouds Y, is a uniform r.v between 0 and 1. The temperature Z, is a uniform r.v between 0 and 1. The probability of rain X, is given by the average of Y and Z. Unfortunately Jean only has access to a machine that measures the value Y+2Z.

XV. Continuous Distributions

Which Distribution and which value?

1. Due to human impact on climate and habitats, fewer and fewer Roan antelope can be seen in South Africa’s Kruger National Park. If there are an average of 2 sightings of Roan antelope per day, what is the probability that there are 7 sightings in a week?

2. The average score on a recent exam was 70%. Only 10 students scored above 90%. How many students are expected to have scored below 50%?

3. On average, there are 6 independent attempts to pull a fire alarm during a single 61A exam that takes 3 hours. Assuming that the first attempt will always succeed, how far into the exam do we expect to go, until the fire alarm rings?

XVI. Markov Chains

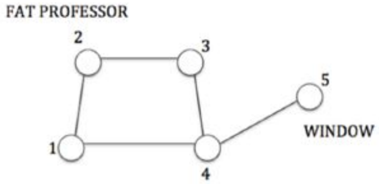

Figure 2: A fly wanders randomly on a graph.

1. Suppose that the fly wanders as follows: if it is at node i at time n, then it chooses one of its neighbors j of i uniformly at random, and then wanders to node j at time n + 1. For times n = 0, 1, 2, … let X_n be the fly's position at time n. Argue that $\{X_n, n \geq 0\}$ is a Markov Chain, and find the invariant distribution.

2. Now for the process in part (a), suppose that the (not-to-be-named) professor sits at node 2 reading a heavy book. The professor is very fat, so he/she doesn't move at all, but will drop the book on the fly if it reaches node 2 (killing it instantly). On the other hand, node 5 is a window that lets the fly escape. What is the probability that the fly escapes through the window supposing that it starts at node 1?

3. Now suppose that the fly wanders as follows: when it is at node i at time n, it chooses uniformly from all neighbors of node i except for the one that it just came from. For times n = 0, 1, 2, … let Y_n be the fly's position at time n. Is this new process $\{Y_n, n \geq 0\}$ a Markov chain? If it is, write down the probability transition matrix; if not, explain why it does not satisfy the definition of Markov chains.

XVII. Conditional Expectation

You're hoping to get a high score on the CS70 final, but not surprisingly, the number of questions you get wrong depends on how much you study. Say you study in 1 hour sessions, where the probability of the number of sessions, S, you study decreases linearly such that Pr[S] = 2/576(24-S).

The number of questions Q you get wrong is distributed over Binom(100, p), where p=½S.

Find E[Q|S=12].