CS 70 Mock Final

This final is definitely too long for the time given, like the real final probably will be. Questions from Fall 2017.

1 True/False

1. (True/False) If $d \mid mn$ then $d \mid n$ or $d \mid m$.

2. (True/False) $\neg P \implies Q$ implies that $\neg Q \implies \neg P$.

3. (True/False) If there are two different stable pairings in a stable marriage instance, then it cannot be the case that all men have the same preference list.

4. (True/False) If men ask women in reverse order of preference, we still get a stable pairing, but this time it is female optimal.

5. (True/False) The length of every cycle in the hypercube is even.

6. (True/False) If $cov(X, Y) = 0$, then X, Y are independent. If false, give a counterexample.

7. (True/False) If $X \sim \text{Geom}(p)$, then $E[X + m \mid X > n] = m + n + E[X]$.

8. (True/False) The CLT can be used to bound the probability that a random variable is far from its mean.

2 Short Answer: Discrete Math

1. $\neg(\forall x)P(x) \vee Q(X) \equiv \exists x,$ __________ (Fill in the blank.) Make sure the negation is fully distributed.

2. What is the size of $\{a(\text{mod } n), 2a(\text{mod } n), 3a(\text{mod } n), \ldots (n-1)a(\text{mod } n)\}$ if $\gcd(a,n) = 1$?

3. What is $2^{16}(\text{mod } 7)$?

4. What is the size of $\{ay(\text{mod } pq) \colon y \in 1, \ldots, pq-1\}$ when a is a multiple of p (but not q)?

5. What is $5^{60}(\text{mod } 77)$?

6. What is the minimum number of degree 1 vertices in an n-vertex tree for $n > 1$? (Answer could be an expression that involves n).

7. What is the maximum number of degree 1 vertices in an n-vertex tree? (Answer could be an expression that involves n).

8. Consider Professor Rao's public key (N, e) and secret key d, which has 512 bits. Professor Rao wants to share the secret key with his three children where any two can recover the secret d.

   (a) What degree polynomial should he use?

   (b) How big should the field over which we are working be? (That is, how big should the modulus be for the modular arithmetic that we use).

9. We have to assign 750 students to rooms in CS 70.

   (a) How many ways are there to do this in 3 rooms of capacity 240, 250, and 260?

   (b) How many ways are there to do this in 3 rooms of capacity 250, 260, and 270? (Notice there will be a total of 30 empty seats in this case.) (May involve summations)

3 A Probability Proof

You have n coins $C_1, C_2, \ldots, C_n$ for $n \in \mathbb{N}$. Each coin is weighted differently so that the probabilit ythat coin C_i comes up heads is $\frac{1}{2i+1}$. Prove by induction that if the n coins are tossed, then the probability of getting an odd number of heads is $\frac{n}{2n+1}$.

1. Base case.

2. State your inductive hypothesis.

3. Do the inductive step.

4 Short Answer: Probability

1. In a class of 24 students, what is the probability that at least two students have the same birthday? (Assume 365 days in a year. No need to simplify)

2. Two real numbers are chosen uniformly from the unit interval. What is the probability that their sum is less than or equal to 1 given that one of them is less than or equal to 1/2?

3. X, Y are independent continuous-valued random variables uniform in the interval $[0, 1]$. What is $E[X|X + Y = 1.5]$?

4. I want to take a student poll to find the popularity of CS 70 (assume each student independently likes it with probability p), and I need to pay each student \$1 to get their opinion. Suppose I want to estimate p within 1 percent accuracy with a 95% confidence level, I want to find how much money I need to find my estimate.

   (a) What estimator could you use for p from a set of samples, $X_1, X_2, \ldots, X_n$? (it should have expectation p).

   (b) What is an upper bound on the variance of your estimator that does not depend on p?

   (c) How much money would I need to spend if I use the CLT?

   (d) How about if I use the Chebyshev bound?

5. A relatively rare disease afflicts 1 in 100 people in the population. Screening for the disease has a missed detection rate of 1% (i.e. there is a 1% chance that a person has the disease but the test doesn't catch it), and a false alarm rate of 5% (i.e. there is a 5% chance that the test comes out positive for the disease even though the person does not have it), then if a test comes out positive, what is the probability that the person has the disease? (The answer can be an expression with numbers, no need to simplify into a single number.)

6. This problem is about collecting Marvel superheros.

   (a) You want to collect 20 Marvel superheroes, which come randomly in cereal boxes. How many cereal boxes do you expect to buy before you collect them all?

   (b) Now let's say you can only buy 20 cereal boxes. What's the expected number of superheroes you'll get?

   (c) What is the variance of the number of distinct superheros that you collect in part b)?

7. X and Y are continuous random variables and are uniformly distributed with pdf $f(x, y) = c$ over their region of support (the area where they have even a chance of taking on values). Their region of support is the following: $\{1 < X < 2, 1 < Y < 4\} \cup \{2 < X < 3, 2 < Y < 3\}$.

   (a) Find c. Are X and Y independent?

   (b) Find the marginal distributions of X and Y.

   (c) Find the MMSE estimate of Y given X.

8. There are N passengers boarding a full flight. They have assigned seats but they have all lost their boarding passes, so they choose to sit in random seats.

   (a) What is the expected number of passengers who sit in their assigned seats?

   (b) What is the probability that i passengers sit in their assigned seats?

9. Let $X \sim N(0, 1)$ and $Y \sim N(1, 1)$ be independent Gaussian random variables. You get an observation $z = 0.6$ that is equally likely to come from X and Y. You want to decide whether z came from X or Y by evaluating which decision leads to a larger probability of being right.

   (a) If you decide z came from X, what is the probability you are right?

   (b) Should you decide z is from X or from Y to get a larger probability of being right?

10. A hard-working GSI is holding her office hours (OH) for EECS 70 students. A random number of students enter and leave her office during her OH. Let us break up time into 1-minute segments, and assume that there is either 0 or 1 student in each time segment, and that this discrete arrival/departure process is well modeled by a 2-state Markov Chain. For each time segment, the transition probabilities are 0.8 for going from 0 students to 1 student in the OH, and 0.4 for going from 1 student to 0 students in the OH.

   (a) If the GSI starts off her office hours at $t = 0$ with 1 student, what is the probability that she has 0 student at time $t = 2$?

   (b) What does the probability go to, as t gets large, that there is 1 student?

5 Longer Probability Question

Alice (A) and Bob (B) play a one-on-one pickup game of basketball. Each made basket counts as 1 point. A beats B by a score of 51-49. We want to find the probability that A leads from start to finish (i.e. the game is never tied at any time other than at the start of the game) but we will lead up to this with some helpful hint questions that should help you find the answer.

Let us first define the following useful events:

- T_A: Game had at least one tie and A got the first point
- T_B: Game had at least one tie and B got the first point
- N: Game had no tie.

1. How are $P(T_A), P(T_B)$, and $P(N)$ related? (write an equation)

2. Given the final score, what is the probability that B got the first point?

3. How is $P(T_B)$ related to the probability that B got the first point?

4. How are $P(T_A)$ and $P(T_B)$ related? (Hint: use symmetry)

5. Using the above, what is the probability that A leads from start to finish (i.e. the game is never tied except at the start of the game)?