Mock CS70 Test

August 6, 2015

120 Points on test, with 27 as bonus points. You will have 120 minutes to complete the exam. Good Luck!

Problems

1. **Probability and Counting**

   (a) (6 points) Question: Alice and Bob have three fair coins. Alice flips the three coins, and Bob flips the three coins. What is the probability that Alice gets the same number of heads as Bob?

   (b) (6 points) I have a six-sided fair dice. You can pay me one dollar to roll it. Once it lands, you can either choose to keep the number on the dice and I give you the equivalent number of dollars, or pay one dollar to roll it again. What is your strategy for this game?

   (c) (8 points) In a baseball game, the teams play best out of 5 games. This means if a team has won 3 games already, it automatically wins. Assume team A and team B are playing and each have the same probability $\frac{1}{2}$ of winning. What is the probability that there will be a 5th game?

2. **Modular**

   (a) (8 points) Question: Prove or disprove that $x^2 + y^2 = 2015$ has no solution in integers x, y.

   (b) (10 points) Prove that for a prime p, if $a^k \equiv 1 \mod p$ and a is not 1 mod p, then k and $p-1$ are NOT relatively prime.

3. **Lots of induction**

(a) (5 points) Prove that for all pairs of positive reals a, b, we have that:

$$\frac{a+b}{2} \geq \sqrt{ab}$$

(b) (7 points) Prove by induction, for all positive integers n, that:

$$\frac{a_1 + a_2 + ... + a_{2^n}}{2^n} \geq (a_1 a_2 ... a_{2^n})^{\frac{1}{2^n}}$$

(c) (10 points) Prove by reverse induction for all positive integers k,

$$\frac{a_1 + a_2 + ... + a_k}{k} \geq (a_1 a_2 ... a_k)^{\frac{1}{k}}$$

4. **Geometry and Variance?!**

(a) (4 points) Pythagorean Theorem, Revisited
Recall from lecture that for random variables X, Y,

$$Var(X+Y) = Var(X) + Var(Y) + 2(E(XY) - E(X)E(Y))$$

where $E(XY) - E(X)E(Y)$ is the $cov(X,Y)$. Recall from your geometry class that

$$c^2 = a^2 + b^2 - 2ab\cos(\theta)$$

for the sides of a triangle. Do you notice a relation? Note that if X, Y are independent, then we have an analogous Pythagorean Theorem. Please write the relation here.

(b) (8 points) Note that the correlation, corr, is defined by $corr(X,Y) = \frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$. If I have variables A, B, C such that $corr(A,B) = 0.5$, $corr(B,C) = 0.5$, what is the maximum value of $corr(A,C)$?

(c) (10 points)Prove that $E(XY) \leq \sqrt{E(X^2)E(Y^2)}$ always, and show that equality holds when $X = c * Y$ for some constant c.

5. **Combinatorial Proof** (10 points) Prove that for all positive integers a, b, k, we have that:

$$\binom{a+b}{k} = \sum_{i=0}^{k} \binom{a}{i} \binom{b}{k-i}$$

6. **Nobody knows each other** (Bonus 15 points) In a group of n people, friendship is mutual. For a person v, denote $d(v)$ the number of people he is friends with in the group. Prove that we can find a set of at least x people such that nobody knows anybody else in the set, where

$$x = \sum_{v} \frac{1}{d(v)+1}$$

(Hint: Use probability and expected value!)

7. **Probability and Prime Numbers**

   (a) (4 points) I pick a random number from 1 to infinity. What's the probability that it is not divisible by 2, 3, and 5?

   (b) (4 points) If I have prime numbers $p_1, p_2, p_3, ...p_k$, what's the probability that my choice is not divisible by any of them?

   (c) (6 points) Prove, by induction, that all natural numbers greater than 1 can be expressed uniquely as a combination of prime factors.

   (d) (6 points) Prove that there are an infinite number of primes.

   (e) (Bonus 12 points) Recall that a geometric series $1 + x + x^2 + ... = \frac{1}{1-x}$ if $x < 1$. Prove that $\frac{1}{(1-\frac{1}{p_1})(1-\frac{1}{p_2})(1-\frac{1}{p_3})...} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + ...$

   (f) (8 points)Note that If we take all primes less than n to be $p_1, p_2, ...p_k$, then an approximation for $\frac{1}{(1-\frac{1}{p_1})(1-\frac{1}{p_2})...(1-\frac{1}{p_k})}$ is $\frac{1}{1} + \frac{1}{2} + ... + \frac{1}{n}$. What does this suggest about the approximate number of primes below n?

Solutions

1. **Probability and Counting**

   (a) Question: Alice and Bob have three fair coins. Alice flips the three coins, and Bob flips the three coins. What is the probability that Alice gets the same number of heads as Bob?

   Solution: We use casework. Since they flip the same number of coins, we sum the square of the probabilities.
   For no heads, this is $\binom{3}{0}\frac{1}{2^3} = \frac{1}{8}$.
   For one head, this is $\binom{3}{1}\frac{1}{2^3} = \frac{3}{8}$.
   For two heads, this is $\binom{3}{2}\frac{1}{2^3} = \frac{3}{8}$.
   For three heads, this is $\binom{3}{3}\frac{1}{2^3} = \frac{1}{8}$.
   Then the total probability is:

$$\frac{1}{8}^2 + \frac{3}{8}^2 + \frac{3}{8}^2 + \frac{1}{8}^2 = \frac{5}{16}$$

   (b) I have a six-sided fair dice. You can pay me one dollar to roll it. Once it lands, you can either choose to keep the number on the dice and I

give you the equivalent number of dollars, or pay one dollar to roll it again. What is your strategy for this game?
Solution: Note that the expected value of one roll is $3.5 - 1 = 2.5$. This means that I should keep rolling until I get a number above the expected value (i.e. $3, 4, 5, 6$.)

(c) In a baseball game, the teams play best out of 5 games. This means if a team has won 3 games already, it automatically wins. Assume team A and team B are playing and each have the same probability $\frac{1}{2}$ of winning. What is the probability that there will be a 5th game?
Solution: Our best option is to take the complement. We can treat winning and losing like flipping a coin. Then using casework on probabbilities: For 3 games only:

$$HHH, TTT = 2/8$$

For 4 games only:

$$8 - (HHHT, TTTH) = 6$$

gives $6/16$ probability. Then $1 - 2/8 - 6/16 = 3/8$.

2. **Modular**

(a) Question: Prove or disprove that $x^2 + y^2 = 2015$ has no solution in integers x, y.
Solution: The key here is to use modular arithmetic. Note that $2015 \equiv 3(\mod 4)$. However, we can see that by case work,
If $x \equiv 0 \mod 4$, then $x^2 \equiv 0 \mod 4$.
If $x \equiv 1 \mod 4$, then $x^2 \equiv 1 \mod 4$.
If $x \equiv 2 \mod 4$, then $x^2 \equiv 0 \mod 4$.
If $x \equiv 3 \mod 4$, then $x^2 \equiv 1 \mod 4$.
Since the values of x^2 and y^2 can therefore only be 0 or 1 in modulo 4, we cannot obtain 3.

(b) Prove that for a prime p, if $a^k \equiv 1 \mod p$ and a is not $1 \mod p$, then k and $p - 1$ are NOT relatively prime.
Solution: We will prove that if $a^k \equiv 1 \mod p$, then $a^{gcd(k,p-1)} \equiv 1 \mod p$ as well, which will solve the problem.
Proof: Note that by Fermat, $a^{p-1} \equiv 1 \mod p$, as well as $a^k 1 \mod p$. If $a^x \equiv 1 \mod p$ and $a^y \equiv 1 \mod p$, where $x > y$, then because all residues except 0 have inverses in prime p, we have that $a^{x-y} = a^x * \frac{1}{a^y} \equiv 1 \mod p$. Thus, now $x = x - y$ and y satisfy the condition, and we may keep applying this operation, which is the same as the gcd algorithm. Thus, let $x = k$, $y = p - 1$, and we are done.

3. **Lots of induction**

(a) Prove that for all pairs of positive reals a, b, we have that:

$$\frac{a+b}{2} \geq \sqrt{ab}$$

Solution: We have initially, $(\sqrt{a} - \sqrt{b})^2 \geq 0$. Then multiplying out and expanding, we see that $a + b - 2\sqrt{ab} \geq 0$, and the rest follows.

(b) Prove by induction, for all positive integers n, that:

$$\frac{a_1 + a_2 + ... + a_{2^n}}{2^n} \geq (a_1 a_2 ... a_{2^n})^{\frac{1}{2^n}}$$

Solution: Since it was proven above when $n = 1$, this is the base case. Now assume it holds for n. To prove it holds for $n + 1$, note that

$$x = \frac{a_1 + a_2 + ... + a_{2^n}}{2^n} \geq (a_1 a_2 ... a_{2^n})^{\frac{1}{2^n}}$$

and

$$y = \frac{a_{2^n+1} + a_{2^n+2} + ... + a_{2^(n+1)}}{2^(n)} \geq (a_{2^n+1} a_{2^n+2} ... a_{2^(n+1)})^{\frac{1}{2^n}}$$

Then using the above,

$$\frac{x+y}{2} \geq \sqrt{xy}$$

, and replacing terms, we prove it is valid for $n + 1$.

(c) Prove by reverse induction for all positive integers k,

$$\frac{a_1 + a_2 + ... + a_k}{k} \geq (a_1 a_2 ... a_k)^{\frac{1}{k}}$$

Solution: Since we've proven it at points $k = 2^1, 2^2, ...2^n, ...$, we can choose an arbitrarily large base then proceed downwards. Assume the inequality holds for k. Then to prove the case at k-1, Let

$$a_k = \frac{a_1 + a_2 + ... + a_{k-1}}{k-1}$$

. Then,

$$\frac{a_1 + a_2 + ... + a_k}{k} = \frac{a_1 + ... + a_{k-1}}{k-1}$$

But because of the RHS,

$$\frac{a_1 + ... + a_{k-1}}{k-1} \geq (a_1 a_2 ... a_{k-1} \frac{a_1 + a_2 + ... + a_{k-1}}{k-1})^{\frac{1}{k}}$$

Now rearranging the $\frac{a_1+a_2+...+a_{k-1}}{k-1}$ to the LHS, we get the desired result.

4. **Geometry and Variance?!**

   (a) Pythagorean Theorem, Revisited
   Recall from lecture that for random variables X, Y,

   $$Var(X+Y) = Var(X) + Var(Y) + 2(E(XY) - E(X)E(Y))$$

   where $E(XY) - E(X)E(Y)$ is the $cov(X,Y)$. Recall from your geometry class that

   $$c^2 = a^2 + b^2 - 2ab\cos(\theta)$$

   for the sides of a triangle. Do you notice a relation? Note that if X, Y are independent, then we have an analogous Pythagorean Theorem. Please write the relation here.
   Solution: $\cos(\theta) = -\frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

   (b) (8 points) Note that the correlation, corr, is defined by $corr(X,Y) = \frac{cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$. If I have variables A, B, C such that $corr(A,B) = 0.5$, $corr(B,C) = 0.5$, what is the maximum value of $corr(A,C)$?
   Solution: The solution is motivated by the fact that correlations are like cosines. Thus the maximum would be $-\cos(2\theta) = 1 - 2cos(\theta)^2$, where $cos(\theta) = 0.5$. We can prove this algebraically by using the variance formula. and using inequalities.

   (c) Prove that $E(XY) \leq \sqrt{E(X^2)E(Y^2)}0$ always, and show that equality holds when $X = c * Y$ for some constant c.
   Solution: If the student knows linear algebra, than he can prove by considering the vector space with random variables as vectors. Let the inner product be E(XY) for (X,Y); it satisfies all the axioms of a inner product, which means Cauchy-Swarchz holds.
   An elementary proof would be by considering the variable $Z = X - cY$. Then

   $$0 \leq E(Z^2) = E(X^2) + c^2 * E(Y^2) - 2c * E(XY)$$

   Solving for c in order for the RHS to look like what we need, we put $c = E(XY)/E(Y^2)$, and the algebra works to what we wanted.

5. **Combinatorial Proof** Prove that for all positive integers a, b, k, we have that:

   $$\binom{a+b}{k} = \sum_{i=0}^{k} \binom{a}{i}\binom{b}{k-i}$$

   Solution: We want to pick k people among $a+b$ people for a committee, where a are in Room 1 and b are in Room 2. Then the LHS clearly shows the simple choosing. However, we may also pick i people from Room 1, and $k-i$ people from Room 2, where i ranges from 0 to k. Thus the RHS shows this, and we have proven the statement.

6. **Nobody knows each other** In a group of n people, friendship is mutual. For a person v, denote $d(v)$ the number of people he is friends with in the group. Prove that we can find a set of at least x people such that nobody knows anybody else in the set, where

$$x = \sum_v \frac{1}{d(v)+1}$$

(Hint: Use probability and expected value!)
Solution: Randomly select a permutation of the people. Now apply this algorithm from left to right: Pick a person into the set if he precedes his neighbors in the ordering. Then nobody will know each other in this set. The probability of this happening is $\frac{1}{d(v)+1}$, and thus the expected value of the number of people in this group over all people is the sum over all such probabilities. Thus by the law of expected value, I can find a set with at least x members.

7. **Probability and Prime Numbers**

   (a) I pick a random number from 1 to infinity. What's the probability that it is not divisible by 2, 3, and 5?
   Solution: $(1-1/2)(1-1/3)(1-1/5)$

   (b) If I have prime numbers $p_1, p_2, p_3, ...p_k$, what's the probability that my choice is not divisible by any of them?
   Solution: Just take the product of the probability that my choice is not divisible by each prime. $(1-1/p_1)(1-1/p_2)...(1-1/p_k)$

   (c) Prove, by induction, that all natural numbers greater than 1 can be expressed uniquely as a combination of prime factors.
   Solution: Base case is simple for 2. Now assume the hypothesis is true for all numbers less than or equal to n. Then for $n+1$, it is either a prime, or a composite number of the form ab, where $a, b \leq n$. Using the inductive hypothesis, since a b can be uniquely represented, multiplying them will still make them uniquely represented, and thus we are done.

   (d) Prove that there are an infinite number of primes.
   Solution: Assume otherwise that we have a finite number of primes $p_1, ...p_k$. Then take the number $n = p_1 p_2 ... p_k + 1$. Clearly it is not divisible by any prime in our set. If it is prime itself, we have a contradiction. If it is composite, this contradicts the result above.

   (e) Recall that a geometric series $1 + x + x^2 + ... = \frac{1}{1-x}$ if $x < 1$. Prove that $\frac{1}{(1-\frac{1}{p_1})(1-\frac{1}{p_2})(1-\frac{1}{p_3})...} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + ...$
   Solution: Note that the LHS is equal to:

$$\frac{1}{1+\frac{1}{p_1}+\frac{1}{p_1^2}+...}\frac{1}{1+\frac{1}{p_2}+\frac{1}{p_2^2}+...}...$$

But since from the above results, every integer can be represented uniquely as a product of prime factors, then after multiplying out the product and collecting terms, each term maps to a distinct $\frac{1}{n}$ for each n, hence we are done.

(f) Note that If we take all primes less than n to be $p_1, p_2, ...p_k$, then an approximation for $\frac{1}{(1-\frac{1}{p_1})(1-\frac{1}{p_2})...(1-\frac{1}{p_k})}$ is $\frac{1}{1}+\frac{1}{2}+...+\frac{1}{n}$. What does this suggest about the approximate number of primes below n?
Solution Note that $\frac{k}{n}$ is approximately the probability of choosing a prime from the range $(1, n)$. But this is also equal to the recipricoal of the LHS, as it is also approximately equal to the probability the selected number is not divisible by any prime below n. Since the harmonic series can be approximated using an integral to get $\ln n$. Then we have $\frac{n}{\ln n}$ primes approximately below n.