EECS 70 Discrete Mathematics and Probability Theory
Spring 2014 Anant Sahai Final Review

1. True/False

Circle the right answer. No justification is needed.

T **F** $\forall x\,(P(x) \Rightarrow \exists y\,Q(y)) \equiv \neg\exists x\,(P(x) \wedge \forall y\,\neg Q(y))$

T **F** $\exists i, \forall j, P(i,j) \implies \forall i, \exists j, \neg P(i,j)$.

T **F** Let $P(x) =$"x is prime" and $Q(x) =$"x is even". It is true that: $\neg\exists x\,(P(x) \wedge Q(x) \Rightarrow x = 2)$.

T **F** For P, Q as above the following is true: $\forall x\,(P(x) \wedge Q(x) \Rightarrow x = 2)$

T **F** 6 has a multiplicative inverse modulo 15.

T **F** The efficient implementation of RSA hinges upon our ability to efficiently check whether a number is prime or not.

T **F** Toby and his 4 friends go to a horror movie and sit together in five consecutive seats. Toby will not sit in the middle seat. The number of ways the 5 friends can be arranged in the 5 seats is 96.

T **F** For any two disjoint events A, B, with $\Pr[B] \neq 0$, it holds that $\Pr[A|B] = 0$.

T **F** For any set of n i.i.d. random variables it holds that $E[X_1 \cdot X_2 \ldots \cdot X_n] = E[X_1]^n$.

T **F** The union bound $\Pr[\bigcup_{i=1}^n A_i] \leq \sum_{i=1}^n \Pr[A_i]$ holds for all events A_i, disjoint or not.

T **F** The number of cereal boxes we have to buy before collecting all n different baseball cards follows the Geometric distribution with parameter $1/n$.

T **F** If you have a set of 11 points, where 7 of them agree with a degree 4 polynomial p_1, and 9 of them agree with a degree 4 polynomial p_2, then p_1 and p_2 must be the same polynomial.

T **F** The set of reals is countable.

T **F** The set of all subsets of size 10 of the integers is countable.

T **F** The set of all subsets of the integers is countable.

T **F** The set of all finite subsets of the natural numbers is uncountable.

T **F** For any events A and B, if $\mathbf{P}[A] \neq 0$, $\mathbf{P}[B] \neq 0$, and A and B disjoint, then A and B are dependent.

T **F** For any two events A and B, if $\mathbf{P}[A] \neq 0$, $\mathbf{P}[B] \neq 0$, and $\mathbf{P}[A|B] = 1$, then $\mathbf{P}[B|A] = 1$.

T **F** For any two events, if $\mathbf{P}[B] \neq 0$ and $\mathbf{P}[\overline{B}] \neq 0$, then $P[A|B] + P[A|\overline{B}] = 1$.

T **F** For any three events, A, B, C, if $\mathbf{P}[A] \neq 0$, $\mathbf{P}[B] \neq 0$, and A and B are independent, then A and B are conditionally independent on C.

2. What Number? (Fall 2004)

Short answer. Justification is not needed.

1. The number of different functions $f : \{1...n\} \to \{1, \ldots, m\}$.

2. The number of ways for k nonnegative integers (zero is nonnegative) to add up to n.

3. The number of ways to pick n fruits from k varieties of fruits where one picks at least $l < k$ different fruits.

4. The probability that bin 1 is empty when m balls are thrown into n bins.

5. The probability that bin 1 has exactly k balls in it.

6. The number of solutions to the equation $2x + 7 = 0 \mod 41$?

7. The number of values in the set $\{x : \exists a,\ x = 5a \mod 41\}$.

8. The number of values in the set $\{x : \exists a,\ x = 2a \mod 15\}$

9. $3^{26} \mod 15$?

10. Least upper bound on the number of zeros of a degree d polynomial over a field.

11. Least upper bound on the number of intersections of two distinct degree d polynomials.

12. The number of polynomial on a field with p elements of degree d.

13. The number of polynomials of degree $d > 4$ on a field with p points that go through 4 points.

14. The expected number of pairs of people with the same birthday in a room with 30 people in it. (No need to simplify.)

3. Induction

1. Let r be a real number, $r \neq 1$ and $r \neq 0$. Prove by induction that for any $k \in \mathbb{N}$,

$$r^0 + r^1 + \cdots + r^k = \frac{1 - r^{k+1}}{1 - r}.$$

(Summer 2011)

2. Prove that if $x + 1/x$ is an integer, then $\forall n \geq 1,\ x^n + 1/x^n$ is an integer. (Spring 2003)

3. You're a parking attendant. The lot has k consecutive parking spots in a straight line. There are three types of vehicles: VW Bugs, which take up a single parking spot, stretch Limos, which take up two consecutive spots, and SUVs, which also take up two consecutive spots. Parking in Berkeley being as it is, you have an inexhaustible supply of each kind of vehicle.

Let z_k denote the number of different ways to park vehicles in a lot with k parking spots without leaving any empty spots. For instance, $z_3 = 5$, because the valid ways to fill the lot are VVV, VL, VS, LV, and SV (notation: $V =$ VW bug, $L =$ limo, and $S =$ SUV).
Prove that $z_k = (2^{k+1} + (-1)^k)/3$ for all k. (Fall 2001)

4. The so-called Russian multiplication for multiplying two positive integers a, b is defined recursively as follows:

Russian Multiplication

```
algorithm mult(a,b)
  if a = 1 then return(b)
  else return(mult(⌊a/2⌋,2b) + b × (a mod 2))
```

Prove using (strong) induction on the first input a that the algorithm correctly outputs the product ab. (Spring 2000)

4. More short answers! (Fall 2009, Summer 2010)

1. Solve for x in the equation $7x = 43 \mod 48$.

2. What is $2^{147} \mod 21$?

3. Solve for x : $7x + 1 \equiv 4 \mod 23$. Show ALL work.

4. Compute $1^{13} + 2^{13} + \ldots + 12^{13} \mod 13$. (Hint: there is an easier way than repeated squaring!)

5. Let X_1 and X_2 be random variables both binomially distributed with parameters n and p.
   (a) Suppose X_1 and X_2 are independent. Give the mean, variance and distribution of $X_1 + X_2$.

   (b) Suppose $X_1 = X_2$. Give the mean, variance and distribution of $X_1 + X_2$.

6. Consider an instance of a stable-marriage problem on n men and n women. Suppose that each woman's (man's) preference list is a random permutation of the men (or women) and that a pairing is produced according to the traditional propose and reject algorithm.
   (a) What is the probability that *every* man gets his top choice?

   (b) What is the probability that *every* women gets her bottom choice?

7. Consider the following instance of the Stable Marriage problem, in which the men are $\{1,2,3,4\}$ and the women are $\{A,B,C,D\}$.

Men (1-4)					Women (A-D)				
1:	B	D	A	C	A:	2	1	4	3
2:	D	B	C	A	B:	4	3	1	2
3:	B	C	A	D	C:	1	4	3	2
4:	D	A	C	B	D:	2	1	4	3

Run the Traditional Propose & Reject algorithm to compute a stable pairing. Show your work.

8. Let X and Y be the outcomes of two independent rolls of a 6-sided dice. Let $Z \equiv X + Y \bmod 6$. Are the random variables X and Z independent? Justify your answer.

5. Polynomials (Fall 2007, Spring 2003)

1. Let p be a prime. How many distinct polynomials of degree at most 2 are there over $GF(p)$? Explain your answer.

2. How many distinct polynomials of degree at most 2 over $GF(p)$ pass through the point $(0,0)$? Explain your answer.

3. List all polynomials of degree exactly 2 over $GF(5)$ that pass through the points $(0,1)$ and $(1,4)$. Show your work.

4. You are given three points, $(0,2),(1,7)$ and $(2,1)$. What is the probability that a degree 5 polynomial in GF_{11}, with coefficients chosen uniformly at random, passes through these three points?

5. Given the same three points, what is the probability that a degree 1 polynomial in GF_{11}, with coefficients chosen uniformly at random, passes through these three points?

6. Secret sharing (Spring 2009)

Suppose we wish to share a secret among five people, and we decide to work modulo 11. We construct a degree two polynomial $p(x) = ax^2 + bx + s$ by picking the coefficients a and b at random (mod 11); the constant term is the secret s, also a number mod 11. We give shares $p(1), ..., p(5)$ to each of the five people.

1. Suppose that three of the people arrange a meeting and share the information that $p(1) = 9, p(2) = 5$, and $p(4) = 1$. What is the secret s? Use Lagrange interpolation and show your work.

2. Suppose now that the first of these three people wants to discover the secret herself but deceive the others into thinking that the value of the secret is larger by 1 than the true value s. Explain clearly how she can do this; assuming that the other two are honest. (Hint: You should not need to repeat the entire calculation from the earlier question)

7. Event Modulus (Spring 2006)

Answer each of the following questions with a short explanation (one/two sentences should suffice)

1. Let X be uniformly distributed on $0, 1, ..., 100$. What is $P[2X \equiv 21 \bmod 101]$?

2. Let X be uniformly distributed on $0, 1, ..., 99$. What is $P[2X \equiv 21 \bmod 100]$?

3. Let X be uniformly distributed on $0, 1, ..., 36$. What is $P[X^2 \equiv 36 \bmod 37]$?

8. Combinatorial Story, Again... (Fall 2010)

Prove the following identity:
$\binom{n}{r} \cdot \binom{r}{k} = \binom{n}{k} \cdot \binom{n-k}{r-k}$

1. By telling us a story (combinatorial proof)

2. Algebraically

9. Monty is bored (Spring 2008)

Tired of hosting the same game year after year, Monty Hall decided to make some changes to his game. There are still three doors, but now one contains 1000 dollars, one contains 500 dollars, and one contains nothing (0 dollars), with the order of the prizes randomly permuted. The contestant first selects a door. Then she has the choice of paying X dollars for Monty to open, among the two unchosen doors, the one that contains the smaller amount of money. If the contestant paid Monty, she then has the choice of switching to the other unopened door.

1. Suppose the contestant refuses to pay Monty. In this case, what is the expected value of her prize?

2. Suppose the contestant decides to pay, and then Monty opens a door that contains 500. Given this, what is the expected value of her prize if she switches? What is the expected value of her prize if she sticks with her original door?

3. Now for a different scenario: Suppose that the contestant pays, and then Monty opens a door that contains 0. Given this, what is the expected value of her prize if she switches? What is the expected value of her prize if she sticks with her original door?

4. Now suppose a second contestant, Bob, decides in advance that he will always pay and always switch to the unopened door (no matter what he sees behind the door that Monty opens). What is the overall expected value of his prize, with this strategy?

10. Anagrams (Spring 2006)

Recall that an anagram of a word is a string made up from the letters of that word, in any order. (For instance, there are exactly three anagrams of EYE: namely, EEY, EYE, and YEE. Note that anagrams need not form legal words in any language.)

How many different anagrams of SUCCESSFULLY are there given that the two L's are not next to each other?

11. Packets transmission

Suppose Alice encodes 40 packets of data into 100 packets to transmit to Bob, using the polynomial-based error correction code discussed in class.

Suppose that each transmitted packet is dropped independently with probability 0.5.

1. What is the exact probability that Bob can recover the data? (You do not have to evaluate the expression.)

2. Compute a positive lower bound on the probability that Bob will be able to recover the data.

3. Using the Central-Limit Theorem, give an approximation to the probability that Bob can recover the data. How do you compare this approximation to the lower bound you derived in part (b) ?

4. Suppose now Alice wants to send 400 data packets to Bob instead of 40. To guarantee the same probability of recovery of the data, do you think she has to encode into (i) less than, (ii) more than, or (iii) equal to 1000 packets to transmit to Bob? No calculations are required but briefly explain your reasoning.

12. Start-up (Summer 2010)

After completing EECS70, you decide to found a start-up and you're trying to pick a developer out of n candidates that applied for the job. To do that you want to interview the developers; the problem is that immediately after each interview you must decide whether to accept or reject the candidate; rejected candidates cannot be accepted later on, and once you accept a candidate the other candidates are rejected. If you reject the first $n-1$ candidates that show up then you must accept the n-th candidate without an interview.

1. What is the probability of hiring the best candidate if you just randomly pick one of the candidates?

2. Suppose $n = 3$. You come up with the following clever idea: you will interview the first candidate and reject her. Then you interview the second candidate and compare her to the first one; if she is better than the first candidate you hire her, otherwise you hire the third candidate. Show that this strategy has a probability of 1/2 of hiring the best applicant. Is that an improvement over the arbitrary choice? (Hint: think of the problem as a sequence of choices)

   The natural generalization of this idea for n developers is the following: you interview the first k applicants and don't hire anyone. You then hire the first applicant after (and including) the $k+1$-th, that is better than the best of the first k applicants.

3. What is the probability that the best applicant is in position i?

4. What is the probability that the strategy does not select any applicant in positions $k+1$ through $i-1$, where $i \geq k+1$?

5. Are the two events described in 3 and 4 independent? Give a brief explanation.

6. What is the probability of picking the best candidate with the above strategy? What is the optimal choice for k? (You may find the following approximation useful: $\sum_{i=a}^{b-1} \frac{1}{i} \approx \ln \frac{b}{a}$)

13. Beat Stanford! (Spring 2003)

The Cal Bears are playing the Stanford Cardinal in a 2-out-of-3 series, i.e. they play games until one team wins a total of two games. The probability that the Bears win the first game is $\frac{1}{2}$. For subsequent games, the probability of winning depends on the outcome of the preceding game: the team is energized by victory and demoralized by defeat. If the Bears win a game, then they have a $\frac{2}{3}$ chance of winning the next game. On the other hand, if the Bears lose, they have only a $\frac{1}{3}$ chance of winning the next game.

1. What is the probability that the Bears win the 2-out-of-3 series given that they win the first game?

2. What is the probability that the Bears won the first game, given that they won the series?

14. Stirling intensifies (Rozanov, "Probability Theory: A Concise Course")

A full deck of card is divided in half at random. Use Stirling's approximation to estimate the probability that each half contains the same number of red and black cards.

15. Sampling (Spring 2003)

1. Suppose we randomly sample k elements out of $\{1,2,\ldots,n\}$ without replacement. What is the probability that the sequence of elements chosen is strictly increasing?

2. Now suppose that we sample *with* replacement. Now what is the probability that the sequence of elements is strictly increasing?

16. Bad Medicine (Summer 2011)

A hospital receives $\frac{1}{5}$ of its flu vaccine shipments from Pallagen Laboratories and the remainder of its shipments from other companies. Each shipment contains a very large number of vaccine vials. In the shipments from Pallagen Laboratories, 50% of the vials are ineffective, while in all other shipments, 25% of the vials are ineffective. Suppose that the the hospital tests 3 randomly selected vials from an unknown shipment and finds that exactly 1 of the 3 is ineffective. What is the probability that this shipment came from Pallagen Laboratories?

(Hint: You will find it much easier to solve this problem using fractions than with decimals.)

17. Can't trust the machines (Fall 2007)

An unreliable machine consists of four components connected in parallel. The machine fails if and only if *all four* components fail. Component failures are independent events, and occur each time the machine is switched on with probabilities $0.5, 0.5, 0.4$ and 0.2 respectively.

1. What is the probability that the machine functions correctly when it is switched on?

2. Suppose you turn on the machine and it functions correctly. What is the probability that *none* of the four components has failed?

(c) You switch the machine on and off repeatedly until it fails. What is the expected number of times you switch it on?

18. Balls and bins redux (Spring 2003)

Suppose n indistinguishable balls are thrown into n bins independently and uniformly at random.

1. What is the probability of a particular bin containing *exactly* one ball?

2. Compute the expected number of bins with exactly one ball.
   (Hint: Linearity of Expectation).

19. The evolution of a social network (Fall 2011)

(We give a simplified analysis of the connectivity of a social network.)

Say one person in a class of n people knows a secret, perhaps how many problems there are on the final. Occasionally a randomly chosen person A **who doesn't know the secret** calls a randomly chosen person B ($B \neq A$) and learns the secret if B knows it.

Let X_2 be a random variable that represents the number of calls (no two calls are simultaneous) until two people know the secret.

1. What is the distribution of X_2?

2. What is $E[X_2]$?

3. Let X_i be the number of calls needed to go from $i-1$ people knowing the secret to i people. What is $E[X_i]$?

4. What is the expected time for everyone to know the secret?

20. More expectation (Spring 2003, Fall 2006)

1. Define the weight of a subset $S \subseteq \{1,2,\ldots,n\}$ to be the sum of all the elements of S; i.e. $\sum_{x \in S} x$. For example, if the subset is $\{2,5,9,15\}$, then the weight is 31.
   Suppose we pick the subset $S \subseteq \{1,2,\ldots,n\}$ uniformly at random (i.e. each of the 2^n subsets is equally likely). What is the expected weight of the chosen subset?

2. Given a sequence of n numbers $[a_1, a_2, \ldots, a_n]$, we define an *inversion* to be a pair (a_i, a_j) such that $i < j$ but $a_i > a_j$. For example, in the sequence $[5,3,7,2]$ there are three inversions: $(5,3),(5,2),(7,2)$. A sorted array has zero inversions, and the number of inversions is one measure of how "close-to-sorted" the sequence is.
   What is the expected number of inversions in a random sequence of n numbers? Assume that the sequence consists of a n distinct numbers that are randomly permuted.

(You will learn more about the inversion counting problem and how to solve it using divide-and-conquer in EECS170.)

21. Candy Bar (Spring 2000)

A candy bar of total length L is made up of a linear sequence of n equal-length blocks. Assume that n is odd. Suppose you cut the bar at one of the $L-1$ boundaries between two blocks chosen uniformly at random. Let the random variable X be the length of the longer of the two resulting pieces.

1. Compute $E[X]$ in the case $n = 5$.

2. Compute $E[X]$ as a function of n. Check your answer against the value you computed in the earlier part. (Hint: You may want to use the fact that the sum of the first m integers is $\frac{m(m+1)}{2}$.)

3. Compute the variance $Var(X)$ as a function of n. (Hint: You may want to use the fact that the sum of squares of the first m positive integers is $\frac{m(m+1)(2m+1)}{6}$.)

4. Use Chebyshev's Inequality together with the two above parts to derive an upper bound on the probability that the longer piece has length at least $\frac{7L}{8}$.

22. Crime Lord (Fall 2009)

You are a top secret government operative seeking an organized crime lord. You know that he is either hiding in a warehouse or in a power plant. Every day you can only search one location - if you look in the warehouse and the crime lord is actually there, you will find him with probability 0.5. (depending on whether you are lucky that day). If you look in the power plant and the crime lord is actually there, you will find him with only probability 0.3. Prior to your search, you received a tip that the crime lord is in the power plant with probability 0.6.

1. On the first day, should you search the warehouse or the power plant? Justify your answer with a calculation.

2. Suppose you search the warehouse in the first day but cannot find the crime lord. Given this information, would you now search the warehouse or would you switch to the power plant on the second day? Justify your answer with a calculation. (You can assume that the crime lord does not change locations from day to day.)

23. Redo the virtual labs... just kidding! (Spring 2010)

Each flip of a biased coin comes up Heads with probability p and Tails with probability $1-p$, independently of all other flips. For any integer $n \geq 1$, let X_n be the number of coin flips until the n^{th} Heads appears. We denote $\mu_n = E[X_n]$ and $\sigma_n^2 = Var(X_n)$.

1. What is $\mu_1 = E[X_1]$?

2. What is $\mu_n = E[X_n]$?

3. What is $\sigma_n^2 = Var(X_n)$? You may leave your answer in terms of $\sigma_1^2 = Var(X_1)$

4. For $t > 0$, use Chebyshev's Inequality to bound the probability that X_n is outside the range $[\mu_n - t\sigma_n, \mu_n + t\sigma_n]$.

5. For $t > 0$, use the Central Limit Theorem to obtain an approximate bound on the probability that X_n is outside the range $[\mu_n - t\sigma_n, \mu_n + t\sigma_n]$.

24. LoLN (Fall 2010)

The *Law of Large Numbers* is said to hold for a sequence of random variables $S_1, S_2, S_3, S_4, \ldots$ if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr[\,\left|\tfrac{1}{n}S_n - \mathbb{E}(\tfrac{1}{n}S_n)\right| > \varepsilon] = 0.$$

In the course we have shown that the Law of Large Numbers holds if $S_n = X_1 + \cdots + X_n$, where the X_i's are i.i.d. random variables. This problem explores if the Law of Large Numbers holds under other circumstances.

Packets are sent from a source to a destination node over the Internet. Each packet is sent on a certain route. Each route has a failure probability of p and different routes fail independently. If a route fails, all packets sent along that route are lost. You can assume that the routing protocol has no knowledge of which route fails.

For each of the following routing protocols, determine whether the Law of Large Numbers holds when S_n is defined as the total number of received packets out of n packets sent. Circle YES if the Law of Large Number holds, or NO if not. (Whenever convenient, you can assume below that n is even.)

1. YES or NO: Each packet is sent on a completely different route.

2. YES or NO: The packets are split into $n/2$ pairs of packets. Each pair is sent together on its own route (i.e., different pairs are sent on different routes).

3. YES or NO: The packets are split into 2 groups of $n/2$ packets. All the packets in each group are sent on the same route, and the two groups are sent on different routes.

4. YES or NO: All the packets are sent on one route.

25. Countability

1. Prove that the real numbers in the interval $[0, 1]$ are uncountable.

2. Recall that a real number which is not a rational number is called irrational. Are the irrationals countably infinite or uncountably infinite? Prove your answer carefully. (You may use facts we proved in class about the cardinality of the reals and the rationals.)