CS 70 Discrete Mathematics and Probability Theory
Spring 2015 Vazirani Final Exam Solution

PRINT your name: Felix, Yt
(last) (first)

SIGN your name: _____

PRINT your student ID: 3033784 0 x x

CIRCLE your exam room: Wheeler Auditorium 220 Hearst Gym

Name of the person sitting to your left: _____

Name of the person sitting to your right: _____

Please write your name and student ID on every page (**after** the exam starts).

Please write your answers in the spaces provided in the test. There are extra pages at the end that you can use as scratch papers. We will not grade them unless you clearly tell us in the problem's main page to look there.

You may consult three single-sided sheets of handwritten notes. Apart from that, you may not look at books, notes, etc. Calculators and computers are not permitted.

You have 180 minutes. There are 11 questions provided, but you need to do only 9 for a total of 170 points.

- Choose three out of questions 4, 5, 6, and 7.
- Choose three out of questions 8, 9, 10, and 11.

$\frac{162}{170} = 95.3\%$

Clearly mark which three questions you want us to grade (you might even cross out the question you don't want us to grade). Please note that there is no advantage to attempting all four questions — if you do not indicate which three questions you want us to grade, then we will only grade the first three questions.

You may use the number of points as a rough guide for the amount of time to allocate to that question. Note that many of the points are for proofs and justifications for your answers. Please make sure you spend the time to write clear, correct and concise justifications. Also make sure you read the questions carefully. Good luck!

Do not turn this page until your instructor tells you to do so.

Name: _____ SID: _____

-4

1. **True or False (30 points)**

For each question below, circle True if the statement is true, and circle False if the statement is false. No justification needed. 2 points per question; no partial credit.

(a) A proposition and its contrapositive cannot both be true.

Circle one: **True** **False**

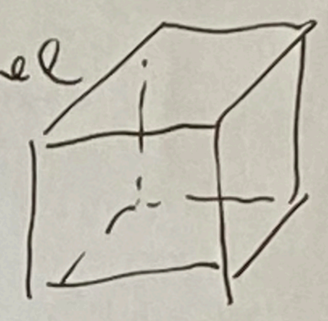
-1 -1

(b) The proposition $(A \wedge B) \vee (\neg A \wedge B) \vee \neg B$ can never be false.

Circle one: **True** **False**

A	B
T	T
F	T

even degree



(c) The hypercube graph always has an Eulerian tour.

Circle one: **True** **False**

? I don't think so.

(d) If $f: A \to B$ is an injective (1-1) function, then there exists a surjective (onto) function $g: B \to A$.

Circle one: **True** **False**

A → 1
B → 2
C → 3
D → 4

gcd(42, 3) = 3

(e) If $\gcd(a,b) = d$, then a has no factor larger than d.

Circle one: **True** **False**

gcd ≠ largest factor.

$n = (p-1)(q-1)$ $13\overline{)91}$ 7 $12 \times 6 = 72$

(f) In RSA with modulus $n = 91$ and encryption power $e = 5$, the decryption power is $d = 73$ because $de = 365 \equiv 1 \bmod 91$.

Circle one: **True** **False**

(g) If the multiplicative inverse $a^{-1} \bmod p$ exists for all $a \in \{1, \ldots, p-1\}$, then p is a prime.

Circle one: **True** **False**

all $\{1 - p-1\}$ coprime.

(h) For any $d \in \mathbb{N}$, the set of polynomials of degree d with integer coefficients is countable.

Circle one: **True** **False**

false.

→ A.

= 73

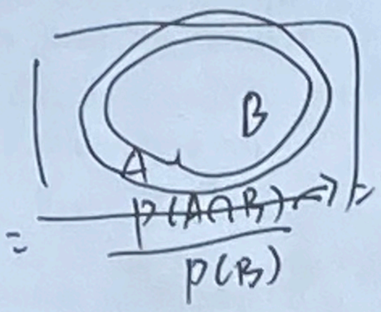
1 2 3 4 5 6

Name: _____ SID: _____

one event → one sample point.
once at a time.

(i) If a sample space Ω has n sample points, then there are 2^n possible events.
Circle one: **True** **False**

(j) If $P(A \mid B) = 1$, then $P(B) \le P(A)$.
Circle one: **True** **False**



$P(A|B) = \frac{P(A \cap B)}{P(B)}$ $\geq P(B)$

(k) For any random variable X, $\mathbb{E}(X^2) \ge \mathbb{E}(X)^2$.
Circle one: **True** **False**

$Var[X] \geq 0$.

(l) For any random variables X and Y, $\mathbb{E}(X - Y) = \mathbb{E}(X) - \mathbb{E}(Y)$.
Circle one: **True** **False**

$E[X]$ $E[\frac{1}{Y}]$

(m) If X and Y are independent random variables, then $\mathbb{E}(X/Y) = \mathbb{E}(X)/\mathbb{E}(Y)$.
Circle one: **True** **False**

$E[\frac{1}{Y}] = p_1 \frac{1}{y_1} + \cdots p_n \frac{1}{y_n}$

$E[Y] = p_1 y_1 + p_2 y_2 \cdots p_n y_n$

(n) If X and Y are independent random variables, then $\mathrm{Var}(X) = \mathrm{Var}(X - Y) - \mathrm{Var}(Y)$.
Circle one: **True** **False**

$E[(X-Y) - (\mu_x - \mu_y)]^2$

(o) If X is an exponential random variable, then $P(X \ge s+t \mid X \ge s) = P(X \ge s+t)$.
Circle one: **True** **False**

100%

$\neq 100\%$.

$E[(X-Y)^2] - E[(X-Y)]^2$

$= E[X^2] - 2E[XY] + E[Y^2] - (E[X] + E[Y])^2$

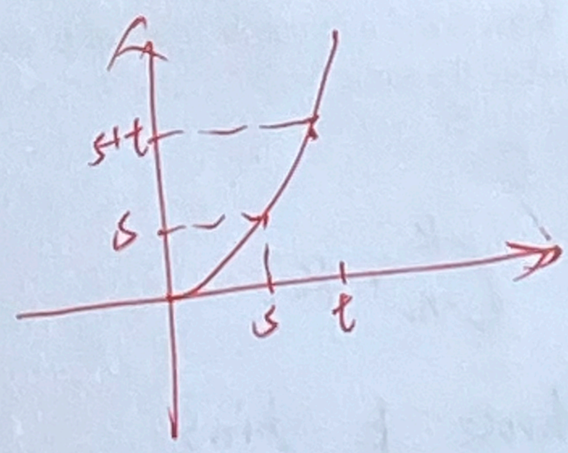
$= (E[X^2] - E[X]^2) + (E[Y^2] - E[Y]^2) + -2(E[XY] + 2E[X]E[Y])$

$P(X \geq s+t \mid X \geq s)$

$= \frac{P(X \geq s+t \cap X \geq s)}{P(X \geq s)}$

$X \geq s$

$\int_{s+t}^{\infty} e^x dx$ $\int_{s}^{\infty} e^x dx$



Name: _____ SID: _____

2. **Short answers (40 points)**

(a) What distribution would best model each of the following scenarios? Choose from binomial, Poisson, geometric, exponential, uniform, or normal distribution. No justification needed.

i. Number of taxis passing the corner of Euclid Ave and Hearst Ave between 5 pm and 6 pm on a weekday.

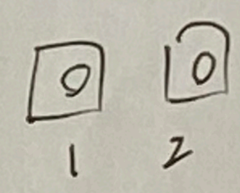
Poisson.

ii. Number of customers who purchase a lottery ticket before someone hits the jackpot.

Geometric

iii. Number of balls in the first urn in a Polya urn process, with two urns each starting with one ball.

uniform



iv. Times of finishers in the NY Marathon.

Normal

v. Number of girls in a family with 6 kids.

Binomial

vi. Number of miles a car can run before the engine fails.

Exponential $P[x>t] = \int_t^\infty \lambda e^{-\lambda x} dx = e^{-\lambda t}$

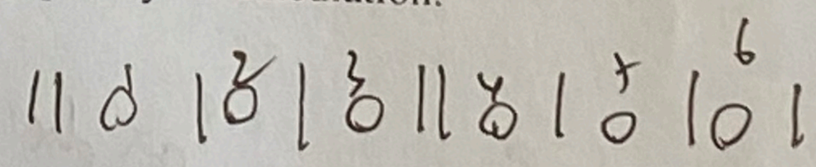
(b) What is the number of poker (5 card) hands with 2 pairs? Explain your calculation.
(A poker hand has 5 cards. The 2 pairs must be of different ranks, and the last card must also be different. So (2♢, 2♣, 4♡, 4♠, 8♢) is an example of a poker hand with 2 pairs, but (2♢, 2♣, 2♡, 2♠, 8♢) and (2♢, 2♣, 4♡, 4♠, 4♢) are not. The ordering of the cards does not matter.)

$C_{13}^2 C_4^2 C_4^2 C_{11}^1 C_4^1$

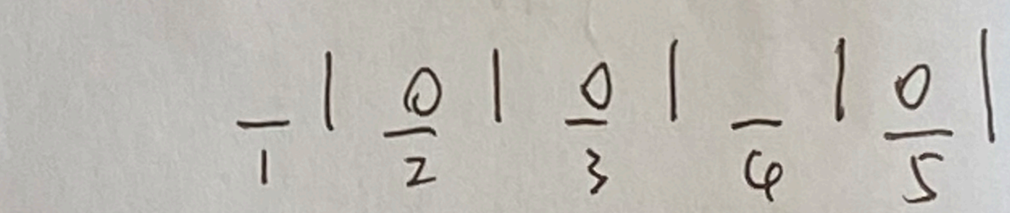
$\overline{2!}$ No need here

(c) What is the number of ways of placing k labelled balls in n labelled bins such that no two balls are in the same bin? Assume $k \le n$. Explain your calculation.

$C_n^k \cdot k!$



n choose k bins to hold k balls. all not in the same bin. then these k balls have order



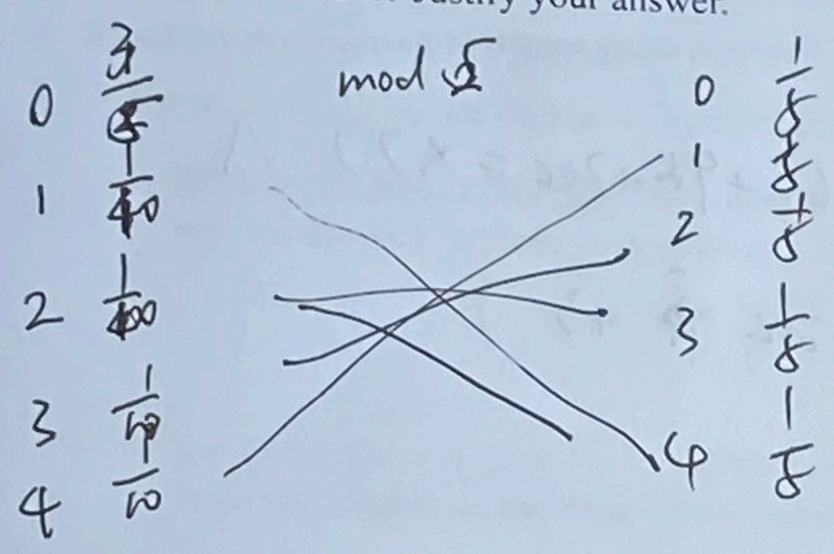
Name: _____ SID: _____

-4

But why.

Y. 0 1 . 2 . 3 . 4 . . . n-1 mod n

? (d) X and Y are independent random variables modulo n. You don't know the distribution of X, but you know that Y is uniformly distributed. What can you say about the distribution of $Z = X + Y \bmod n$? Justify your answer.



0. $\frac{3}{25} + \frac{1}{50} \times 4 = \frac{10}{50}$

1. $\frac{2}{50} + \frac{1}{50} + \frac{1}{50} + \frac{1}{50} = \frac{10}{50}$

2. $\frac{3}{25} + \frac{1}{50} + \frac{1}{50} = \frac{8}{50}$

3. $\frac{3}{25} + \frac{1}{50} + \frac{1}{50} + \frac{1}{50} = \frac{9}{50}$

4. $\frac{3}{25} + \frac{1}{50} + \frac{1}{50} + \frac{1}{50} + \frac{1}{50} = \frac{10}{50}$

Z is also uniformly distributed

(e) X and Y are independent random variables with normal distribution with mean m_1 and m_2 respectively, and variance σ_1^2 and σ_2^2 respectively. Describe the distribution of $Z = X + Y$ (including mean and variance).

$Z \sim N(\mu_x + \mu_y, \sigma_x^2 + \sigma_y^2)$

$p(z) = \frac{1}{\sqrt{2\pi\sigma_z^2}} e^{-\frac{(x-\mu_z)^2}{2\sigma_z^2}}$

Name: $\forall x \geq 44, \exists a.b.c$ SID: _____

6. Chicken McNugget (10 points) $\forall . a . b . c . \quad 6a+9b+20c \neq$

McDonald's sells chicken McNuggets only in 6, 9 and 20 piece packages. This means that you cannot purchase exactly 8 pieces, but can purchase 15. The Chicken McNugget Theorem theorem states that the largest number of pieces you *cannot* purchase is 43 (i.e., you cannot purchase exactly 43 pieces, and 43 is the largest number that you cannot purchase).

Formally state the Chicken McNugget Theorem using quantifiers.

$(\forall x > 43 . (\forall a, b . c \in \mathbb{N}, 6a+9b+20c = x)) \wedge$

$(\nexists a, b, c \in \mathbb{N}, 6a+9b+20c = 43)$

Name: _____ SID: _____

Instruction: Answer any three of the next four questions (questions 4, 5, 6, 7). Clearly indicate which three questions you want us to grade. If you do all four questions, we will only grade the first three.

4. **Random proposition (15 points)**

Suppose $x_1, x_2, \ldots, x_k$ are chosen independently and uniformly at random from $\{\text{True}, \text{False}\}$.

(a) What is the probability that the proposition $Q_1 = x_2 \wedge x_3 \wedge \cdots \wedge x_k$ is true? (Note that Q_1 does not involve x_1.) Explain your answer.

$P = \frac{1}{2^{k-1}}$

All need to be true for Q_1 to be true

$k-1$ terms

(b) Let $Q = Q_1 \vee Q_2 \vee \cdots \vee Q_k$ where $Q_i = x_1 \wedge \cdots \wedge x_{i-1} \wedge x_{i+1} \wedge \cdots \wedge x_k$. (Note that Q_i does not involve x_i, but does involve all other $k-1$ variables.) Prove that $\Pr[Q \text{ is true}] \leq k/2^{k-1}$.

$P[Q \text{ is True}] = \cup_i P[Q_i \text{ is True}]$

Applying the union bound.

$P[Q \text{ True}] = \cup_i P[Q_i \text{ true}]$

$\leq \sum_{i=1}^{k} P[Q_i \text{ true}]$

$= \frac{k}{2^{k-1}}$

Name: _____ SID: _____

5. **Neverloops (15 points)**

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.

(Note: the standard halting problem, which is uncomputable, asks on input P, x whether program P halts on input x.)

$$\text{Neverloops}(P) = \begin{cases} 0 & \text{if } P \text{ does not halt on some input } x \\ 1 & \text{if } P \text{ halts on every input } x. \end{cases}$$

No.

Name: _____ SID: _____

Name: _____ SID: _____

6. **Coin induction (15 points)**

We have n coins $C_1, \ldots, C_n$. The coins are weighted such that coin C_i comes up Heads with probability $\frac{1}{2i+1}$. Prove by induction that if the n coins are tossed independently, the probability of getting an odd number of Heads is $\frac{n}{2n+1}$.

Base case $n = 1$.

$$P(C_1 = H) = \frac{1}{2 \cdot 1 + 1} = \frac{1}{3} = \frac{n}{2n+1} = \frac{1}{3}. \checkmark$$

Inductive Hypothesis:

~~$n = k$~~

~~$P(C_i = H) = \frac{1}{2i+1}$~~

suppose $n = n-1$. the hypothesis holds

then ~~$P[\#H_{n-1}]$~~ $P[\#H_{n-1} \text{ is odd}] = \frac{n-1}{2n-1}$

Inductive step

$n = n$:

odd # heads $\begin{cases} \text{odd \# heads in Coin from 1 to } (n-1) + \text{Tail in } C_n \\ \text{even \# heads in Coin from 1 to } (n-1) + H \text{ in } C_n. \end{cases}$

Then ~~$P[\#H_n]$~~ is

$$P[\# H_n \text{ is odd}] = \frac{n-1}{2n-1} \times \left(1 - \frac{1}{2n+1}\right) + \left(1 - \frac{n-1}{2n-1}\right) \times \frac{1}{2n+1}$$

$$= \frac{n(2n-1)}{(2n-1)(2n+1)}$$

$$= \frac{n}{2n+1}. \checkmark$$

Name: 10 10 SID: _____

7. **Drawers of socks (15 points)**

A chest of drawers has two drawers. 10 different pairs of socks are randomly placed in the two drawers (each of the 20 socks is equally likely to be placed in either drawer).

1. 2. 3. 4. 5. ... 20

(a) Let N be the number of complete pairs of socks in the first drawer. Find the distribution of N. Specify the parameter(s).

Binomial Distribution.

$\frac{1}{2}$ P[1 pair of socks fall in drawer 1] $= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

P[k pairs of socks in drawer 1] $= C_{10}^{k} (\frac{1}{4})^{k} \cdot (1-\frac{1}{4})^{10-k}$

(b) What is the probability that at least one drawer has no complete pairs of socks? Explain your calculation.

[At least 1 drawer has no complete pairs of socks] = P[0 pairs of socks in drawer 1] + P[0 pairs in drawer 2] − P[no complete pairs in drawer 1 and 2]

$= C_{10}^{0} (\frac{1}{4})^{0} (\frac{3}{4})^{10} + C_{10}^{0} (\frac{1}{4})^{0} \cdot (\frac{3}{4})^{10} - (\frac{1}{2})^{10}$

$= 0.1126 - 0.000977$

≈ 0.1116

Name: _____ SID: _____

Instruction: Answer any three of the next four questions (questions 8, 9, 10, 11). Clearly indicate which three questions you want us to grade. If you do all four, we will only grade the first three.

8. **Base disease (15 points)**

$P[D] = 0.01$

Suppose that 1 percent of the population has a certain disease. There is a test for the disease, but it's not always correct.

$P[P|D] = 0.9 \quad P[\bar{P}|D] = 0.1$

- For a randomly chosen person who has the disease, the test comes back positive with probability 0.9 and negative with probability 0.1.
- For a randomly chosen person who doesn't have the disease, the test comes back positive with probability 0.01 and negative with probability 0.99.

$P[P|\bar{D}] = 0.01 \cdot P[\bar{P}|\bar{D}] = 0.99$

(a) The test on a random person comes back positive. What is the probability that the person has the disease?

$$P[D|P] = \frac{P[P \cap D]}{P[P]} = \frac{P[P|D]\,P[D]}{P[P \cap D] + P[P \cap \bar{D}]}$$

$$= \frac{0.9 \times 0.01}{0.9 \times 0.01 + 0.01 \times 0.99}$$

$$\approx 0.4762$$

✓

Name: _____ SID: _____

(b) Suppose that each test is itself probabilistic — if you perform the test twice on the same person with the disease, each time it comes up positive *independently* with probability 0.9. Similarly, if you perform the test twice on the same person who doesn't have the disease, each time it comes up positive independently with probability 0.01.

You choose a person at random and run the test twice on that person. Suppose the first test comes back positive. What is the probability that the second one comes back positive too?

Hint: Use your answer from part (a).

$$P[P_2 \mid P_1] = \frac{P[P_1 \cap P_2]}{P[P_1]}$$

$$= \frac{P[(P_1 \cap P_2) \cap D] + P[(P_1 \cap P_2) \cap \bar{D}]}{P[P_1]}$$

$$= \frac{0.9 \times 0.9 \times 0.01 + 0.01 \times 0.01 \times 0.99}{0.9 \times 0.01 + 0.01 \times 0.99}$$

$$= \frac{0.0081 + 0.000099}{0.0189} \approx 0.43.$$

Note $P[(P_1 \cap P_2) \cap D] = P[(P_1 \cap P_2) \mid D] \cdot P[D]$

$$= 0.9^2 \times 0.01$$

$$P[P_2 \mid P_1] = \frac{10}{21} \times 0.9 + \frac{11}{21} \times 0.01 \approx 0.43$$

Name: _____ SID: _____

9. **Secret sharing (15 points)**

Recall that in a secret sharing scheme the secret $p(0) \bmod q$ 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 $p(0)$. As we saw in lecture, this condition can be formally stated using conditional probability as follows: $\Pr[p(0) = a \mid p(1), p(2), \ldots, p(d)] = 1/q$ for every $a \bmod q$.

Now suppose Alice wishes to share a secret that consists of two numbers a and b, each $\bmod q$. She picks a random degree d polynomial $p(x) \bmod q$ such that $p(0) = a$ and $p(1) = b$. She distributes shares $p(2), \ldots, p(k)$ as with standard secret sharing (where $k \geq d+2$), 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 $p(2), p(3), \ldots, p(d+1)$ reveal absolutely no information about the secret a, b.

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

Name: _____ SID: _____

Name: _____ SID: _____

10. **Umbrella store (15 points)**

Bob has a store that sells umbrellas. The number of umbrellas that Bob sells on a rainy day is a random variable Y with mean 25 and standard deviation $\sqrt{105}$. But if it is a clear day, Bob doesn't sell any umbrellas at all. The weather forecast for tomorrow says it will rain with probability $\frac{1}{5}$. Let Z be the number of umbrellas that Bob sells tomorrow.

(a) Let X be an indicator random variable that it will rain tomorrow. Write Z in terms of X and Y.

$X = \begin{cases} 0 & \text{not rain} \\ 1 & \text{rain} \end{cases}$ $\quad Z = X \cdot Y$.

(b) What is the mean and standard deviation of Z?

$Z = \begin{cases} Y & p = \frac{1}{5} \\ 0 & p = \frac{4}{5} \end{cases}$

$E[Z] = E[XY] = \frac{1}{5} \cdot E[1 \cdot Y] + \frac{4}{5} \cdot E[0 \cdot Y]$

$= 5 + 0$

$= 5$

(c) Use Chebyshev's inequality to bound the probability that Bob sells at least 25 umbrellas tomorrow.

$Var[Z] = E[Z^2] - E[Z]^2$

$= E[X^2 Y^2] - E[Z]^2$

$= \frac{1}{5} E[Y^2] + \frac{4}{5} \cdot 0 - 25$

$= \frac{1}{5}(Var[Y] + E[Y]^2) - 25$

$= 121$

$P[Z \geq 25] = P[Z - E[Z] \geq 25 - E[Z]$

$= P[Z - 5 \geq 20]$

$\leq P[|Z - 5| \geq 20] \leq \frac{Var[Z]}{20^2} = \frac{121}{400} = 0.3025$

Name: _____

SID: _____

11. **To infinity and beyond (15 points)**

You are the captain of the Bimillennial Eagle, a spaceship that has just returned from hyperspace to ordinary space, only to encounter the debris of a recently destroyed planet. Your maneuvering jets are temporarily out of order. The expected number of pieces of debris in any km^3 of space is $1/10^6$. You reckon that your spaceship has a cross section of area $1/1000$ km^2, and you must travel 10^5 km before you are clear of the debris.

Model the debris field by a Poisson distribution and calculate your chances of getting all the way through the debris field without a collision.

Total travel volume $= 10^{-3} \times 10^5 = 10^2$ km^3.

$$P[X=i] = \frac{\lambda^i}{i!} e^{-\lambda}$$

$\lambda = E[\#\text{collision in travel volume}] = \frac{1}{10^6} \times 10^2 = 10^{-4}$.

$$P[X=0] = \frac{(10^{-4})^0}{0!} e^{-10^{-4}}$$

$\approx 99.999\%$

25