CS 70 Discrete Mathematics and Probability Theory
Fall 2023 Tal and Rao

Final Solutions

PRINT Your Name: Oski Bear

SIGN Your Name: *OSKI*

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

1. **Pledge.**

Berkeley Honor Code: As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others.

In particular, I acknowledge that:

- I alone am taking this exam. Other than with the instructor and GSIs, I will not have any verbal, written, or electronic communication about the exam with anyone else while I am taking the exam or while others are taking the exam.
- I will not have any other browsers open while taking the exam.
- I will not refer to any books, notes, or online sources of information while taking the exam, other than what the instructor has allowed.
- I will not take screenshots, photos, or otherwise make copies of exam questions to share with others.
- I won't work for people from Stanford. [This is optional.]

Signed: ______________________________

2. **Propositions. Basics.**

1. Let P, Q and R be propositions.
   (a) $P \wedge \neg P \equiv$ True.
   Answer: False. Both P and $\neg P$ can't be true, so the entire statement is false.
   (b) Consider the statement: $(P \implies R) \wedge (P \implies \neg R) \equiv$ ______. (Fill in the blank to make it true.)
   Answer: $\neg P$.
   (c) $(P \wedge Q \implies \neg R) \equiv (\neg R \vee \neg P \vee \neg Q)$.
   Answer: True. $(P \wedge Q \implies R) \equiv (\neg R \vee \neg(P \wedge Q)) \equiv (\neg R \vee \neg P \vee \neg Q)$
2. Let $P(x,y)$, $Q(x,y)$ be predicates on x and y.
   (a) $\neg(\forall x)(\exists y)(Q(x,y)) \equiv (\exists x)(\forall y)(\neg Q(x,y))$.
   Answer: True. This is De Morgan's over quantifiers.
   (b) If $(\forall x)(\exists y)(Q(x,y)) \implies (\forall x)(\forall y)(P(x,y))$, and $P(a,b)$ is false for some a and b, then $(\exists x)$ such that $(\forall y)$, $Q(x,y)$ is ______.
   Answer: False. This uses the contraposition which states the left hand expression must be false, which means there is some x, where for every y, $Q(x,y)$ is false.

3. **Quick Proofs!**

Prove or give a counterexample.

1. (5 points) For $a,b,c \in \mathbb{N}$, if $a \mid b$ and $b \mid c$, then $a \mid c$.
   Answer: $b = ka$ and $c = k'b$ and therefore $c = kk'a$ which implies $a \mid c$.
2. (7 points) Show that if the sum of digits of $n \in \mathbb{N}$ is divisible by 9, then n is divisible by 9.
   Answer:

We show by induction on n that $n \mod 9$ equals the sum of digits mod 9. For $n \leq 9$, clear. For $n \geq 10$, $n = 10a + b$ where $b < 10$. $n \equiv 10a + b \equiv a + b \mod 9$. By induction, the sum of digits of a mod 9 equals $a \mod 9$, and thus the sum of digits of n modulo 9 is $a + b \mod 9$.

Alternate Solution: Express n (in base 10 expanded form) as

$$n = \sum_{k \geq 0} a_k \cdot 10^k.$$

Then, we see that

$$\begin{aligned} n &\equiv \sum_{k \geq 0} a_k \cdot 10^k \pmod 9 \\ &\equiv \sum_{k \geq 0} a_k \cdot 1^k \pmod 9 \\ &\equiv \sum_{k \geq 0} a_k \pmod 9 \end{aligned}$$

which means that the remainder of n when divided by 9 is the same as the remainder of the sum of the digits of n divided by 9. Thus, if the latter is divisible by 9 then so is the former.

4. Stable Matchings.

1. If a job is paired to its favorite candidate (the candidate that is first in its preference list) in some stable pairing, it must be matched to its favorite candidate in every stable pairing.

   Answer: False. Consider a two job, two candidate solution with different job and candidate optimal stable pairings.

2. Consider a job that is last on every candidate's preference list, then it must be rejected $n - 1$ times in the job-propose matching algorithm (for an n job, n candidate instance.)

   Answer: False. Consider the two job, two candidate situation where each job has a seperate favorite. The algorithm terminates without any rejections.

3. If a job proposes to its least preferred candidate in the job-proposes matching algorithm, then the algorithm finishes on that day.

   Answer: True. Every candidate has been proposed to, and must have a job on the string. Thus, the algorithm terminates.

5. Graphs.

1. Consider G_e to be result of adding an edge e to a graph G. Then G_e either contains a cycle containing e or it has fewer connected components than G.

   Answer: True. Either e is between two connected components of G, or between two vertices in the same connected component. In the former case, the number of connected components drops by 1; in the latter, there is a cycle consisting of e and a path in the connected component between its endpoints.

2. Consider a tree T with $n > 0$ vertices.

   (a) How many edges does T have?

   Answer: $n - 1$. By definition, or one can prove that a connected acyclic graph must have $n - 1$ edges, and can have no more, since more will create a cycle.

(b) Removing a subset of k edges from T forms a graph with _______ connected components on the vertices of T.
Answer: $k+1$. We start with one connected component and every edge removed increases the number of connected components by 1.

(c) If T has exactly 2 vertices of degree 5 then it must have at least ______ vertices of degree one. (Give a tight bound.)
Answer: 8. The total degree is $2n-2$, and 2 vertices have degree 5 so the total degree of the other $n-2$ vertices is $2n-12$. Say $n-2-k$ other vertices have degree greater than one and k have exactly 1. Then the total degree of the other vertices is at least $2(n-2-k)+k$ and exactly $2n-12$ which means $k \geq 8$.

3. Consider removing a simple cycle from an n vertex Eulerian graph (graph that contains an Eulerian tour) resulting in a graph with c components.

(a) Each connected component of the resulting graph is Eulerian.
Answer: True. Each vertex either is not in the simple cycle or has its degree changed by 2. Thus, every vertex has even degree. And each component is connected and is therefore Eulerian.

(b) The length of the simple cycle is at least _______. (Perhaps in terms of c and n; give a tight bound.)
Answer: c. If one collapses the connected components, the (remaining) edges of the simple cycle must connect all the components into a cycle.

4. A hypercube of dimension 4 has an Eulerian tour.
Answer: True. It has even degrees and is connected. All hypercubes are bipartite as every cycle has an even length.

5. Given a graph $G=(V,E)$ that can be vertex-colored with k colors and $G'=(V,E')$ that is bipartite, then $G''=(V,E\cup E')$ can be vertex-colored with ______ colors. (Give a tight bound.)
Answer: $2k$. The bipartite graph can be two colored. And a color can be assigned to vertex v, that is (a,b) where a is the coloring for G' and b is the color for G. The coloring is valid as any edge from G has a different values for a in its endpoints, and any edge from G' has different values for b in its endpoints.

6. Consider a connected simple planar graph with n vertices.

(a) If it has $2n$ edges, what is the number of faces in terms of n?
Answer: $n+2$. Euler's formula is $n+f=e+2$, and thus $f=n+2$

(b) Given that the average face size is 5, what is the number of edges in terms of n?
Answer: $\frac{5}{3}n-\frac{10}{3}$. We have that $2e=5f$ as each edge partitipates in two faces, or $f=\frac{2}{5}e$, and $n=\frac{3}{5}e+2$.

6. Modular Arithmetic

1. Consider the equation $47(a)+49(b)=1$, for integers a and b. Find a solution of the form (a,b) that satisfies the above equation.
Answer: One solution is (a=24, b=-23). One can run Extended Euclid's algorithm. Alternatively, notice that (47-49) = -2, so 23(47-49) = -46. Thus, we add one more 47 to get 1.

2. What is $2^{30} \pmod{35}$?
Answer: $29 \pmod{35}$. $2^{30}=2^{(6)(4)+6}=2^6=64=29 \pmod{35}$

3. Let p, q be distinct primes.

(a) Simplify $q^p \pmod{pq}$. (Hint: use CRT.)
Answer: q. $q^p=q \pmod{p}$ and $q^p \pmod{q}=0$, thus $q^p=q \pmod{pq}$.

(b) There are $pq-(\_\_\_\_\_\_)$ elements of $\{0,\ldots,pq-1\}$ that have inverses under arithmetic modulo pq.

Answer: $(p+q)-1$. Subtract the ones that are divisible by p or q and add back 1 for the element 0 which is divisible by both p and q.

7. Mod: Proof.

1. (5 points) Prove that for a prime p, any $a \in \{2,\ldots,p-2\}$ satisfies $a^{-1} \not\equiv a \pmod{p}$. (Think about the equation $x^2 \equiv 1 \pmod{p}$)

   Answer: The number of roots is at most 2 for the polynomial x^2-1, thus only 1 and -1 have the property that they are their own inverse.

2. What is the product of all elements of $\{1,\ldots,p-1\}$ modulo p? (Simplify.)

   Answer: -1 or $p-1 \pmod{p}$.

3. (5 points) For distinct primes $p,q>2$, show that there exists an $a \in \{2,\ldots,pq-2\}$ where $a^{-1} \equiv a \pmod{pq}$. (Hint: perhaps use CRT.)

   Answer: Consider a where $a=-1 \pmod{p}$ and $a=1 \pmod{q}$, it is neither $1 \pmod{pq}$ nor $-1 \pmod{pq}$ by CRT and $a^2=1 \pmod{pq}$ since $a^2=1 \pmod{p}$ and $a^2=1 \pmod{q}$ by CRT.

8. Polynomials.

1. Give a polynomial of degree 2 modulo 5, with roots at $x=1$ and $x=0$.

   Answer: $x(x-1) \pmod{5}$. Plug in 0 or 1 and you get 0.

2. Let $P(x)$ be a polynomial of degree exactly d_P and let $Q(x)$ be a distinct polynomial of degree exactly d_Q. What is the maximum number of distinct values of x that satisfy $P(x)=Q(x)$?

   Answer: $\max(d_P,d_Q)$. It's the number of roots of $P(x)-Q(x)$ which could be a degree $\max(d_P,d_Q)$ polynomials with d roots.

3. Rao constructed a degree 1 polynomial $P(x)=mx+b \pmod{p}$ where p is prime, where b is his bank account password, and m is a random value. Rao has three kids; he gives the oldest kid $P(1)$, gives the second child $P(2)$, and gives the youngest the coefficient m.

   (a) Any two kids can find Rao's password.

   Answer: True. The two older kids can use Lagrange interpolation to reconstruct $P(x)$. See the next part for when the youngest one and an older kid collaborate.

   (b) Give an expression for finding the password b, given m and $P(1)$. (Answer should be possibly in terms of m, $P(1)$ and/or their inverses or constants modulo p.)

   Answer: $P(1)-m \pmod{p}$. $m+b=P(1) \implies b=P(1)-m \pmod{p}$

9. Countability/Computability

1. The set of subsets of the prime numbers is countable.

   Answer: False. The power set of of countably infinite set is uncountable. The set of prime numbers is infinite.

2. Given a set of sets, $\mathscr{A}=\{A_1,A_2,\ldots,A_n\}$, and the set $D=\{i \mid i \notin A_i\}$, then $D \notin \mathscr{A}$.

   Answer: True. D cannot be set i, for any set i.

3. The set of programs that halt on the input 1 is countable.

   Answer: True. It is a subset of the binary strings.

4. There is a program that takes a program P, an input x, and a natural number n, and determines whether $P(x)$ runs for more than n steps.

   Answer: True. One can run the program and keep track of the memory and if it ever uses more than m space, one returns No. If it halts without using more than space m, one says Yes. If one ever reproduces the same state, which consists of the location in the program and the memory contents, then it must be in an infinite loop, and will thus never use more than m space. The last case can be detected in $O(|P|2^m)$ time where $|P|$ is the size of the program P.

10. Counting.

1. How many distinct ways can you rearrange the letters in the word "CAT"?

   Answer: 6.

2. A histogram with k bins and n observations assigns a nonnegative integer to each bin, such that the values add up to n. How many such histograms are there? (Your answer should be in terms of n and k.)

   Answer: $\binom{n+k-1}{k-1}$. There are n stars and $k-1$ bars to split the stars into k groups.

3. Consider a staircase of height n which consists of k steps, where each step is of height 1 or 2. (Answers should be in terms of n and k.)

   (a) How many steps of height 2 are there exactly?
   Answer: $n-k$. Let s be the number of height 2 steps. Then $2s+(k-s)=k+s=n$ to get to height n.

   (b) How many height n staircases are there with k steps?
   Answer: $\binom{k}{n-k}$. Choose 2 foot steps out of sequence of steps.

11. Discrete Probability, Conditional Probability, Independence

1. There are two bags of marbles. The first bag has 4 red marbles and 4 blue marbles. The second bag has 3 red marbles and 5 blue marbles. You pick one bag uniformly at random and draw 2 marbles from it without replacement. Let E be the event that both marbles you drew were blue. Let A be the event that you picked the first bag.

   (a) Compute $\mathbb{P}[E \mid A]$.
   Answer: $\mathbb{P}[E|A] = \frac{4}{8}\cdot\frac{3}{7} = \frac{3}{14}$.

   (b) Compute $\mathbb{P}[A \mid E]$.
   Answer: $\mathbb{P}\left[E|\bar{A}\right] = \frac{5}{8}\cdot\frac{4}{7} = \frac{5}{14}$.

$$\mathbb{P}[A|E] = \frac{\mathbb{P}[A]\,\mathbb{P}[E|A]}{\mathbb{P}[A]\,\mathbb{P}[E|A]+\mathbb{P}\left[\bar{A}\right]\mathbb{P}\left[E|\bar{A}\right]} = \frac{0.5\cdot 3/14}{0.5\cdot 3/14+0.5\cdot 5/14} = \frac{3}{8}.$$

2. You toss two fair dice (with labels 1 through 6).

   (a) Let X be their sum modulo 3. What is the probability that $X=1$?
   Answer: Let Z be the sum of the two dice.

$$\mathbb{P}[X=1] = \mathbb{P}[Z=4 \vee Z=7 \vee Z=10] = \frac{3+6+3}{36} = \frac{12}{36} = 1/3$$

In fact, the distribution of X is uniform over $\{0,1,2\}$.
Alternative Solution: Let X_1 and X_2 be the results of the first and second tosses. For any $i \in \{0,1,2\}$ and any choice of X_1, there are exactly 2 out of the 6 options that would make $X_1+X_2 \equiv i(\bmod 3)$. Thus, $\mathbb{P}[X=i]=1/3$.

(b) Let Y be their sum modulo 4. What is the probability that $Y=3$?
Answer: Let Z be the sum of the two dice.

$$\mathbb{P}[Y=3]=\mathbb{P}[Z=3 \vee Z=7 \vee Z=11]=\frac{2+6+2}{36}=\frac{10}{36}$$

(c) What is $\mathbb{P}[(X=1)\cap(Y=3)]$?
Answer: By the Chinese-Remainder-Theorem, $(X=1)\cap(Y=3)$ if and only if $Z=7$ which happens with probability $1/6$.

(d) What is $\mathbb{P}[(X=1)\cup(Y=3)]$?
Answer: By inclusion-exclusion formula

$$\begin{aligned}\mathbb{P}[X=1\cup Y=3]&=\mathbb{P}[X=1]+\mathbb{P}[Y=3]-\mathbb{P}[X=1\cap Y=3]\\&=\frac{1}{3}+\frac{10}{36}-\frac{1}{6}=\frac{12+10-6}{36}=\frac{16}{36}=\frac{4}{9}.\end{aligned}$$

3. You play "Let's Make a Deal" but the host, Monty Hall, doesn't know where the prize is. The setting is as follows:

- Carol picks where to put the prize uniformly at random from $\{1,2,3\}$, but tells no one (including Monty).
- You first pick door number 1.
- Monty picks one of the other doors $\{2,3\}$ at random and opens it. In the case he opens the door with the prize, you automatically lose. Otherwise, he offers to let you switch to the other closed door.

Let A be the event that Monty opened a door *without* the prize. Let B be the event that door number 1 has the prize.

(a) Compute $\mathbb{P}[A]$.
Answer: 2/3. Denote by Ω as the sample space of the game. $\Omega=\{(i,j)|i\in\{1,2,3\},j\in\{2,3\}\}$ where i denotes Carol's choice and j denotes Monty's choice. Each outcome is equally likely and happens with probability $1/6$. Thus,

$$\mathbb{P}[A]=\mathbb{P}[(1,2)]+\mathbb{P}[(1,3)]+\mathbb{P}[(2,3)]+\mathbb{P}[(3,2)]=2/3.$$

(b) Compute $\mathbb{P}[B \mid A]$.
Answer: 1/2. By Bayes $\mathbb{P}[B|A]=\frac{\mathbb{P}[A|B]\cdot\mathbb{P}[B]}{\mathbb{P}[A]}=\frac{1\cdot 1/3}{2/3}=\frac{1}{2}$.

12. Discrete Random Variables

1. We throw n distinguishable balls into n distinguishable bins. Let X be the number of balls in bin 1.

(a) Compute $\mathbb{P}[X=0]$.
Answer: $(\frac{n-1}{n})^n$

(b) What is the distribution of X?
Answer: $X \sim Binom(n, \frac{1}{n})$

(c) As n grows to infinity, what does $\mathbb{P}[X = 10]$ converge to? Your answer should not contain a limit.
Answer: As n grows to infinity, X converges to $Pois(\lambda = E[X] = 1)$, so $P[X = 10] = e^{-1}\frac{1^{10}}{10!}$. Alternatively, you can compute the limit of the binomial PMF expression.

2. We flip a fair coin 100 times. Let X be the number of heads minus the number of tails.

(a) Compute $\mathbb{P}[X = 0]$.
Answer: $\mathbb{P}[X = 0] = \mathbb{P}[50 \text{ heads}] = \frac{\binom{100}{50}}{2^{100}}$.

(b) What is $\mathbb{E}[X]$?
Answer: $X = X_1 + \ldots X_{100}$ where $X_i = 1$ if the i-th flip was H and $X_i = -1$ otherwise. Since heads and tails are equally as likely, $\mathbb{E}[X_i] = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$. Thus, $\mathbb{E}[X] = \sum_{i=1}^{100} \mathbb{E}[X_i] = 0$.

(c) (4 points) Let $Y = X^2$. Are X and Y independent? (Justify briefly)
Answer: X and Y are not independent. For example, $\mathbb{P}[X = 0 \mid Y = 0] = 1$ but $\mathbb{P}[X = 0] = \frac{\binom{100}{50}}{2^{100}} < 1$.

(d) What is $\mathrm{Cov}(X,Y)$?
Answer: Surprisingly, X and Y are uncorrelated. $\mathbb{E}[XY] = \mathbb{E}\left[X^3\right] = \sum_i i^3 \cdot \mathbb{P}[X = i]$. Because of symmetry, $\mathbb{P}[X = i] = \mathbb{P}[X = -i]$ and the two cancel each other, making the sum 0. From (b), $\mathbb{E}[X] = 0$. Thus, $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0 - 0 = 0$.

13. Expectation, Variance, Covariance

1. Gavin has n different pairs of socks (n left socks and n right socks, for $2n$ individual socks total) and is doing his laundry. He notices that the laundry machine spits out a uniformly random permutation of the $2n$ socks.

Let X be the number of matching pairs that are placed next to each other.

As an example, for the outcome $1_L 3_R 1_R 2_R 2_L 3_L$, $X = 1$ since only the 2nd pair of socks are placed next to each other.

(a) What is the probability that the 1st pair of socks are placed next to each other?
Answer: Consider the ith matching pair as a single, condensed unit. As an example, in for $n = 3$, an original permutation could look like 132213. Let us condense both the 2's together, and label it as B. Then, a resulting string would look like $13B13$. Then, there are $2n-1$ 'units' left that we can order, and thus $(2n-1)!$ ways to order them. Also, when we condensed them, either the left sock or the right sock could've came first, so there are 2 ways to condense this pair. Thus, the probability is $\frac{2(2n-1)!}{(2n)!} = \frac{1}{n}$.
Alternative Solution: The left sock is in the side (first or last) with probability $1/n$. If it is on the side, the right sock is placed next to it with probability $1/(2n-1)$. Otherwise, the right sock is placed next to it with probability $2/(2n-1)$. Overall, we get

$$\mathbb{P}[\text{1st pair is next to each other}] = \frac{1}{n} \cdot \frac{1}{2n-1} + (1 - \frac{1}{n}) \cdot \frac{2}{2n-1} = \frac{1 + 2(n-1)}{n(2n-1)} = \frac{1}{n}.$$

(b) What is $\mathbb{E}[X]$?
Answer: Let X_i be an indicator for the ith pair of matching socks, $X_i = 1$ if the socks are placed together, and 0 otherwise. Then, $X = \sum_{i=1}^{n}(X_i)$, since there are n pairs. Thus, $\mathbb{E}(X) = \sum_{i=1}^{n}\mathbb{E}(X_i) = n \cdot \mathbb{P}[X_i = 1] = \frac{2n(2n-1)!}{(2n)!} = 1$.

(c) What is the probability that both the 1st pair are placed next to each other and the 2nd pair are placed next to each other?

Answer: Again, we consider condensing both pairs. There are $2^2 = 4$ ways to condense both pairs. Once condensed, there are $2n-2$ units left, and thus $(2n-2)!$ ways to order them, so the probability becomes $4\frac{(2n-2)!}{(2n)!} = \frac{4}{(2n)(2n-1)}$.

(d) Let p be the answer to (a), and let q be the answer to (c). What is $\text{Var}(X)$? Feel free to leave your answer in terms of p and q.

Answer: We have

$$\mathbb{E}[X^2] = \mathbb{E}[(X_1+\cdots+X_n)^2] = \sum_{i=1}^{n}\mathbb{E}[X_i^2] + \sum_{i\neq j}\mathbb{E}[X_iX_j] = n\mathbb{E}[X_1^2] + n\cdot n(n-1)\mathbb{E}[X_1X_2] = np + n(n-1)q$$

where p is the answer to part a, and q is the answer to part c. So,

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = np + n(n-1)q - n^2p^2 = 1 + \frac{4n(n-1)}{2n(2n-1)} - 1 = \frac{2n-2}{2n-1}.$$

2. In a class with k students, each student has a uniformly random birthday in $\{1,\ldots,365\}$, and the students' birthdays are independent.

(a) What is the expected number of pairs of students that share a birthday?

Answer: $X = \sum_{i<j} X_{i,j}$ where $X_{i,j}$ is the indicator that student i and student j share the same birthday. $\mathbb{E}[X] = \sum_{i<j}\mathbb{E}[X_{i,j}] = \binom{k}{2}\cdot\frac{1}{365}$.

(b) What is the expected number of triplets of students that share a birthday?

Answer: $X = \sum_{i<j<\ell} X_{i,j,\ell}$ where $X_{i,j,\ell}$ is the indicator that students i, j and ℓ share the same birthday. $\mathbb{E}[X] = \sum_{i<j<\ell}\mathbb{E}\left[X_{i,j,\ell}\right] = \binom{k}{3}\cdot\frac{1}{365^2}$.

(c) What is the variance of the number of pairs of students that share a birthday?

Answer: $X = \sum_{i<j} X_{i,j}$. We observe that the random variables $X_{i,j}$ are pairwise independent. This is since $\mathbb{E}\left[X_{i,j}X_{i',j'}\right] = 1/365^2$ if $\{i,j\}$ and $\{i',j'\}$ are disjoint and also in the case they share an index, i.e., $i=i', j\neq j'$. Thus $\text{Var}(X) = \sum_{i<j}\text{Var}(X_{i,j}) = \binom{k}{2}\frac{1}{365}\frac{364}{365}$.

14. Concentration Inequalities, Confidence Intervals, LLN

We flip a biased coin with unknown heads probability p, and want to estimate $\frac{1}{p}$. Let X be the number of flips until we get k heads.

1. (5 points) Show that $Y = \frac{X}{k}$ is an unbiased estimator for $\frac{1}{p}$; that is, $\mathbb{E}[Y] = \frac{1}{p}$.

Answer: $X = X_1 + \ldots + X_k$ where $X_i \sim Geo(p)$ and thus $\mathbb{E}[X] = k\cdot\mathbb{E}[X_1] = k/p$. Thus, $\mathbb{E}[Y] = \frac{1}{p}$

2. Compute $\text{Var}(Y)$.

Answer: $X = X_1 + \ldots + X_k$ where $X_i \sim Geo(p)$ are independent and thus $\text{Var}(Y) = \frac{k\cdot\frac{1-p}{p^2}}{k^2} = \frac{1-p}{kp^2}$.

3. Use Chebyshev's inequality to upper bound $\mathbb{P}\left[\left|Y-\frac{1}{p}\right| \geq \varepsilon\cdot\frac{1}{p}\right]$.

Answer: Using Chebyshev's inequality, we have $\mathbb{P}\left[|Y-\frac{1}{p}| \geq \varepsilon\cdot\frac{1}{p}\right] \leq \frac{\text{Var}(Y)}{\varepsilon^2\cdot\frac{1}{p^2}} = \frac{(1-p)/(kp^2)}{\varepsilon^2/p^2} = \frac{1-p}{\varepsilon^2 k}$

4. Suppose p is unknown, and we only know $p \geq 0.5$. What value of k should we pick to ensure $\mathbb{P}\left[|Y-\frac{1}{p}| < 0.1\cdot\frac{1}{p}\right] \geq 0.95$?

Answer: Since we know $p \geq 0.5$, the above calculation gives $\mathbb{P}\left[|Y - \frac{1}{p}| \geq 0.1\frac{1}{p}\right] \leq \frac{1-p}{0.1^2 k} \leq \frac{0.5}{0.1^2 k} = \frac{50}{k}$. If we pick $k = 50 \cdot 20$ we get $\mathbb{P}\left[|Y - \frac{1}{p}| \geq 0.1\frac{1}{p}\right] \leq 0.05$. Thus, the probability of the complement event $|Y - \frac{1}{p}| < 0.1\frac{1}{p}$ is at least 0.95.

15. Continuous Probability

1. (10 points) You drive from Berkeley to Los Angeles. The distance is 360 miles. You choose a speed (in miles per hour) uniformly at random from the interval $[40,60]$ and drive. What is the expected driving time in hours? (Show your work.)

   Answer: If $S \sim U([40,60])$ is the speed, then $T = 360/S$ is the driving time. We seek to calculate $\mathbb{E}[T] = \mathbb{E}[360/S]$. We observe that the pdf of S is $f_S(s) = 1/20$ for $s \in [40,60]$ and $f_S(s) = 0$ otherwise. We get

$$\mathbb{E}[360/S] = \int_{-\infty}^{\infty} \frac{360}{s} \cdot f_S(s) ds = \frac{360}{20} \int_{40}^{60} \frac{1}{s} ds = 18 \cdot (\ln(60) - \ln(40)) = 18 \ln(1.5) \approx 7.29.$$

2. Let $X, Y, Z \sim U([0,1])$ be mutually independent.

   (a) What is the CDF $F_1(t)$ of $\max(X,Y,Z)$, for $t \in [0,1]$?
   Answer: For $t \in [0,1]$, the c.d.f. is $\mathbb{P}[\max(X,Y,Z) \leq t] = t^3$.

   (b) What is the PDF $f_1(t)$ of $\max(X,Y,Z)$, for $t \in [0,1]$?
   Answer: Thus, the p.d.f is the derivative of the c.d.f., which is $3t^2$.

   (c) What is the expectation of $\max(X,Y,Z)$?
   Answer: $\mathbb{E}[\max(X,Y,Z)] = \int_0^1 t \cdot 3t^2 dt = \frac{3}{4} t^4 \Big|_0^1 = \frac{3}{4}$.

   (d) What is the PDF $f_2(t)$ of $\min(X,Y,Z)$, for $t \in [0,1]$?
   Answer: For $t \in [0,1]$, the c.d.f. is $\mathbb{P}[\min(X,Y,Z) \leq t] = 1 - \mathbb{P}[\min(X,Y,Z) \geq x] = 1 - (1-t)^3$. Thus, the p.d.f. is the derivative which is $3(1-t)^2$.

   (e) The median of three numbers is the middle number after sorting, e.g., $\text{median}(0.6, 0.3, 0.9) = 0.6$. What is the expectation of $\text{median}(X,Y,Z)$?
   Answer: $\mathbb{E}[\text{median}(X,Y,Z)] = \mathbb{E}[X+Y+Z - \max(X,Y,Z) - \min(X,Y,Z)] = 1.5 - \mathbb{E}[\max(X,Y,Z)] - \mathbb{E}[\min(X,Y,Z)]$. It is easy to see that $\mathbb{E}[\max(X,Y,Z)] = \mathbb{E}[1 - \min(X,Y,Z)]$ since they have the same p.d.f, thus $1.5 - \mathbb{E}[\max(X,Y,Z)] - \mathbb{E}[\min(X,Y,Z)] = 1.5 - (1 - \mathbb{E}[\min(X,Y,Z)]) - \mathbb{E}[\min(X,Y,Z] = 0.5$.
   Alternative solution I: By symmetry $\mathbb{E}[\text{median}(X,Y,Z)] = \mathbb{E}[\text{median}(1-X, 1-Y, 1-Z)]$ and their sum is always 1, thus $\mathbb{E}[\text{median}(X,Y,Z)] = 1/2$.
   Alternative solution II: For small dt, $\mathbb{P}[\text{median}(X,Y,Z) \in [t, t+dt]] \approx 3! \cdot dt \cdot t \cdot (1-t)$. Thus, the p.d.f. is $6t(1-t)$. The expectation is

$$\int_0^1 6t^2(1-t) dt = (2t^3 - 1.5t^4)\Big|_0^1 = 2 - 1.5 = 0.5$$

3. Let $T_S \sim N(4, 0.2)$ and $T_J \sim N(4, 0.3)$ be two independent random variables.

   (a) What's the distribution of $T_S + T_J$?
   Answer: The sum of independent Gaussians is a Gaussian. Specifically, $T_S + T_J \sim N(8, 0.5)$.

(b) What's the distribution of $T_S - T_J$?
Answer: $-T_J \sim N(-4, 0.3)$ since its mean is -4 and its variance is 0.2. Since the sum of two independent Gaussians is Gaussia, we see that $T_S - T_J \sim N(0, 0.5)$.

(c) Compute $\mathbb{P}[T_S > T_J]$.
Answer: We see that $T_S - T_J \sim N(0, 0.5)$. Hence,

$$\mathbb{P}[T_S > T_J] = \mathbb{P}[T_S - T_J > 0] = \frac{1}{2}$$

by symmetry.

16. Conditional Expectation/Estimation

1. Suppose a store has 3 circular dartboards with radii a, b and c. Shreyas is in a rush so he buys one dartboard at random and leaves the store. Suppose his throws are uniformly distributed across the entire dartboard. Shreyas throws a dart at random at his board.

(a) What is the dart's expected distance to the center? Your answer may depend on the variables a, b, and c.
Answer: $\frac{2(a+b+c)}{9}$.

(b) (5 points) Justify your answer.
Answer: We use conditional expectation. Denote the random variable Y as the radius of the dartboard Shreyas chooses and X as the random variable representing the distance of the throw from the center. If the board has radius $Y = y$ then the cdf of the throw is

$$F_{X|Y=y}(x) = \mathbb{P}[X \le x] = \frac{\pi x^2}{\pi y^2} = \frac{x^2}{y^2}.$$

Differentiating gives us the PDF

$$f_{X|Y=y}(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_{X|Y=y}(x) = \frac{2x}{y^2}.$$

Thus, we can compute

$$\mathbb{E}[X \mid Y = y] = \int_{-\infty}^{\infty} x f_{X|Y=y}(x)\mathrm{d}x = \int_0^y x\frac{2x}{y^2}\mathrm{d}x = \left.\frac{2x^3}{3y^2}\right|_0^y = \frac{2y}{3}.$$

Finally, since Y is uniform, $\mathbb{E}[Y] = \frac{a+b+c}{3}$. Thus, by Wald's identity,

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}\left[\frac{2Y}{3}\right] = \frac{2}{3}\mathbb{E}[Y] = \frac{2(a+b+c)}{9}.$$

2. Suppose we have a random variable Y with expectation 7 and variance 2. We also have $Z \sim N(3, 5)$, where Y and Z are independent. Let $X = Y + Z$.

(a) (3 points each) Given that $Y = y$, what is the function $\hat{X}$ in terms of y that minimizes the mean squared error $\mathbb{E}\left[(X - \hat{X})^2 \mid Y = y\right]$? What is the corresponding mean squared error?

$\hat{X} =$ Mean Squared Error $=$

Answer: $\hat{X} = \mathbb{E}[X|Y=y] = y + \mathbb{E}[Z] = y+3$.
$MSE(\hat{X}) = \text{Var}(X|Y=y) = \text{Var}(y+Z) = \text{Var}(Z) = 5$

(b) What is the linear least squares estimator $L[X \mid Y]$?
Answer: The MMSE estimator from the previous part was already linear, thus the LLSE is equal to the MMSE. Thus, $LLSE(X|Y) = Y+3$

3. (7 points) Suppose the height of a person (in inches) is distributed like $X+Y$ where X and Y are independent; $X \sim N(69,3)$ is a genetic component, and $Y \sim N(0,1)$ is an environmental component.

Two identical twins are separated at birth and raised in different environments.

The random variables $Z_1 = X+Y_1$ and $Z_2 = X+Y_2$ are the heights of twin A and twin B respectively (note that they share the genetic component). Here, $X \sim N(69,3)$, $Y_1 \sim N(0,1)$, and $Y_2 \sim N(0,1)$ are mutually independent.

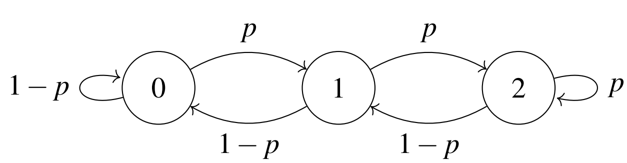
What is the linear least squares estimator of Z_2 given Z_1? (i.e., $L[Z_2 \mid Z_1]$)

Answer:

$$
\begin{aligned}
L[Z_2|Z_1] &= \mathbb{E}[Z_2] + \frac{\text{Cov}(Z_1,Z_2)}{\text{Var}(Z_1)}(Z_1 - \mathbb{E}[Z_1]) \\
&= 69 + \frac{\text{Cov}(X+Y_1, X+Y_2)}{4}(Z_1 - 69) \\
&= 69 + \frac{\text{Cov}(X,X) + \text{Cov}(X,Y_2) + \text{Cov}(Y_1,X) + \text{Cov}(Y_1,Y_2)}{4}(Z_1 - 69) \\
&= 69 + \frac{3+0+0+0}{4}(Z_1 - 69) \\
&= 69 + \frac{3}{4}(Z_1 - 69)
\end{aligned}
$$

17. Markov Chains

1. Consider the following Markov chain:



Let X_n be the random variable representing the current state of the Markov chain at step n.

(a) (5 points) Let $p = \frac{1}{3}$. Write, but do not solve, a system of linear equations to find the stationary distribution of this Markov chain.

Answer: The stationary distribution of the Markov chain must satisfy:

$$\begin{aligned}\pi(0) &= \frac{2}{3}\pi(0) + \frac{2}{3}\pi(1)\\ \pi(1) &= \frac{1}{3}\pi(0) + \frac{2}{3}\pi(2)\\ \pi(2) &= \frac{1}{3}\pi(1) + \frac{1}{3}\pi(2)\\ 1 &= \pi(0) + \pi(1) + \pi(2)\end{aligned}$$

Suppose we choose a p such that the stationary distribution of the Markov chain is

$$\pi = \begin{bmatrix} \frac{1}{13} & \frac{3}{13} & \frac{9}{13} \end{bmatrix}.$$

Let Y_n be a new random variable dependent on X_n such that $Y_n = 0$ if $X_n \in \{0,1\}$, and $Y_n = 1$ if $X_n = 2$. Suppose that in this experiment we only observe the value of Y_n, and we'd like to estimate the value of X_n given our observations. Further, suppose that in our observations, the Markov chain has already been evolving over a long period of time, and X_n has converged to the stationary distribution.

(b) Compute the minimum mean squared estimate of X_n given Y_n, i.e. $\mathbb{E}[X_n \mid Y_n]$.
Answer: If $Y_n = 0$, then we must have either $X_n = 0$ or $X_n = 1$. Similarly, if $Y_n = 1$, then we must have either $X_n = 2$. These conditional probabilities are:

$$\begin{aligned}\mathbb{P}[X_n = 0 \mid Y_n = 0] &= \frac{\frac{1}{13}}{\frac{1}{13} + \frac{3}{13}} = \frac{1}{4}\\ \mathbb{P}[X_n = 1 \mid Y_n = 0] &= \frac{\frac{3}{13}}{\frac{1}{13} + \frac{3}{13}} = \frac{3}{4}\\ \mathbb{P}[X_n = 2 \mid Y_n = 1] &= 1\end{aligned}$$

We'd like to compute $\mathbb{E}[X_n \mid Y_n]$, so when $Y_n = 0$, we have

$$\mathbb{E}[X_n \mid Y_n = 0] = 0 \cdot \frac{1}{4} + 1 \cdot \frac{3}{4} = \frac{3}{4}.$$

When $Y_n = 1$, we have

$$\mathbb{E}[X_n \mid Y_n = 1] = 2.$$

Together, we have

$$\mathbb{E}[X_n \mid Y_n] = \begin{cases} \frac{3}{4} & Y_n = 0 \\ 2 & Y_n = 1 \end{cases}.$$

(c) Compute the linear least squares estimate of X_n given Y_n, i.e. $L[X_n \mid Y_n]$. Your answer should be in the form $aY_n + b$, for some constants a and b. (Hint: there should not be much work involved.)
Answer: One thing to notice with the expression for $\mathbb{E}[X_n \mid Y_n]$ is that it is only defined by two points. This means that it is actually a linear function of Y_n; interpolating through the two points, we can write

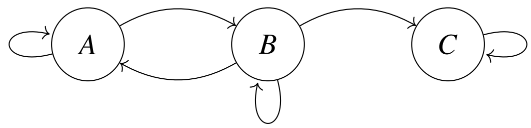
$$\mathbb{E}[X_n \mid Y_n] = \frac{5}{4}Y_n + \frac{3}{4}.$$

Since the MMSE is a linear function of Y_n, the LLSE must be equal to the MMSE, and we have

$$L[X_n \mid Y_n] = \frac{5}{4}Y_n + \frac{3}{4}$$

as well.

2. Suppose you have a fair die (with labels 1 through 6) and you keep rolling the die until the product of the last two rolls is 15. Consider the following Markov chain modeling this scenario.



Here, A is the starting state, as well as the state in which the previous roll was in $\{1,2,4,6\}$, B is the state in which the previous roll was in $\{3,5\}$ but the product of the last two rolls is not 15, and C is the state in which the product of the last two rolls was 15.

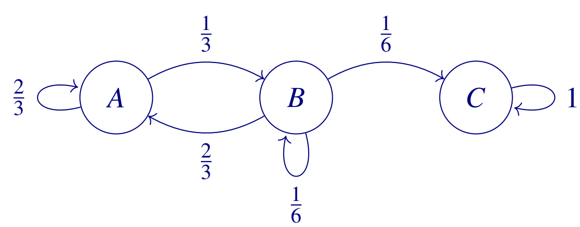
(a) (1 point each) What are the transition probabilities in this Markov chain? You only need to provide the transition probabilities of the drawn edges in the diagram above.

$\mathbb{P}[X_n = A \mid X_{n-1} = A] =$ $\qquad$ $\mathbb{P}[X_n = B \mid X_{n-1} = B] =$
$\mathbb{P}[X_n = B \mid X_{n-1} = A] =$ $\qquad$ $\mathbb{P}[X_n = C \mid X_{n-1} = B] =$
$\mathbb{P}[X_n = A \mid X_{n-1} = B] =$ $\qquad$ $\mathbb{P}[X_n = C \mid X_{n-1} = C] =$

Answer: We have the following transition probabilities:

$$\mathbb{P}[X_n = A \mid X_{n-1} = A] = \frac{2}{3}$$
$$\mathbb{P}[X_n = B \mid X_{n-1} = A] = \frac{1}{3}$$
$$\mathbb{P}[X_n = A \mid X_{n-1} = B] = \frac{2}{3}$$
$$\mathbb{P}[X_n = B \mid X_{n-1} = B] = \frac{1}{6}$$
$$\mathbb{P}[X_n = C \mid X_{n-1} = B] = \frac{1}{6}$$
$$\mathbb{P}[X_n = C \mid X_{n-1} = C] = 1$$

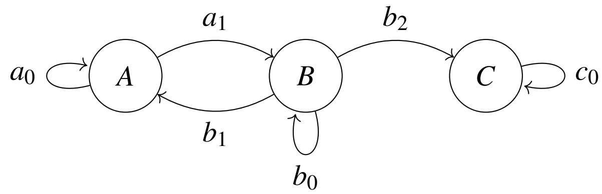
In diagram form, we have



(b) This Markov chain is irreducible.
Answer: False. We cannot get to states A and B from state C.

(c) This Markov chain is aperiodic.
Answer: True. Since each state has a self-loop with nonzero probability, the period of each state is 1, so the chain is aperiodic.

(d) (6 points) Let $\beta(i)$ denote the expected number of rolls you make until the last two rolls multiply to 15, given that you are currently in state i of the Markov chain. Write, but do not solve, a system of linear equations involving $\beta(i)$ for $i \in \{A,B,C\}$. (You should use the variables in the following diagram as the coefficients in your equations.)



Answer: We have the following system of equations:

$$\begin{aligned}
\beta(A) &= 1 + a_0\beta(A) + a_1\beta(B) \\
\beta(B) &= 1 + b_1\beta(A) + b_0\beta(B) + b_2\beta(C) \\
\beta(C) &= 0
\end{aligned}$$

The system above is what we expected for full credit.

With the actual probabilities, we have

$$\begin{aligned}
\beta(A) &= 1 + \frac{2}{3}\beta(A) + \frac{1}{3}\beta(B) \\
\beta(B) &= 1 + \frac{2}{3}\beta(A) + \frac{1}{6}\beta(B) + \frac{1}{6}\beta(C) \\
\beta(C) &= 0
\end{aligned}$$

If we actually solve the system we get $\beta(A) = 21, \beta(B) = 18, \beta(C) = 0$ which means that on average we expect it to take 21 rolls for two consecutive rolls to multiply to 15.