CS 70 Discrete Mathematics and Probability Theory

Spring 2019 Ayazifar and Rao

Midterm 2 Solutions

PRINT Your Name: Oski Bear

SIGN Your Name: $\mathscr{OSKI}$

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

1. TRUE or FALSE? 2 points each part, 26 total.

For each of the questions below, answer TRUE or FALSE. No need to justify answer.

Please fill in the appropriate bubble!

Answer: Note that the answers provide explanations for your understanding, even though no such justification was required

1. $(P \implies (R \wedge \neg R)) \implies \neg P$

   Answer: True. This is the form of a proof by contradiction of $\neg P$.

2. Let $\mathbb{Z}$ be the integers, and $P(i)$ be a predicate on integers,
   $(P(0) \wedge ((\exists i \in \mathbb{Z})\ P(i) \wedge P(i+1)) \implies (\forall i \in \mathbb{Z})\ ((i \geq 0) \implies P(i)))$

   Answer: False. Just because a particular i makes $P(i) \wedge P(i+1)$ true does not mean $P(i)$ is true for every i. Consider the predicate P such that $P(0)$, $P(4)$ and $P(5)$ are true, and false everywhere else. Then this predicate satisfies the left hand side of the implication but not the right hand side.

3. Let $\mathbb{R}$ be the real numbers, $(\forall x, y \in \mathbb{R})((x < y) \implies ((\exists z \in \mathbb{R})\ (x < z < y))))$

   Answer: True. There is always a real number between any two real numbers.

4. Let $\mathbb{Q}$ be the rational numbers, $(\forall x, y \in \mathbb{Q})((x < y) \implies ((\exists z \in \mathbb{Q})\ (x < z < y))))$

   Answer: True. There is always a rational number between any two rational numbers.

5. Any stable pairing that is optimal for one man is optimal for all men.

   Answer: False. Construct a stable marriage instance from two instances by taking the union of the men and women and extending the preference list where the people in the other instance are disliked. The only stable pairings pair people in the original instances, and the pessimal pairing in one instance and the optimal pairing in the other is stable, since there will be no rogue couple consisting of people from different instances. This pairing is optimal for half the men, and pessimal for the other half.

6. Any graph with no triangles is two colorable.

   Answer: False. Consider a single cycle of length 5.

7. There is a graph with average degree 2 that does not have a cycle.

   Answer: False. Consider the connected components. At least one has average degree greater than 1, which suggests that the number of edges is at least n, which says it is connected and not a tree and therefore must have a cycle.

8. The length of any Eulerian tour of a graph is even.

   Answer: False. Consider a triangle.

9. There is a program that takes a program P and input x and number of steps, s and returns YES if and only if P run on x halts in s steps.

   Answer: True. One can just simulate the program P run on x for s steps.

10. If one can write a program that solves a problem P using the halting problem as a subroutine then the problem P is undecidable.

    Answer: False. This is the wrong way. It could be easy to solve. To show the problem P is undecidable one would show how to solve the halting problem given a program to solve the problem P.

11. There is a bijection between the powerset of rational numbers and the real numbers. (The powerset of set S is the set of all subsets of S.)

    Answer: True. There is a bijection from the rationals to the integers and the power set of the integers has a bijection with the real numbers.

12. If $\Pr[A \cup B] = \Pr[A] + \Pr[B]$ then A and B are independent.

    Answer: False. If A and B have non-zero probability, this implies that $\Pr[A \cap B] = 0$, which is clearly not $\Pr[A] \times \Pr[B]$ which means they are not indpendent.

13. Given n balls being thrown into n bins, the event "the first bin is empty" and the event "the second bin is empty" are independent.

    Answer: False. If the first bin is empty, the second is less likely to be empty.

2. Quick proof. 7 points.

Prove that $\displaystyle\sum_{i=1}^{n} \frac{1}{i(i+1)} = \frac{n}{n+1}$.

Answer: Solution 1 (Induction): We perform induction on n.

Base Case: For $n = 1$, we have $\frac{1}{1\cdot 2} = \frac{1}{1+1}$ which is a true statement.

Inductive Hypothesis: Suppose for some k, we have $\sum_{i=1}^{k} \frac{1}{i(i+1)} = \frac{k}{k+1}$.

Inductive Step: We see that

$$\begin{aligned}
\sum_{i=1}^{k+1} \frac{1}{i(i+1)} &= \frac{1}{(k+1)(k+2)} + \sum_{i=1}^{k} \frac{1}{i(i+1)} \\
&= \frac{1}{(k+1)(k+2)} + \frac{k}{k+1} \qquad \text{(Inductive Hypothesis)} \\
&= \frac{1 + k(k+2)}{(k+1)(k+2)} \\
&= \frac{(k+1)\cancel{(k+1)}}{\cancel{(k+1)}(k+2)} \\
&= \frac{k+1}{k+2}
\end{aligned}$$

which concludes our induction.

Solution 2: Note that $\frac{1}{i(i+1)} = \frac{1}{i} - \frac{1}{i+1}$. Therefore,

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \cdots + \frac{1}{n(n+1)} = \frac{1}{1} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \cdots + \frac{1}{n-1} - \frac{1}{n}.$$

The terms in the middle all cancel and we are left with $1 - \frac{1}{n} = \frac{n}{n+1}$.

3. Short Answer: Discrete Math. 3 points each part, 48 points total.

1. What is the number of faces in a planar drawing of a planar graph with n vertices where every vertex has degree 3?

   Answer: $n/2+2$. e is the sum of degrees divided by 2 or $3n/2$ and Euler says $v+f=e+2$ so $f=e-n+2=n/2+2$.

2. Given a graph $G=(V,E)$ with k connected components, what is the minimum number of edges one needs to add to ensure that the resulting graph is connected?

   Answer: $k-1$. One can repeatedly add an edge between two components, reducing the number of components to 1 after $k-1$ edges.

3. The hypercube graph for dimension d has an Eulerian tour when $d=$ ____ $\pmod 2$.

   Answer: 0. Degree is d, which should be even.

4. For a dimension d hypercube with a Eulerian tour of length L and a Hamiltonian cycle of length ℓ, what is L/ℓ?

   Answer: $d/2$. The Eulerian path uses every edge of which there are $2^d \cdot d/2$, and the Hamiltonion path is of length 2^d, since there are 2^d vertices.

5. What is the minimum number of odd degree vertices in a connected acyclic graph?

   Answer: 2. There is at least one degree 1 vertex, and the number of odd degree vertices in any graph is even. Moreover, a path has two degree 1 vertices.

6. What is $2^{10} \pmod{11}$?

   Answer: 1. One can use Fermat's Theorem.

7. For distinct primes p,q,r and $N=pqr$, how many elements of $\{0,1,\ldots,N-1\}$ are relatively prime to N?

   Answer: $(p-1)(q-1)(r-1)$. We can use inclusion/exclusion: Start with pqr then subtract pr for the multiples of p, and qr for the multiples of p, and pq for the multiples of r, and add back p for multiples of qr, q for multiples of pr, and r for multiples of pq and subtract 1 for multiple pqr.

   This turns out to be $pqr-pr-qp-qr+p+q+r-1=(p-1)(q-1)(r-1)$.

   Alternatively, there are $p-1$ choices for what $x \pmod p$ is, $q-1$ choices for what $x \pmod q$ is, and $r-1$ choices for what $x \pmod r$ is. By CRT, any choice of what x is modulo p,q,r yields a unique value of $x \pmod{pqr}$, so the final answer is $(p-1)(q-1)(r-1)$.

8. Consider N and the set $S=\{x\in\{0,\ldots N-1\} : gcd(x,N)=1\}$ where $k=|S|$.
   For $a\in S$, we define T = $\{ax \pmod N : x\in S\}$. What is $|T|$? Answer may include N and k.

   Answer: k. The function $f(x)=ax \pmod N$ is a bijection since a is relatively prime to N. It is injective because

   $$f(x)=f(y) \Rightarrow ax=ay \pmod N \Rightarrow a^{-1}ax=a^{-1}ay \pmod N \Rightarrow x=y \pmod N.$$

   This function is surjective because for each $y\in S$, we know $f(a^{-1}y)=y$ and $a^{-1}y\in S$.

9. For a prime p, what is a positive integer x that guarantees $a^x=1 \pmod{p^2}$ for all a relatively prime to p? Answer may include p.

   Answer: $p(p-1)$. Let S be the set of integers between 0 and p^2-1 inclusive that are relatively prime to p. Then, $|S|=p^2-p$ and $f:S\to S$ where $f(x)=ax \pmod{p^2}$ is a bijection. Therefore,

   $$\prod_{i\in S} i=\prod_{i\in S} ai \pmod{p^2} \Rightarrow a^{|S|}=1 \pmod{p^2}.$$

$|S| = p(p-1)$ yields the desired conclusion.

Note: any positive multiple of $p(p-1)$ was given credit, since our proof also proves that $a^{kp(p-1)} = 1 \pmod{p^2}$ for any positive integer k.

It also turns out that $p(p-1)$ is the best that we can do. The proof of this is beyond the scope of this class.

10. For distinct primes p,q,r, what is $a^{(p-1)(q-1)(r-1)} \pmod{pqr}$, where a is relatively prime to pqr. Answer may include p,q,r.

    Answer: $1 \pmod{pqr}$. By FLT, we see $a^{(p-1)(q-1)(r-1)} \equiv 1 \pmod{p}$, $a^{(p-1)(q-1)(r-1)} \equiv 1 \pmod{q}$, and $a^{(p-1)(q-1)(r-1)} \equiv 1 \pmod{r}$, so by CRT, we must have $a^{(p-1)(q-1)(r-1)} = 1 \pmod{pqr}$.

11. Jonathan wants to tell Emaan how many chicken nuggets he ate today, which we will call c. He doesn't want the world to know, so he encrypts it with Emaan's public key (N,e), which yields the ciphertext x. Jerry intercepts the message, and wants to make it look like Jonathan actually ate 5 times as many chicken nuggets. What message should she send to Emaan? Answer may include x, N, and e. You may not include c.

    Answer: $5^e \cdot x \pmod{N}$. This works because $x = c^e \pmod{N}$, so $5^e x = (5c)^e \pmod{N}$.

    For the following parts consider two non-zero polynomials $P(x)$ and $Q(x)$ of degree d over $GF(p)$ (modulo p), with r_p roots and r_q roots respectively.

12. What is the maximum number of roots for the polynomial $P(x)Q(x)$? Answer may include d, r_p, and r_q. (Your answer should be achievable for any valid d, r_p and r_q.)

    Answer: $\min(p, r_p + r_q)$. One can only have roots at a root of $P(x)$ or $Q(x)$ which yields $r_p + r_q$ as an upper bound. However, since we are working in $GF(p)$, any polynomial can only have at most p roots.

13. What is the minimum number of roots for the polynomial $P(x)Q(x)$? Answer may include d, r_p, and r_q.

    Answer: $\max(r_p, r_q)$. The roots could completely overlap.

14. Let $S = \{(x_1,y_1),\ldots,(x_{n+2k},y_{n+2k})\}$ be a set of $n+2k$ points where the x_i are distinct. If $P(x)$ and $Q(x)$ are polynomials where $P(x_i) = y_i$ for at least $n+k$ points in S and $Q(x_j) = y_j$ for at least $n+k$ points in S, what is the minimum number of points that $P(x)$ and $Q(x)$ must agree on in S? Answer may include n and k.

    Answer: n. There are $2n+2k$ points contained in both polynomials and only $n+k$ points, so they must both contain at least $2n+2k-(n+2k) = n$ points. This is from error correction.

15. Working over $GF(5)$, describe a degree *exactly* 2 polynomial where $P(1) = 1$ and $P(2) = 2$.

    Answer: All answers of the form $c(x-1)(x-2)+x$ are accepted, for $c \in \{1,2,3,4\}$. Note that $P(x)-x$ has roots at 1 and 2, so therefore $P(x)-x = c(x-1)(x-2)$.

16. Let $P(x)$ be a degree $d = n-1$ polynomial over $GF(p)$ (p is prime) that contains all but $\ell \le k$ of $n+2k$ points which are given. In this situation, recall that the Berlekamp-Welsh procedure can reconstruct $P(x)$ by assuming the existence of an error polynomial $E(x)$ of degree exactly k and leading coefficient of 1, and a polynomial $Q(x) = P(x)E(x)$. How many possible pairs of $Q(x)$ and $E(x)$ are consistent with the Berlekamp-Welsh procedure? Answer may include ℓ, k, d, n, and p.

    Answer: $p^{k-\ell}$. We know that $E(x)$ is divisible by $(x-e_1)(x-e_2)\cdots(x-e_\ell)$, so $E(x) = (x-e_1)(x-e_2)\cdots(x-e_\ell)R(x)$ for some polynomial $R(x)$, where the degree of $R(x)$ is $k-\ell$ and its leading coefficient is 1. There are $p^{k-\ell}$ such $R(x)$.

4. Short Answer: Counting. 3 points each. 12 points total.

1. What is the number of ways to place n distinguishable balls into k distinguishable bins?
   Answer: k^n. Each of n balls has k possibilities.
2. What is the number of ways to place n distinguishable balls into k distinguishable bins where no two balls are placed in the same bin? You may assume that $n \le k$.
   Answer: $\frac{k!}{(k-n)!}$. k ways to choose the first one, $k-1$ ways to choose the second and so on.
3. What is the number of ways to divide d dollar bills among p people? Assume dollar bills are indistinguishable and people are distinguishable.
   Answer: $\binom{d+p-1}{p-1}$. This is stars and bars where the dollars correspond to d stars and the people correspond to $p-1$ stars.
4. How many $(x_1,\dots,x_k,y_1,y_2,\dots,y_k)$ are there such that all x_i, y_i are non-negative integers, $\sum_{i=1}^{k} x_i = n$, and $y_i \le x_i$ for $1 \le i \le k$? Answer may *not* include any summations.
   Answer: $\binom{n+2k-1}{2k-1}$. Define $z_i = x_i - y_i$ for each i. Then, we see that $z_i \ge 0$ and

$$\sum_{i=1}^{k} z_i + \sum_{i=1}^{k} y_i = n$$

   From stars and bars, there are $\binom{n+2k-1}{2k-1}$ ways to pick the y_i and the z_i. x_i can be uniquely constructed from y_i and z_i, so this is our final answer.

5. Short Answer: Probability. 3 points each part, 18 points total.

1. Given two tosses of a fair coin, what is Pr[heads on the second coin|at least one heads in the two tosses].
   Answer: $2/3$. Each of the four possibilities HH,TH,HT,TT are equally likely. And out of the three with a heads have heads for the second coin.
2. Consider two events, A and B with $\Pr[A \cup B] = \frac{3}{4}$, and $\Pr[A] = \frac{1}{2}$, and $\Pr[B] = \frac{4}{5}$, what is $\Pr[A \cap B]$?
   Answer: $\frac{11}{20}$. Inclusion/Exclusion. $\frac{1}{2} + \frac{4}{5} - \frac{3}{4} = \frac{11}{20}$.
3. Alice and Bob both try to a climb a rope. Alice and Bob will get to the top of the rope with probability $1/3$ and $1/4$ respectively. Given that exactly one person got to the top, what is the probability that the person is Alice?
   Answer: $\frac{3}{5}$. The probability that exactly one person makes it to the top is $\frac{1}{3}\cdot\frac{3}{4} + \frac{2}{3}\cdot\frac{1}{4} = \frac{5}{12}$. Then, the probability that it was actually Alice and not Bob is $\frac{1}{3}\cdot\frac{3}{4} = \frac{1}{4}$. Dividing $1/4$ by $5/12$ yields $3/5$.
4. Given $X \sim \text{Geom}(p)$, what is $\Pr[X = i | X > j]$? Assume $i > j$.
   Answer: $p(1-p)^{(i-j-1)}$. One needs a success and $(i-j-1)$ failures to get from j to i.
5. Given independent $X,Y \sim \text{Bin}(n,p)$, what is $\Pr[X+Y=i]$?
   Answer: $\binom{2n}{i}p^i(1-p)^{2n-i}$. $X+Y \sim \text{Bin}(2n,p)$
6. Consider a random variable X where $E[X^4] = 5$, give as good upper bound on $\Pr[X \ge 5]$ as you can.
   Answer: $\frac{1}{125}$.
   $\Pr[X \ge 5] = \Pr[X^4 \ge 5^4] \le \frac{E[X^4]}{5^4} = \frac{1}{5^3}$.
   The random variable $X = 5$ with probability $\frac{1}{125}$ and 0 demonstrates that this is the best possible bound.

6. Concepts through balls in bins. 3 points each part, 18 points total.

Consider throwing n balls into n bins uniformly at random. Let X be the number of balls in the first bin.

1. What is the expected value of X?

   Answer: 1. One can use linearity of expectation: $X_1 + \cdots X_n$, where X_1 indicates that ball i falls into bin 1 and $E[X_1] = 1/n$.

2. Use Markov's inequality to give an upper bound on $\Pr[X \geq k]$.

   Answer: $1/k$. Markov says $\Pr[X \geq k] \leq \frac{E[X]}{k}$.

3. What is the variance of X?

   Answer: $1 - \frac{1}{n}$.

   Take $X = X_1 + \ldots + X_n$ where X_i is an indicator random variable for choosing ball i choosing bin 1. From the fact that each X_i has variance $\frac{1}{n}(1 - \frac{1}{n})$ and the fact that X_is are independent from each other, we get $n\left(\frac{1}{n}\left(1 - \frac{1}{n}\right)\right) = (1 - 1/n)$.

4. Use Chebyshev's inequality to give an upper bound on $\Pr[X \geq k]$.

   Answer: $\frac{1-\frac{1}{n}}{(k-1)^2}$. From Chebyshev's, we know that

$$\Pr[|X - 1| \geq k - 1] \leq \frac{1 - 1/n}{(k-1)^2}$$

   Furthermore, we know $\Pr[|X - 1| \geq k - 1] = \Pr[X \geq k \cup X \leq 2 - k] \geq \Pr[X \geq k]$.

5. Now let Y be the number of balls in the second bin. What the joint distribution of X, Y, i.e., what is $\Pr[X = i, Y = j]$?

   Answer: $\binom{n}{i}\binom{n-i}{j}\left(\frac{1}{n}\right)^{i+j}\left(1 - \frac{2}{n}\right)^{n-i-j}$. There are $\binom{n}{i}$ ways to choose which balls will go into the first bin, and $\binom{n-i}{j}$ ways to pick which balls go into the second bin. After picking, the probability that the i balls actually end up in the first bin is $\frac{1}{n^i}$, the probability that the j balls end up in the second bin is $\frac{1}{n^j}$, and the probability that the other $n - i - j$ balls end up in not the first and second bins is $\left(1 - \frac{2}{n}\right)^{n-i-j}$.

6. What is $\Pr[X = i | Y = j]$?

   Answer: $\binom{n-j}{i}\left(\frac{1}{n-1}\right)^i\left(\frac{n-2}{n-1}\right)^{n-i-j}$. We can pretend the second bin doesn't exist, since it will have j balls already. Then, there are $n - j$ balls left, thrown into $n - 1$ bins.

7. Lots of chicken nuggets. 5 points each part, 15 points total.

We will model the number of customers going into Emaan's and Jonathan's favorite McDonalds within an hour as a random Poisson variable, i.e., $X \sim \text{Poisson}(\lambda)$.

1. Given that $\lambda = 5$, what is the probability that 5 people come in during the hour that Emaan and Jonathan are eating chicken nuggets?

   Answer: $e^{-5}\frac{5^5}{5!}$.

2. If λ is unknown but is definitely at most 10, how many hours do Emaan and Jonathan need to be at McDonalds to be able to construct a 95% confidence interval for λ that is of width 2. (You should use Chebyshev's inequality here. Recall for $X \sim \text{Poisson}(\lambda)$ that $Var(X) = \lambda$)

   Answer: 200 hours. For the $X \sim \text{Poisson}(\lambda)$, the variance is λ. If we observe n samples $X_1, X_2, \ldots, X_n$, we can estimate λ as $Y := \frac{X_1 + X_2 + \cdots + X_n}{n}$. Then we have $E[Y] = \lambda$ and $Var(Y) = \frac{\lambda}{n}$.

   Chebyshev says $\Pr[|Y - \mu| \geq t] \leq \frac{Var(Y)}{t}$. We wish this probability to be less than $\frac{1}{20}$, and we have $t = 1$. Plugging in, we get $\frac{10}{n} \leq \frac{1}{20}$. Or $n = 200$.

   That's a lot of time to eat chicken nuggets.

3. Solve the previous problem but now assume you can use the Central Limit Theorem. (*Hint:* You may want to use the table in the back of the exam).
   Answer: 40 hours. Or more precisely $\lceil 10 \cdot 1.96^2 \rceil = 39$ hours. We assume now that $\frac{Y-\lambda}{\sqrt{\lambda/n}}$ follows a normal 0,1 distribution. From the table in the back of the exam, we see that
   $$\Pr\left[-1.96 \leq \frac{Y-\lambda}{\sqrt{\lambda/n}} \leq 1.96\right] \approx 0.95$$
   Therefore, the width of the confidence interval is $2 \cdot 1.96 \cdot \sqrt{\lambda/n}$. Setting this equal to 2 and using the fact that $\lambda \leq 10$ yields $10 \cdot 1.96^2$.

8. Not so dense density functions. 5 points each (sub)part, 15 points total.

1. Consider a continuous random variable whose probability density function is cx^{-3} for $x \geq 1$, and 0 outside this range. What is c?
   Answer: 2. $1 = \int_1^\infty cx^{-3} = -\frac{cx^{-2}}{2}|_1^\infty = \frac{c}{2}$, which implies $c = 2$.
2. Consider random variables X, Y with joint density function $f(x,y) = cxy$ for $x, y \in [0,1]$, and 0 outside that range.
   (a) What is c?
   Answer: 4. $\int_0^1 \int_0^1 cxydxdy = \frac{c}{4} = 1$, so $c = 4$.
   (b) What is $\Pr[|X-Y| \leq 1/2]$?
   Answer: $\frac{41}{48}$. The density function is symmetric around the line $y = x$, so we will only consider the part above $y = x$. Moreover, the integration is a bit easier if one considers the complement or when $|X-Y| \geq 1/2$. This corresponds the $2\int_{1/2}^1 \int_0^{1/2-y} 4xydxdy$, where the factor of 2 is where we use the symmetry. Integrating gives $\frac{7}{48}$. Taking the complement yields $\frac{41}{48}$.

9. This is Absolutely Not Normal! 6 points each part, 12 points total.

Consider a standard Gaussian random variable Z whose PDF is
$$\forall z \in \mathbb{R}, \quad f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}.$$
Define another random variable X such that $X = |Z|$.

(a) Determine a reasonably simple expression for $f_X(x)$, the PDF of X. It may be helpful to draw a plot. Place your final expression in the box below.
Answer: We'll solve this part in two ways.
Method I For $x < 0$, the CDF of X is
$$F_X(x) = \Pr[X \leq x] = 0.$$
For $x \geq 0$ the CDF is given by
$$\begin{aligned} F_X(x) &= \Pr(-x \leq Z \leq x) \\ &= \Phi(x) - \Phi(-x) \\ &= \Phi(x) - \underbrace{[1 - \Phi(x)]}_{\Phi(-x)} \\ &= 2\Phi(x) - 1, \end{aligned}$$

where $\Phi(x) = \int_{-\infty}^{x} f_Z(z)\,dz$ denotes the CDF of the standard Gaussian random variable.

Method II This method is longer for the problem at hand, but it uses mixture probabilities and the Law of Total Probability, so it's worth describing. In particular, we recognize that

$$X = \begin{cases} Z & \text{if } Z \geq 0 \\ -Z & \text{if } Z < 0. \end{cases}$$

Using the Law of Total Probability, we can write the PDF for X as a *mixture PDF*—namely,

$$f_X(x) = f_{X|Z\geq 0}(x)\mathsf{Pr}(Z \geq 0) + f_{X|Z<0}(x)\mathsf{Pr}(Z < 0).$$

We note that $\mathsf{Pr}(Z \geq 0) = \mathsf{Pr}(Z < 0) = 1/2$. We also recognize that for $x < 0$, it must be the case that $f_X(x) = 0$. However, for $x \geq 0$, we have:

$$\begin{aligned} f_X(x) = f_{X|Z\geq 0}(x) &= \frac{f_{X,Z\geq 0}(x)}{\mathsf{Pr}(Z \geq 0)} \\ &= \frac{f_Z(x)}{\mathsf{Pr}(Z \geq 0)} \\ &= \sqrt{\frac{2}{\pi}}e^{-x^2/2}. \end{aligned}$$

Therefore,

$$f_X(x) = \begin{cases} 0 & \text{if } x < 0 \\ \sqrt{\frac{2}{\pi}}e^{-x^2/2} & \text{if } x \geq 0 \end{cases}$$

(b) Determine a reasonably simple expression for $E(X)$, the mean of X. Place your final answer in the box below.

Answer: The mean of X is given by

$$\begin{aligned} \mathsf{E}(X) &= \int_0^\infty x\, f_X(x)\,dx \\ &= \sqrt{\frac{2}{\pi}}\int_0^\infty x e^{-x^2/2}\,dx. \end{aligned}$$

Let $s = x^2/2$, so that $ds = x\,dx$. We then have

$$\mathsf{E}(X) = \sqrt{\frac{2}{\pi}}\int_0^\infty e^{-s}\,ds,$$

which leads to

$$\mathsf{E}(X) = \sqrt{\frac{2}{\pi}}.$$

10. Joint Distributions with Kyle and Lara. 6 points each part, 18 points total.

Kyle and Lara arrive in Saint Petersburg randomly and independently, on any one of the first five (5) days of May 2019. Let K be the day that Kyle arrives, and let L be the day that Lara arrives. (Note that K and L will both be in $\{1,2,3,4,5\}$).
Whoever arrives first must wait for the other to arrive before going on any kind of excursion in the city.

(a) Determine $E[|K-L|]$, the expected wait time in days.

Answer: If K and L arrive in Saint Petersburg randomly then their PMFs are uniform over the set $\{1,2,3,4,5\}$. Since K and L are independent, the joint PMF will be the uniform PMF over the range $k \in \{1,2,3,4,5\}$ and $\ell \in \{1,2,3,4,5\}$:

$$p_K(k) = \begin{cases} \frac{1}{5} & \text{if } k \in \{1,2,3,4,5\} \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad p_L(\ell) = \begin{cases} \frac{1}{5} & \text{if } \ell \in \{1,2,3,4,5\} \\ 0 & \text{otherwise.} \end{cases}$$

$$p_{K,L}(k,\ell) = \begin{cases} \frac{1}{25} & \text{if } k,\ell \in \{1,2,3,4,5\} \\ 0 & \text{otherwise.} \end{cases}$$

We determine the mean wait time $E[|K-L|]$ in two different ways:

Method I

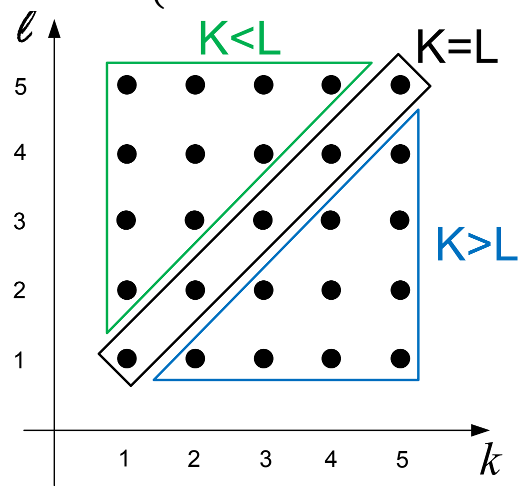
$$\begin{aligned} E[|K-L|] &= \sum_{k=1}^{5}\sum_{\ell=1}^{5}\left(\frac{1}{25}\right)|k-\ell| \\ &= 0\left(\frac{5}{25}\right)+1\left(\frac{8}{25}\right)+2\left(\frac{6}{25}\right)+3\left(\frac{4}{25}\right)+4\left(\frac{2}{25}\right) \\ &= \frac{8}{5} \end{aligned}$$

Method II Let's define three mutually exclusive, collectively exhaustive events

$$K > L, \qquad K < L, \qquad \text{and} \qquad K = L,$$

which are depicted in the figure below:

$$p_{K,L}(k,\ell) = \begin{cases} 1/25 & k,\ell \in \{1,2,\ldots,5\} \\ 0 & \text{elsewhere.} \end{cases}$$



It's clear from the sample space that

$$P(K>L) = P(K<L) = 10/25 = 2/5 \qquad \text{and} \qquad P(K=L) = 5/25 = 1/5.$$

We can then compute the expected value of the wait time by invoking the *Law of Total Expectation* as follows:

$$E[|K-L|] = E\left[|K-L|\middle|A\right]P(A) + E\left[|K-L|\middle|B\right]P(B) + E\left[|K-L|\middle|C\right]P(C).$$

When Event C occurs, $K=L$, so $|K-L|=0$. Furthermore, by symmetry it must be true that

$$E\left[|K-L|\middle|A\right]P(A) = E\left[|K-L|\middle|B\right]P(B).$$

So, $E[|K-L|] = 2E\left[|K-L|\middle|A\right]P(A)$. If event A occurs, $K>L$, we can simply remove the absolute value:

$$\begin{aligned} E[|K-L|] &= 2E\left[K-L\middle|A\right]P(A) \\ &= 2\left[1\cdot\left(\frac{4}{10}\right) + 2\cdot\left(\frac{3}{10}\right) + 3\cdot\left(\frac{2}{10}\right) + 4\cdot\left(\frac{1}{10}\right)\right]\left(\frac{2}{5}\right) = \frac{8}{5}. \end{aligned}$$

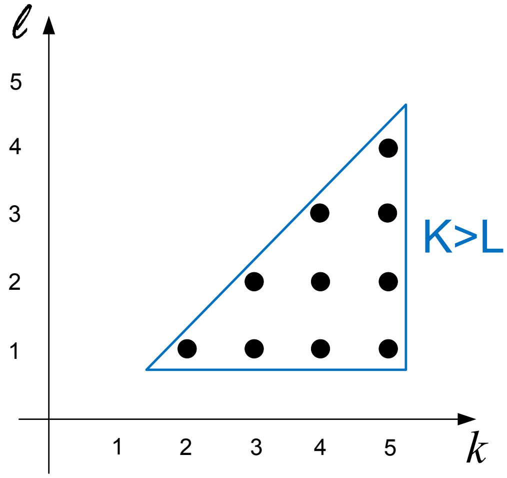
(b) Given that Kyle arrives *at least a day later* than Lara:

(i) Determine the conditional probability mass function for Kyle's arrival day, $p_{K|(K>L)}(k)$

(ii) Provide a well-labeled plot of $p_{K|(K>L)}(k)$.

Answer: Conditioning the sample space on the Event $K>L$ we obtain the following conditional joint PMF of K and L:

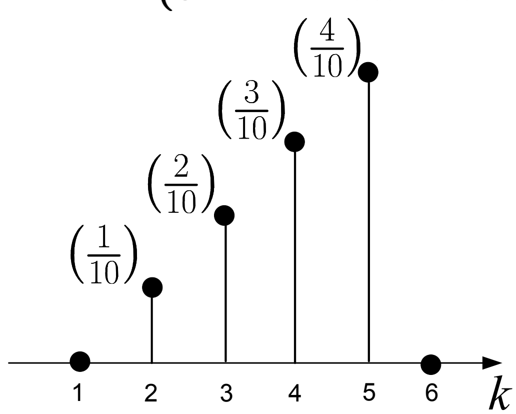
$$p_{K,L|(K>L)}(k,\ell) = \begin{cases} 1/10 & k \in \{2,\dots,5\} \\ & \ell \in \{1,\dots,4\} \\ 0 & \text{elsewhere.} \end{cases}$$



The marginal conditional PMF of K is now obtained by summing the joint conditional PMF values along each column—i.e., along the ℓ direction. In particular,

$$p_{K|(K>L)}(k) = \sum_{\ell} p_{K,L|(K>L)}(k,\ell).$$

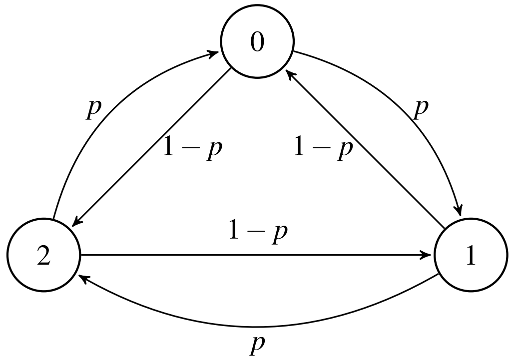
$$p_{K|(K>L)}(k) = \begin{cases} (k-1)/10 & k \in \{2,\dots,5\} \\ \\ 0 & \text{elsewhere.} \end{cases}$$



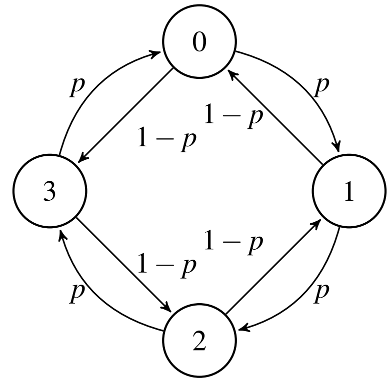
11. Markov Chains 3 points for each part, 18 points total.

Consider the two Markov Chains represented by the following state transition diagrams.

Markov Chain I



Markov Chain II



(a) For Markov Chain I:

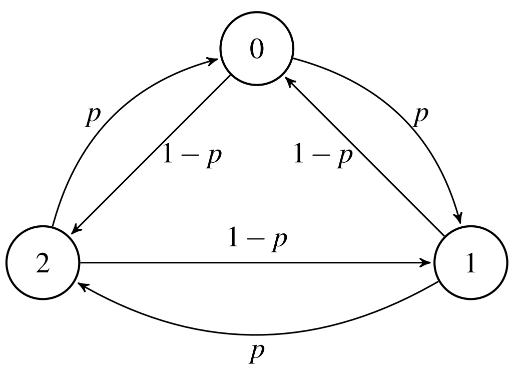
(i) Do the n-step transition probabilities—defined by $r_{ij}(n) = \Pr\left(X_n = j \middle| X_0 = i\right)$ —converge as $n \to \infty$?

Answer: Converges. the chain is aperiodic since the random walker can come back to its original state in either 2 or 3 steps. Thus $r_{ij}(n)$ converges as $n \to \infty$.

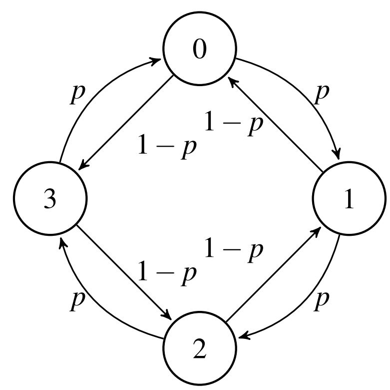
(ii) If so, determine the corresponding limit to which each transition probability converges, and explain whether and why the limit depends on the initial state (i.e., the state at which the walker was stationed initially). If you assert that the transitional probabilities do not converge, explain why no limit exists.

Answer: By structural symmetry, the limiting values are $\frac{1}{3}$ for all i, j. The limit does NOT depend on the initial state. This is because the chain consists of only a single recurrent class.

Markov Chain I



Markov Chain II



(b) For Markov Chain II:

(i) Do the n-step transition probabilities—defined by $r_{ij}(n) = \Pr\left(X_n = j \middle| X_0 = i\right)$ —converge as $n \to \infty$?

Answer: Does not converge

(ii) If so, determine the corresponding limit to which each transition probability converges, and explain whether and why the limit depends on the initial state (i.e., the state at which the walker was stationed initially). If you assert that the transitional probabilities do not converge, explain why no limit exists.

Answer: For the case $m = 4$, the n-step probabilities do not converge since the Markov chain is periodic. To see this, note that the walker can come back to its origin only in an even number of steps.

Alternatively, you can look at the sequence $r_{ii}(n)$, say for $i = 0$: $0, r_{00}(2), 0, r_{00}(4), 0, r_{00}(6), \ldots$. This sequence converges if, and only if, $r_{00}(2k)$ goes to 0 as $k \to \infty$, which is impossible, since the walk cannot oscillate forever between nodes 1, 2, and 3. Thus the n-step transition probabilities do not converge in this case. Yet another way to see the periodicity of the four-node chain is to note that we can group the nodes 0 and 2 into one subgraph, and group the nodes 1 and 3 into another subgraph. Clearly, the graph is bipartite, with each transition occurring across the subgraphs so defined.

(c) (Points) Consider Markov Chain I above. Determine t_0^*, the *mean recurrence time* for State 0.

The mean recurrence time for a state s is the expected number of steps up to the first return to state s, starting from state s. In other words,

$$t_s^* = E\big[\min(n \geq 1 \text{ such that } X_n = s) \mid X_0 = s\big].$$

In particular,

$$t_s^* = 1 + \sum_i p_{si} t_i,$$

where t_i, which denotes the *mean first passage time* from State i to State s, is given by

$$t_i = E\big[\min(n \geq 0 \text{ such that } X_n = s) \mid X_0 = i\big].$$

(i) Write the system of equations that you would solve in the box below. Use t_0^*, t_1, t_2, and p.

(ii) Set p to $1/2$ and write your final answer for the value of t_0^* in the box below.

Answer: For t_0^*, t_1 and t_2, we have the following equations

$$\begin{aligned} t_0^* &= 1 + pt_1 + (1-p)t_2 \\ t_1 &= 1 + pt_2 \\ t_2 &= 1 + (1-p)t_1. \end{aligned}$$

Setting $p = \frac{1}{2}$ yields $t_0^* = 3$.

12. Derive Magic from a Uniform PDF. 5 points per part. 15 points.

A random-number generator produces sample values of a continuous random variable U that is uniformly distributed between 0 and 1.

In this problem you'll explore a method that uses the generated values of U to produce another random variable X that follows a desired probability law distinct from the uniform.

(a) Let $g : \mathbb{R} \to [0,1]$ be a function that satisfies all the properties of a CDF. Furthermore, assume that g is invertible, i.e. for every $y \in (0,1)$ there exists a unique $x \in \mathbb{R}$ such that $g(x) = y$.

Let random variable X be given by $X = g^{-1}(U)$, where g^{-1} denotes the inverse of g. Prove that the CDF of X is $F_X(x) = g(x)$.

Answer: The CDF of X is

$$F_X(x) = \Pr(X \leq x) = \Pr\left[g^{-1}(U) \leq x\right].$$

Since g is strictly increasing, we know that $g^{-1}(U) \leq x$ if, and only if,

$$g\left[g^{-1}(U)\right] = U \leq g(x).$$

Therefore,

$$F_X(x) = \mathsf{Pr}[U \le g(x)] = F_U[g(x)] = g(x).$$

If g is differentiable, we can obtain the PDF of X as well:

$$f_X(x) = \frac{dF_X(x)}{dx} = \frac{dg(x)}{dx}.$$

Upshot: If we want to simulate a random variable X that obeys a desired CDF $F_X(x)$, which is invertible over a range $\mathsf{S} = \{x \mid 0 < g(x) < 1\}$ of interest, we can generate random variable U uniformly distributed in $[0,1)$, and let $X = F_X^{-1}(U)$.

(b) A random variable X follows a *double-exponential* PDF given by

$$\forall x \in \mathbb{R}, \qquad f_X(x) = \frac{\lambda}{2} e^{-\lambda|x|},$$

where $\lambda > 0$ is a fixed parameter.

Using the random-number generator described above (which samples U), we want to generate sample values of X. Derive the explicit function that expresses X in terms of U. In other words, determine the expression on the right-hand side of

$$X = g^{-1}(U).$$

To do this, you must first determine the function g. From part (a) you know that $g(x) = F_X(x)$, so you must first determine $F_X(x)$. It might help you to sketch the PDF of X first. Place your expression for g^{-1} in the box below.

Answer: For $x < 0$,

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt = \frac{\lambda}{2}\int_{-\infty}^{x} e^{\lambda t}\,dt = \frac{1}{2}e^{\lambda x}.$$

For $x \ge 0$,

$$F_X(x) = F_X(0) + \int_0^x f_X(t)\,dt = \frac{1}{2} + \frac{\lambda}{2}\int_0^x e^{-\lambda t}\,dt = \frac{1}{2} + \frac{1}{2}\left[1 - e^{-\lambda x}\right] = 1 - \frac{1}{2}e^{-\lambda x}.$$

That is,

$$g(x) = F_X(x) = \begin{cases} \frac{1}{2}e^{\lambda x} & \text{if } x < 0 \\ 1 - \frac{1}{2}e^{-\lambda x} & \text{if } x \ge 0. \end{cases}$$

To determine $X = g^{-1}(U)$, we consider two ranges of U separately: $0 \le U < 1/2$ and $1/2 \le U < 1$. We do this because for each of these two ranges the CDF $F_X(x)$ takes on distinct functional forms.

If $0 \le U < 1/2$, we let $F_X(X) = \frac{1}{2}e^{\lambda X} = U$. Solving for X, we obtain

$$X = \frac{1}{\lambda}\ln(2U).$$

If $1/2 \le U < 1$, we let $F_X(X) = 1 - \frac{1}{2}e^{-\lambda X} = U$. Solving for X, we obtain

$$X = -\frac{1}{\lambda}\ln[2(1-U)].$$

Accordingly, we generate sample values of X as follows:

$$X = \begin{cases} \dfrac{1}{\lambda}\ln(2U) & \text{if } 0 \le U < 1/2 \\ -\dfrac{1}{\lambda}\ln[2(1-U)] & \text{if } 1/2 \le U < 1. \end{cases}$$

Introduction to Probability, 2nd Ed, by D. Bertsekas and J. Tsitsiklis, Athena Scientific, 2008

	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

The standard normal table. The entries in this table provide the numerical values of $\Phi(y) = \mathbf{P}(Y \le y)$, where Y is a standard normal random variable, for y between 0 and 3.49. For example, to find $\Phi(1.71)$, we look at the row corresponding to 1.7 and the column corresponding to 0.01, so that $\Phi(1.71) = .9564$. When y is negative, the value of $\Phi(y)$ can be found using the formula $\Phi(y) = 1 - \Phi(-y)$.