CS 70 Discrete Mathematics and Probability Theory
Spring 2017 Rao

Final Solutions

1. True/False/Short Answer: Discrete Math (mostly). 3 points by 10 parts.

Please write your answer in the provided box, or bubble in the corresponding option. **This is what is to be graded. No need to justify!**

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

1. (True/False) $A \implies B$ is logically equivalent to $\neg B \implies \neg A$.
   Answer: True. It is the contrapositive.
2. (True/False) $P(0) \wedge (\forall n \in \mathbb{N}, P(n) \implies P(n+1))$ is logically equivalent to $\forall n \in \mathbb{N}, P(n)$.
   Answer: True. It is the principle of induction.
3. (True/False) If $d = gcd(y,z)$ then $gcd(\frac{yz}{d^2}, d) = 1$.
   Answer: False. Example: $gcd(18,12) = 6$ and $(18)(12)/6^2 = (6)$ $gcd(6,6) = 6 \neq 1$.
4. Give an example of x and m where x has no multiplicative inverse $\pmod m$.
   Answer: $x = 2, m = 6$ since $gcd(2,6) = 2 \neq 1$
5. If there is at least one solution to the equation $ax = b \pmod m$ where $d = gcd(a,m)$, how many total solutions in $\{0,\dots,m-1\}$ are there?
   Answer: d. Any solution of the form $x + i\frac{m}{d} \pmod m$. Each of these values is distinct modulo m until $i\frac{m}{d} = m$, or when $i = d$.
6. What is the maximum number of solutions in $\{0,\dots,m-1\}$ for $ax = 1 \pmod m$ for any natural numbers a and m?
   Answer: 1. There is only one multiplicative inverse modulo m.
7. What is $\Delta_1(x)$ in Lagrange interpolation for the points $(1,3)$,$(2,3)$, $(3,4)$ modulo 5, for a degree 2 polynomial? (Factored form is fine.)
   Answer: $\frac{(x-2)(x-3)}{(1-2)(1-3)} = 3(x-2)(x-3)$.
8. Polynomial Related Questions. In the following, recall that a polynomial, $P(x)$, contains a point (a,b) when $P(a) = b$. Two polynomials, $P(x)$ and $Q(x)$, intersect at a point (a,b) when $P(a) = Q(a) = b$.
   (a) Working modulo p where p is prime, what is the maximum number of times a polynomial $P(x)$ of degree exactly $d > 0$ and $d < p$ can take on a value v modulo p?
   Answer: d. Suppose for the sake of contradiction $P(x)$ takes on the value v, $d+1$ times. Then the polynomial $P(x) - v$ takes on the value 0, $d+1$ times, so it must be the zero polynomial, contradicting the fact that $P(x) - v$ must also have a non-zero coefficient for x^d.
   (b) What is the smallest number of degree exactly $0 < d \leq p$ polynomials all modulo p for a prime p, whose product is the zero polynomial. (A product of polynomials $P(x)$, $Q(x)$, $R(x)$, is the function $F(x) = P(x)Q(x)R(x)$. A function is the zero function if $F(x) = 0$, for all x. Your answer may be in terms of p and d.)
   Answer: $\lceil \frac{p}{d} \rceil$. Each polynomial can only have d zeros. The product needs to eventually have p zeros.
9. What is $2^{27} \pmod{15}$?
   Answer: $8 \pmod{15}$. $pq = (3)(5) = 15$, and $(p-1)(q-1) = (2)(4) = 8$ and $27 = (3)(8) + 3$ and finally $2^{(3)(8)+3} = (2^8)^3 2^3 = (1)(8) = 8 \pmod{15}$
10. The RSA scheme only has $m^{ed} = m \pmod N$ for messages m where $gcd(m,N) = 1$. (Recall (N,e) is the public, and d is the private key.) (True/False) **Answer:** False. One always gets back the message.

2. Short Arguments: Mostly Discrete Math. 10/12 points
Provide a clear and concise justification of your answer.

1. Stable Marriage/Simple Proof.

   Consider a stable room-mates problem on $2n$ people where all people have consistently ordered preference lists. That is, if $a > b$ in one list, this is true in every preference list that contains both a and b. Prove or disprove, that there is always a stable pairing. (Recall that a stable pairing of $2n$ people is a partition of the people into n pairs where there are no rogue couples. In this case, a rougue couple is two people not currently paired, who both prefer each other to their current partners).

   Answer: There is always a stable pairing. A proof by induction proceeds by observing that the pair for $n = 2$ is stable. And for larger n, we can pair everyone's favorite with everyone's second favorite, and then recursively make a stable pairing for the $2n-2$ people by induction as the remaining lists still are consistently ordered.

   The two favorites clearly like each other more than all others, so they do not partipate in a rogue couple. The others do not participate in a rogue couple by induction.

2. Colorings/Graphs. Recall a vertex coloring of a graph is an assignement of colors to vertices where the endpoints of each edge are different colors.

   (a) Give an example of a graph with maximum degree 4 that requires 5 colors to be vertex-colored. (No or very brief justification.)
   Answer: A clique with 5 vertices.

   (b) What is the least number of colors needed to color any n-vertex tree? (No or very brief justification.)
   Answer: 2. It is bipartite.

   (c) Prove that for any connected graph with all vertices having degree at most d, that all but one vertex can be colored using only d colors. (That is, the removal of a single vertex leaves a graph that can be colored with d colors.)
   Answer: Consider the method of remove a vertex, recursively color the graph and then replace the vertex and color it differently from the already colored neighbors. As long as the number of colors available is greater than the current degree, d', of the vertex when you replace it, one can color it this vertex with choosing from a set of $d'+1$ colors.
   If a graph is connected. After removing the first vertex, one can always remove a vertex that is incident to a removed vertex. Thus its degree to the other vertices is at most $d-1$ if the original maximum degree was d. Thus, the only vertex that requires more than d colors is the last one.

3. Counting outfits. (10/10) points

This problem should have brief justification.

1. You have 5 shirts, 3 pairs of pants, 4 dresses, and 5 pairs of shoes. Every day you choose to wear either a dress and shoes, or a shirt, pants and shoes.

   (a) How many outfits do you have?
   Answer: There are 4×5 dress/shoe outfits, and $5 \times 3 \times 5$ shirt, pants, shoe combos, for a total of 95.

   (b) If on each day of a 100 day semester, you choose a possible outfit to wear uniformly at random and independently of all other days, what is the expected number of outfits that are worn more than once over the semester? (No need to simplify.)

Answer: The situation corresponds to throwing 100 balls randomly into 95 bins, where a bin is an outfit and a ball is a day. The question asks for the expected number of bins with more than one ball in it.
The probability of a given bin having more than one ball is $1-(\frac{1}{95})(\frac{94}{95})^{99}\binom{100}{1}-(\frac{94}{95}))^{100}\binom{100}{0})$ and by linearity of expectation the answer is $95(1-(\frac{1}{95})(\frac{94}{95})^{99}\binom{100}{1}-(\frac{94}{95})^{100}\binom{100}{0}))$

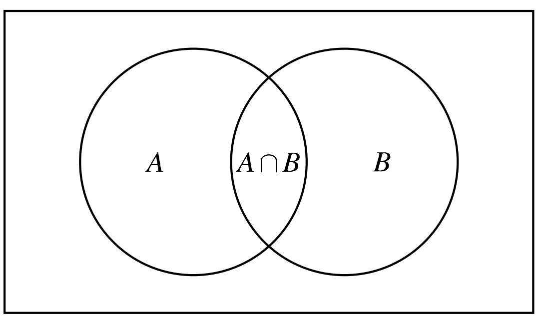
2. You have 5 pairs of socks. Each pair is a different color (and so is distinguishable), but the two socks within a pair are indistinguishable. (Note your left foot, right foot and nose are all distinguishable.)

   (a) How many ways are there for you to choose two socks to wear on your feet?
   Answer: 25. You have 5 choices of what sock to wear on your left foot, and no matter which one you choose, you have 5 choices of what sock to wear on your right foot. Thus, by the first rule of counting, our total number of options is $5*5 = 25$.

   (b) You want to be super cool, so you are also going to wear a third sock on your nose. Now how many ways are there for you to choose socks?
   Answer:
   120. There are two cases to consider: either the socks on your feet match or they do not. If they do, there are 5 choices for what sock to wear on your left foot, 1 choice for what sock to wear on your right foot (since you have to wear the second of the pair), and 4 choices for what sock to wear on your nose (since you have used up an entire one of the pairs). So by the first rule of counting, there are $5*1*4 = 20$ ways to wear socks such that the socks on your feet match.
   We can use a similar procedure to count the number of ways to choose socks such that your feet do not match. There are 5 choices for what sock to wear on your left foot, 4 choices for what sock to wear on your right foot (since you can't choose the same one twice), and 5 choices for what sock to wear on your nose (since you haven't fully used up any pair). So by the first rule of counting, there are $5*4*5 = 100$ ways to wear socks such that your feet do not match. Adding this to the 20 ways we got from before, we get that there are 120 ways to choose socks.

4. Probability Warmup: Zen with Venn. 7 parts. 3 points each.

Consider the diagram, not drawn to scale:



Let x be $Pr[A]$, y be $Pr[B]$ and $z = Pr[A\cap B]$.

In terms of x, y and z, what is

1. $Pr[A|B]$?
   Answer: $\frac{z}{y}$
2. $Pr[B|A]$?
   Answer: $\frac{z}{x}$
3. $Pr[B|\overline{A}]$? **Answer:** $\frac{y-z}{1-x}$
4. $Pr[A\cup B]$? **Answer:** $x+y-z$
5. If A and B are independent, what is z in terms of x and y? (Henceforth, the diagram may be misleading.)
   Answer: $z = xy$
6. If A and B are disjoint, what is z? **Answer:** $z = 0$

7. If A and $\overline{B}$ are disjoint, what is z in terms of x and y? **Answer:** $x = Pr[A] = Pr[A \cap \overline{B}] + Pr[A \cap B] = 0 + z$. Therefore, $z = x$.

5. Probability: Short Answers/True/False. 12 parts. 3 points each.

Please write your answer in the provided box, or bubble in the corresponding option. **This is what is to be graded. No need to justify!**

1. Given a random variable, X, with expectation $\mu = E[X]$, what value a minimizes $E[(X-a)^2]$? (Answer is an expression.)

   Answer: μ.

   The expectation is the value that minimizes the square error.

2. Given random variables, X and Y, with $E[X] = 1$ and $E[Y] = 2$, and $cov(X,Y) = 1$, and $var(X) = 2$, what is the LLSE prediction for Y if $X = 2$?

   Answer: 2.5. LLSE predict the expecion plus the $E[Y] + cov(X,Y)/var(X)(X - E[X])$.

3. (True/False) The covariance of X and Y is always less than the variance of X.

   Answer: False. If $Y = 2X$, we have that $E[XY] - E[X]E[Y] = 2(E[X^2] - E[X]^2) = 2Var(X)$.

4. For independent exponentially distributed random variables, X and Y, both with parameter λ, the covariance of $Z = \min(X,Y)$ and $W = \max(X,Y) - \min(X,Y)$ is positive, negative or zero?

   Answer: Zero. Exponential distributions are memoryless, so we have W being an exponential with parameter λ regardless of the value of Z.

5. For independent exponentially distributed random variables, X and Y, with different parameters λ_X and λ_Y, the covariance of $Z = \min(X,Y)$ and $W = \max(X,Y) - \min(X,Y)$ is positive, negative or zero? **Answer:** Zero. Without loss of generality, we assume X has parameter λ_X and Y has parameter λ_Y We call the event that Z is determined by X, D.

   $$Pr[D|Z \in (z, z+\delta)] = \frac{Pr[X \in (z,z+\delta) \cap Y > z]}{Pr[X \in (z,z+\delta) \cap Y > z] + Pr[X \in (z,z+\delta) \cap Y > z]}$$

   $$= \frac{\lambda_X e^{-\lambda_X z} e^{-\lambda_Y z} \delta}{\lambda_X e^{-\lambda_X z} e^{-\lambda_Y z} \delta + \lambda_Y e^{-\lambda_X z} e^{-\lambda_Y z} \delta} = \frac{\lambda_X}{\lambda_X + \lambda_Y}.$$

   This does not depend on the value of z, thus the distribution of W does not depnd on the value of z. They are independent.

6. (True/False:) Let X, Y, and Z be random variables. Then $\mathbf{E}[(Y - L[Y \mid X])L[Z \mid X]] = 0$. **Answer:** True, by the projection property.

7. (True/False) $\mathbf{E}[(Y - L[Y \mid X])^2] \geq \mathbf{E}[(Y - L[Z \mid X])^2]$. **Answer:** False. $L[Y \mid X]$ is the best linear estimator of Y given X so it should have smaller error.

8. (True/False:) If Z is a linear function of Y and Y is a linear function of X, then $L[Z \mid X] = L[Z \mid L[Y \mid X]]$. **Answer:** True, since $L[Z \mid X] = Z$ and $L[Z \mid L[Y \mid X]] = L[Z \mid Y] = Z$.

9. What is the probability density function for a continuous random variable with $Pr[X \leq x] = 1 - 1/x$, for $x \geq 1$ and $Pr[X \leq x] = 0$, for $x < 1$?

   Answer: $f_X(x) = x^{-2} 1\{x \geq 1\}$

10. What is the Covariance of X and X^3 where X is a uniformly distributed variable on the interval $[0,1]$. ($X \sim U[0,1]$)

    Answer: 3/40.

    $Cov(X) = E[XY] - E[X]E[Y]$

    $E[X] = 1/2$, $E[X^3] = \int_0^1 x^3 dx = \frac{1}{4}$, $E[XY] = \int_0^1 x^4 dx = \frac{1}{5}$.

    $Cov(X) = 1/5 - 1/8 = 3/40$.

11. (Conditional/Wald) In a certain casino game, you either gain one dollar or lose five dollars with equal probability. You roll a fair 6-sided die and play the game that number of times. What is the expected amount of money that you lose?

Answer: 7 dollars.

To find this, we first calculate the expected loss per game. Since we lose five dollars with probability $\frac{1}{2}$ and gain one dollar (which is equivalent to losing negative one dollars) with probability $\frac{1}{2}$, our expected loss is $(\frac{1}{2}*5)+(\frac{1}{2}*(-1))=2$.

Thus, we expect to lose two dollars each time we play the game. So if we play the game c times, our expected loss is $2c$. If we define L to be the random variable representing our total loss and T to be the number of times we play, we have just shown that $\mathbb{E}[L|T]=2T$.

And we can apply the law of iterated expectation to say that $\mathbb{E}[L]=\mathbb{E}[\mathbb{E}[L|T]]=\mathbb{E}[2T]=2\mathbb{E}[T]$.

And we know that the number of times we play has the same distribution as a fair die roll, so $\mathbb{E}[T]=3.5$, meaning that $\mathbb{E}[L]=2*3.5=7$.

12. Let $\Phi(z)$ be the CDF of the standard normal distribution, i.e. $\Phi(z)=\int_{-\infty}^{z}(2\pi)^{-1/2}\exp(-x^2/2)\,dx$. It is known that $\Phi(2)-\Phi(-2)\approx 0.95$. Moreover, the CLT states that when taking n samples from a distribution, that the average, A_n, shifted by its mean and scaled by the standard deviation converges to the standard normal. Assume the CLT holds for n, what is a good upper bound on the probability that A_n is *larger* than the mean of the distribution by 2 standard deviations?

Answer: 2.5%. The probability that it is outside of two standard deviations is bounded by 5%. We divide by 2 as the normal is symmetric.

6. Some confidence! 10 points.

You take n samples $X_1,\cdots,X_n$ from an exponentially distributed variable X with known parameter λ, and calculate the average $A_n=\frac{X_1+\ldots+X_n}{n}$. What value of n do we need to ensure that A_n is within 0.1μ of $\mu=E[X]$ with 95% probability? (You may use Chebyshev or the CLT.)

Answer: The variance of the exponential is $1/\lambda^2$.

Using this fact, we see that $A_n=(X_1+\cdots X_n)/n$, is in $[\mu-4.5\sigma/\sqrt{n},\mu+4.5\sigma/\sqrt{n}]$ with 95% probability.

Thus, we wish $4.5\sigma/\sqrt{n}\leq .1\mu$, taking $\sigma=1/\lambda$ and $\mu=1/\lambda$, we see that $n\geq 45^2$ or $n\geq 2025$ suffices. ($n\geq 2000$ works too as 4.5 is a rounded version of $\sqrt{20}$.

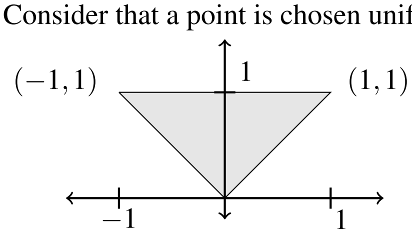
CLT solution.

By the CLT, $A_n\in[\mu-2\sigma/\sqrt{n},\mu+2\sigma/\sqrt{n}]$ with probability 0.95. Since $\sigma=\mu$, we set $2\mu/\sqrt{n}\leq 0.1\mu$ to obtain $\sqrt{n}\geq 20$, or $n\geq 400$.

7. Probability Again: continuous similar to discete. Breathe, just breathe. 12/12/12 points

1. (Continuous Joint distribution: conceptual.)

Consider that a point is chosen uniformly in the area corresponding to the figure below.



Let X be the x-coordinate of the chosen point and Y be the y-coordinate.

(a) What is $Pr[Y > X]$?
Answer: 1. In the shaded region $y \geq x$, the measure of the line is 0.

(b) What is $E[X]$?
Answer: 0. By symmetry.

(c) What is $E[Y]$? **Answer:** $\frac{2}{3}$. If you draw a horizontal line at $Y = y$, the area of the triangle below is $\frac{1}{2}(2y)y = y^2$, so $Pr[Y > y] = 1 - y^2$.
Now, $E[Y] = \int_0^1 Pr[Y > y]dy = (y - \frac{y^3}{3})|_0^1 = \frac{2}{3}$.

(d) What is $E[Y|X]$? (Describe a real valued function whose domain is $[-1,1]$.)
Answer: $\frac{1}{2}(1 + |x|)$.

(e) What is $L[Y|X]$?
Answer: $E[X|Y] = 0$ by symmetry, which implies that the LLSE for X given Y is 0, as it is the MMSE and it is linear. Since the slope is 0, this implies that $cov(X,Y) = 0$, which in turn implies that $L[Y|X] = E[Y] = \frac{2}{3}$.
Alternate Solution:
We can calculate the covariance directly; $cov(X,Y) = E[XY] - E[X]E[Y] = E[XY]$, since $E[X] = 0$. Then, $E[XY] = E[E[XY|Y]] = E[YE[X|Y]]$, and since $E[X|Y] = 0$ by symmetry, then $E[XY] = E[Y \cdot 0] = 0$. Since $cov(X,Y) = 0$, then $L[Y|X] = E[Y] = 2/3$.

2. You pick a real number from the range $[0,1]$ using the uniform distribution. Then Alvin independently picks a real number uniformly at random from the range $[0,2]$.

(a) What is the probability that your two numbers differ by no more than 1?
Answer: 3/4. We can define random variables X for your number and Y for Alvin's and the event D be the pairs (x,y) where x and y differ by more than 1. The joint distribution is the uniform distribution over the rectangle whose corners are $(0,0)$ and $(1,2)$, and the event D is the triangle by $(0,1) - -(0,2) - -(1,2)$. The area of this triangle is $1/2$, and the probability of landing in it is $1/4$ as the density function has value $1/2$ in the rectangle as the area of the rectangle is 2. The question asks for $\overline{D}$ which is then $1 - 1/4 = 3/4$.

(b) You pick a real number from the range $[0,1]$ this time with pdf $f(x) = 2x$. Then Alvin picks a real number uniformly at random from the range $[0,2]$. What is the probability that your two numbers differ by no more than 1?
Answer: $\frac{5}{6}$. Again, let's define some variables: let X be the number you pick, Y be the number Alvin picks, and D be the event that the two numbers differ by no more than 1. We know that $P(D) = \int_0^1 f_X(x)P(D|X = x)dx$. We already have $f_X(x)$, so we just have to find $P(D|X = x)$. But this just amounts to asking what fraction of the interval $[0,2]$ is within 1 of a specific x value. We can see that this proportion is $\frac{1}{2}$ at $x = 0$, 1 at $x = 1$, and increases linearly between those two. Thus, $P(D|X = x) = \frac{1}{2} + \frac{x}{2}$. Plugging this in, we get that $P(D) = \int_0^1 2x(\frac{x+1}{2})dx = \int_0^1 (x^2 + x)dx$. This integral comes out to $\frac{1}{3}x^3 + \frac{1}{2}x^2$ evaluated from 0 to 1, which is just $\frac{1}{3} + \frac{1}{2} = \frac{5}{6}$.
Alternate Solution:
So the bad region is still the same as the previous part, it's the triangle in the upper left corner of the rectangle, only we have to integrate against the density x now (the joint density is $2x \cdot 1/2 = x$). Now since the integrand does not depend on y, we can translate the triangle downwards to have vertices $(0,0),(1,1),(0,1)$. Now the integral is $\int_0^1 \int_x^1 x\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 x(1-x)\,\mathrm{d}x = 1/2 - 1/3 = 1/6$, so the probability of not landing in the bad region is $5/6$.

3. Darts (again.) An ok player hits a circular dartboard of radius 1 centered at $(0,0)$, with uniform probability over the area of the dartboard, a good player has distance from the center that is uniform over $[0,1]$. Say we pick a player who is good with probability 1/2 and ok with probability 1/2 and she throws a dart.

(a) What is the pdf for the random variable, X, corresponding to the distance from the center?
Answer: $f(x) = \frac{1}{2} + \frac{1}{2}2x$.
One way to get this:
$Pr[X \in (x, x+\delta)] = Pr[X \in (x, x+\delta)|good]Pr[good] + Pr[X \in (x, x+\delta)|ok]Pr[ok]$.
and observing that $Pr[X \in (x, x+\delta)|good] = \delta$ since the density function is 1, and that $Pr[X \in (x, x+\delta)|ok] = 2x\delta$ since the density function for distance to center is $2x$ as we observed in class. Divide the δ out gives the density function.

(b) If the dart lands at the point $(\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}})$, what is the probability that the player is good.
Answer: 1/2.
The value of X is $1/2$. Let event A be $X \in (x, x+\delta)$.

$$Pr[good|X \in (x, x+\delta)] = \frac{Pr[A \cap good]}{Pr[A \cap good] + Pr[A \cap ok]} = \frac{\frac{1}{2}\delta}{\frac{1}{2}\delta + \frac{1}{2}(2\frac{1}{2})\delta} = \frac{1}{2}$$

Uh oh. This is 1/2. Sometimes life gives you no guidance!

(c) If that player from part (b) throws another dart, what is the expected distance from the center for this dart.
Answer: $\frac{7}{12}$
Let's call this second toss Y.
$E[Y] = E[Y|good]Pr[good] + E[Y|ok]Pr[ok]$.
In this case $Pr[good] = Pr[ok] = 1/2$ as before.
Thus the expected distance is
$E[Y] = \frac{1}{2}(\int_0^1 x dx + \int_0^1 2x^2 dx)$.
or $E[Y] = 7/12$.

8. It's all about that Chain. True/False Part 1 is 15 points. Part 2 is 12 points.

No justification for True/False.

1. True/False.

(a) (True/False:) A finite Markov chain with a transition matrix where every column sums to 1 is irreducible. **Answer:** False, consider the identity matrix.

(b) (True/False:) A finite Markov chain which is not irreducible does not converge in distribution. **Answer:** (False, consider the chain that transitions from 0 to 1 and remains there forever. The limiting distribution is $\begin{bmatrix} 0 & 1 \end{bmatrix}$.)

(c) (True/False:) If the fraction of time spent in state i is q, and $P(i, j) = 0.3$, then the fraction of time spent in state j is at least $0.3q$. **Answer:** True, consider the balance equation $\pi(j) = 0.3\pi(i) + \cdots \geq 0.3\pi(i)$.

(d) (True/False:) If the fraction of time spent in state i is q and $P(j,i) = 0.3$, then the fraction of time spent in state j is at most $q/0.3$. **Answer:** True. $\pi(i) = 0.3\pi(j) + \cdots \geq 0.3\pi(j), so \pi(j) \leq \pi(i)/0.3$.

(e) (True/False:) A two-state aperiodic Markov chain has a self-loop. **Answer:** True. If there is no self-loop, the Markov chain has period 2.

2. Let n be a positive integer. Take $X_0 = n$, and for each $k \geq 0$, let X_{k+1} have the discrete uniform distribution from 0 to X_k, inclusive. If $X_k = 0$, then it follows that $X_{k+1} = 0$.

(a) The sequence (X_k) is a Markov chain. Describe its state space and its transition probabilities (a diagram would suffice).

Answer: The state space is the naturals from 0 to n. The transition matrix is defined by $P(i,j) = \frac{1}{i+1}$ for $i \geq 0$ and $j \leq i$ and $P(i,j) = 0$ otherwise.

(b) Compute the expected number of steps until the sequence (X_k) first reaches 0. Partial credit will be awarded for a correct linear system of equations. Solve the system for full credit. **Answer:** We write down hitting time equations:

$$
\begin{aligned}
h(0) &= 0 \\
h(1) &= \frac{1}{2}h(0) + \frac{1}{2}h(1) + 1 \\
h(2) &= \frac{1}{3}h(0) + \frac{1}{3}h(1) + \frac{1}{3}h(2) + 1 \\
&\vdots \\
h(k) &= \sum_{j=0}^{k} \frac{1}{k+1} h(j) + 1
\end{aligned}
$$

One way to solve this system is to note that for $k > 1$,

$$h(k) = \frac{k}{k+1}h(k-1) + \frac{1}{k+1}h(k) + \frac{1}{k+1}.$$

Thus we obtain the recursive formula $h(k) = h(k-1) + \frac{1}{k}$. We can compute $h(1) = 2 = 1 + \frac{1}{1}$. In general, $h(k) = 1 + H_k$ where H_k is the k-th harmonic number.