CS 70 Discrete Mathematics and Probability Theory
Summer 2023 Huang, Suzani, and Tausik

Final

1 Good Luck!

a. Nikki is taking a test with 14 problems. She picks one uniformly at random. What is the probability that the problem she picks is Problem 1? **[2 points]**

Answer: $\frac{1}{14}$. Uniform probability, with 14 options.

b. How many ways are there to separate 319 indistinguishable exams into three distinct rooms? **[2 points]**

Answer: $\binom{321}{2}$. Stars and bars – two dividers, 319+2 objects to order.

c. Let P be writing an SID on all pages, and Q be signing the pledge. Suppose we know that $P \implies Q$. If a student didn't sign the pledge, do we know anything else about them? If so, what? **[2 points]**

Answer: They didn't write their SID on all pages. The contrapositive of $P \implies Q$ is $\neg Q \implies \neg P$, meaning not signing the pledge implies they didn't write their SID.

2 Where's My Stable Match?

a. We run propose-and-reject on an n-job and n-candidate stable matching instance. The job-proposing algorithm ends in one day, and the candidate-proposing algorithm also ends in one day. Are the resulting pairings always the same? Why or why not? **[4 points]**

Answer: No, all it means is that all jobs have a different top choice and all candidates also have a different top choice. For example, for any job J_k let its top choice be C_k. For any candidate C_k, let its top choice be J_{n-k}. Both algorithms end in one day, but the resulting pairings are not the same.

b. Consider another n-job and n-candidate stable matching instance where all jobs prefer candidate C_1 the most. Let C_2 be some other candidate. We find there are n stable matchings for this instance. Can C_2 be paired with a different job in all n matchings? Why or why not? **[4 points]**

Answer: No. Suppose C_1 most prefers job J, then (C_1, J) must always be a pairing in the result, otherwise there would be a rogue couple. Thus, there are only $n-1$ jobs with which C_2 can be paired in any stable matching. So, we conclude by the Pigeonhole Principle that if there are n distinct stable matchings, there must be some pair of matchings that match C_2 to the same job.

3 Edgy Questions

a. Consider an n-dimensional hypercube for $n \geq 2$. We modify the hypercube; instead of each vertex being connected to vertices that differ in exactly one bit position, they are connected to vertices that differ in exactly two bit positions. How many edges are in this construction? **[3 points]**

Answer: There are 2^n vertices, which each have $\binom{n}{2}$ connections. The degree is then $2^n\binom{n}{2}$ So the total number of edges is $\frac{2^n\binom{n}{2}}{2}$

b. Let G be any simple, connected graph with 12 vertices. Can all graphs G with 11 edges be colored with two colors? Show why or why not. **[3 points]**

Answer: 11 edges in a 12-vertex graph is $n-1$ edges, so G is a tree. All trees are bipartite, and thus can be colored with two colors.

c. Assume you have a simple, connected, planar graph where each face of the graph is adjacent to at least 6 edges. If the graph has 18 vertices, what is the maximum number of edges it can have? **[4 points]**

Answer: We know that every edge is adjacent to two faces, and each face is adjacent to at least 6 edges. So then we have $6f \leq 2e$, and since the graph is planar we know $v+f=e+2$. So $e=v+f-2$ and $e \leq v+\frac{1}{3}e-2$ meaning $\frac{2}{3}e \leq v-2$ and $e \leq \frac{3}{2}v-\frac{6}{2}$. Plugging in $v=18$, we get $e \leq 27-3$ and so the maximum number of edges is 24.

4 Modular Computation

a. Consider an RSA scheme where $N = pq$ with $p = 7$ and $q = 11$. Can e be 5? Why or why not? **[3 points]**

Answer: No, e has to be coprime to $p-1$ and $q-1$ but $\gcd(q-1, e) = \gcd(10, 5) \neq 1$.

b. Let e be 7. What is the private key k where $0 \leq k \leq 59$? **[3 points]**

Answer: We want to find $e^{-1} \pmod{pq} = 7^{-1} \pmod{60}$. Using Euclid's we have $\gcd(7, 60) = \gcd(7, 4) = \gcd(3, 4) = \gcd(1, 3) = 1$. So, building back up, we get $1 = 3 - 2(1) = 3 - 2(4 - 3) = -2(4) + 3(3) = -2(4) + 3(7 - 4) = -5(4) + 3(7) = -5(60 - 8 \cdot 7) + 3(7) = -5(60) + 43(7)$. This means 7's inverse in $\pmod{60}$ is 43, so the private key is 43.

c. Find an integer $x \pmod{91}$ such that $x \equiv 1 \pmod{7}$ and $x \equiv 4 \pmod{13}$ **[3 points]**.

Answer: We first compute a, b that satisfy $a \equiv 1 \pmod{7}$, $a \equiv 0 \pmod{13}$, $b \equiv 0 \pmod{7}$, $b \equiv 1 \pmod{13}$. $a = 13 \cdot (13^{-1} \pmod{7}) = -13$ as $13^{-1} = (-1)^{-1} = -1 \pmod{7}$. $b = 7 \cdot (7^{-1} \pmod{13}) = 14$ as $7 \cdot 2 = 14 = 1 \pmod{13}$, $7^{-1} = 2 \pmod{13}$. Hence, $x = a + 4b = -13 + 56 = 43$

5 Sumthing

[5 points] Using induction, prove that for all $n \geq 2$,

$$(1^3 - 1^2) + (2^3 - 2^2) + \ldots + (n^3 - n^2) \geq \sum_{x=1}^{n-1} \sum_{y=x+1}^{n} xy$$

(If it helps, the summation expands to:)

$$(1 \cdot 2 + 1 \cdot 3 + \ldots + 1 \cdot n) + (2 \cdot 3 + 2 \cdot 4 + \ldots + 2 \cdot n) + \ldots + ((n-1) \cdot n)$$

Answer: Base Case: When $n = 2$, we get that $1^3 - 1^2 + 2^3 - 2^2 \geq 1 \cdot 2$, which is true. For our induction hypothesis, we can assume that the inequality holds for some $n = k$. For our inductive step, we can now assume that $n = k+1$. We then have

$$1^3 - 1^2 + \ldots + (k+1)^3 - (k+1)^2 \geq 1 \cdot 2 + \ldots + 1 \cdot (k+1) + 2 \cdot (k+1) + \ldots + (k) \cdot (k+1)$$

Using our induction hypothesis, we can reduce this to

$$\begin{aligned}
&(k+1)^3 - (k+1)^2 \geq 1 \cdot (k+1) + 2 \cdot (k+1) + \ldots + (k) \cdot (k+1) \\
\iff\ &(k+1)^2(k+1) - (k+1)^2 \geq \sum_{i=1}^{k} i \cdot (k+1) \\
\iff\ &(k+1)^2(k) \geq (k+1) \sum_{i=1}^{k} i \\
\iff\ &(k+1)^2(k) \geq (k+1) \frac{(k)(k+1)}{2} \\
\iff\ &(k+1)^2(k) \geq \frac{(k+1)^2(k)}{2}
\end{aligned}$$

This inequality will always be true, so we have solved the inductive step, and thus proved the original inequality.

6 Polyinfinite

a. Let P be a degree ≤ 2 polynomial in $GF(13)$ where $P(0) = 1$ and $P(1) = 2$. How many possible values of $P(2)$ are there? **[3 points]**

Answer: 13, since we're working in GF(13) and 3 points define a polynomial of degree 2.

b. Let P be a degree ≤ 2 polynomial in $GF(13)$ where $P(0) = 1$ and $P(1) = 2$. How many possible pairs of $(P(2), P(3))$ are there? **[3 points]**

Answer: Also 13, since three points define a polynomial. So once you set P(2), P(3) is set.

c. Let P_n be the set of all polynomials of degree n with integer coefficients. Note we are not working in a Galois Field. Is P_n finite, countably infinite, or uncountably infinite? Justify your answer. **[4 points]**

Answer: P_n is countably infinite, since we can map each polynomial to its set of coefficients, and the coefficients are all within the integers. So you have a finite Cartesian Product of the integer sets, which is countably infinite.

7 How to Beat Magnus Carlsen

a. Youngmin decides to participate in a chess tournament with a prize of \$100. The probability that Youngmin wins the tournament without cheating is $\frac{1}{4}$. However, Youngmin can pay \$100 to cheat, which increases his probability of winning to $\frac{1}{3}$. The probability Youngmin cheats is $\frac{1}{7}$. What is Youngmin's expected profit? **[4 points]**

Answer: It's 0 in all cases except when he doesn't cheat and wins, or he cheats and loses. So that's $\frac{6}{7} \cdot \frac{1}{4} \cdot 100 + \frac{1}{7} \cdot \frac{2}{3} \cdot -100 = \frac{-200}{21} + \frac{600}{28}$.

b. Given that Youngmin has a net profit of zero dollars after the tournament, what is the probability that he cheated? **[4 points]**

Answer: Let A be the event that Youngmin cheated, and B be the event that Youngmin wins the tournament, and C be the event that Youngmin has a net gain of zero dollars. Then, the probability required is

$$\mathbb{P}[A|C] = \frac{\mathbb{P}[A \cap C]}{\mathbb{P}[C]} = \frac{\mathbb{P}[B|A]\mathbb{P}[A]}{\mathbb{P}[B|A]\mathbb{P}[A] + \mathbb{P}[B^c|A^c]\mathbb{P}[A^c]} = \frac{\frac{1}{21}}{\frac{1}{21} + \frac{9}{14}} = \frac{2}{29}$$

8 Ethical Value

Consider a die with $n+1$ sides, where $n \geq 3$. We play the following game, starting with zero points:
Step 1. Roll the die, add the value of the roll to your score.
Step 2. If the die landed on $n+1$, we end the game.
Step 3. Otherwise, go back to Step 1.

a. Let T be the number of rolls made, including the roll that lands on $n+1$. What is the distribution of T? **[2 points]**

Answer: $T \sim \text{Geom}(\frac{1}{n+1})$ since you're rolling until you get $n+1$.

b. Let X_i be the number of times we roll i for $1 \leq i \leq n$. What is $E[X_1|T]$? As a reminder, T is the number of rolls made, including the roll that lands on $n+1$. **[4 points]**

Answer: Start by computing $\mathbb{E}[X_1|T=t]$. This says that if we have t total rolls, what is the expected number of times we roll a 1? Since we roll uniformly at random, the probability of rolling a 1 in each roll before the last is $\frac{1}{n}$, and the probability of rolling a 1 in the last roll is 0 (since you had to roll n+1). So we have that $\mathbb{E}[X_1|T=t] = \frac{t-1}{n}$. Thus, $\mathbb{E}[X_1|T] = \frac{T-1}{n}$.

c. Find the expected value of your score at the end of the game. You must show your work for full credit. (Hint: Depending on approach, it may be helpful to know that $\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$.) **[4 points]**

Answer:
Let X denote the score for the sum of rolls except the $n+1$th roll. Then, $X = \sum_{i=1}^{n} i \cdot X_i$. We note that by linearity of expectation and symmetry, we can calculate $\mathbb{E}[X|T]$:

$$\mathbb{E}[X|T] = \sum_{i=1}^{n} i \cdot \mathbb{E}[X_i|T] = \frac{n(n+1)}{2}\mathbb{E}[X_1|T] = \frac{n(n+1)}{2}\frac{T-1}{n} = \frac{(n+1)(T-1)}{2}$$

Using $T \sim Geom(\frac{1}{n+1})$ and the law of iterated expectations, we get:

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|T]] = \mathbb{E}[\frac{(n+1)(T-1)}{2}] = \frac{n+1}{2}(\mathbb{E}[T]-1) = \frac{n+1}{2}(n+1-1) = \frac{n(n+1)}{2}$$

Therefore, the answer is $\frac{n(n+1)}{2} + (n+1)$ since we have to add the points for the last roll as well.

Alternatively, we may use Wald's Identity, expressing $X = X_1 + X_2 + \ldots + X_{T-1} + (n+1)$ where X_i is instead the result of the ith dice roll (and the last dice roll has to be $n+1$). Each $\mathbb{E}[X_i] = \frac{n}{2}$ as they are uniform from 1 to n. By Wald's, $\mathbb{E}[X] = \mathbb{E}[X_i]\mathbb{E}[T-1] + (n+1) = \frac{n}{2}(n+1-1) + (n+1) = \frac{n(n+1)}{2} + (n+1)$.

9 Gamer Laptop

Gavin's keyboard is extremely crusty - in fact, some of the keys have stopped working. His keyboard has n i.i.d. keys, each of which has a p probability of working.

a. What is the distribution of the number of keys that work? For the remainder of this problem, let us denote this quantity as X. **[2 points]**

Answer: $X \sim Binom(n, p)$.

b. If $p = 0.6$, use Markov's Inequality to find an upper bound on the probability that at least 70% of Gavin's keys work. **[3 points]**

Answer: $\mathbb{E}(X) = np = 0.6n$. We know X is nonnegative, so we can use Markov's, since you can't have a negative number of keys that work. So $\mathbb{P}\left[X \geq \frac{7n}{10}\right] \leq \frac{\mathbb{E}(X)}{7n/10} = \frac{0.6n \cdot 10}{7n} = \frac{6}{7}$.

c. Certain distributions are symmetric around their mean; that is, such distributions follow $P[X = E[X] + c] = P[X = E[X] - c]$ for all c. Show that X is symmetric if $p = 0.5$. **[3 points]**

Answer: Assume $p = 0.5$. By definition using the fact that X is a binomial, $\mathbb{E}[X] = \frac{n}{2}$.

$$\mathbb{P}\left[X = \frac{n}{2} + c\right] = \binom{n}{n/2+c} 0.5^{n/2+c} 0.5^{n/2-c} = \binom{n}{n/2+c} 0.5^n$$

since $p = 1 - p = 0.5$.

$$\mathbb{P}\left[X = \frac{n}{2} - c\right] = \binom{n}{n/2-c} 0.5^{n/2-c} 0.5^{n/2+c} = \binom{n}{n/2+c} 0.5^n = \binom{n}{n/2-c} 0.5^n = \mathbb{P}\left[X = \frac{n}{2} + c\right]$$

thus X is symmetric.

d. If $p = 0.5$, show that the probability that at least 70% of Gavin's keys work is at most $\frac{25}{8n}$. **[4 points]**

Answer: $\mathbb{P}[X \geq 7n/10] = \mathbb{P}[X \leq 3n/10]$ due to symmetry of the distribution.

$\mathbb{P}[X \geq 7n/10] + \mathbb{P}[X \leq 3n/10] = \mathbb{P}[|X - n/2| \geq 2n/10]$, and thus using the previous equation, $\mathbb{P}[X \geq 7n/10] = \frac{1}{2}\mathbb{P}[|X - n/2| \geq 2n/10]$.

We proceed with Chebyshev's. Note that the variance of X is $n/4$. Thus,

$$\mathbb{P}[|X - n/2| \geq 2n/10] \leq \frac{n/4}{(2n/10)^2} = \frac{25}{4n}.$$

Thus, $\mathbb{P}[X \geq 7n/10] \leq 0.5 \cdot \frac{25}{4n} = \frac{25}{8n}$.

10 The Tortoise and The Hare

We define a program `HaltsBefore`(P_1, P_2, x_1, x_2) that takes in 4 parameters: P_1 and P_2 are programs, and x_1 and x_2 are inputs for P_1 and P_2, respectively. `HaltsBefore` runs both $P_1(x_1)$ and $P_2(x_2)$ at the same time, and returns 0 if $P_1(x_1)$ halts before $P_2(x_2)$ halts or they halt at the same time, 1 if $P_2(x_2)$ halts before $P_1(x_1)$ halts, and 2 if both programs run forever. Show that `HaltsBefore` is uncomputable. **[5 points]**

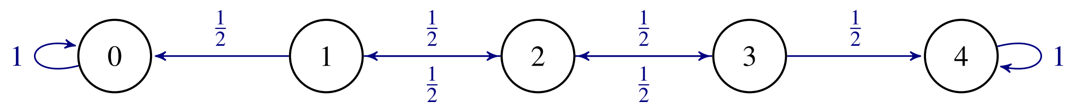
Answer: We can reduce the halting problem to HaltsBefore. We run HaltsBefore(P, G, x, x), where $P(x)$ is the program we are analyzing, and G is a program that loops forever regardless of input. If HaltsBefore returns 0, then $P(x)$ halts. Also, if HaltsBefore returns 2, then $P(x)$ loops forever. Thus, this solves the Halting Problem.

11 Budget Crisis

In order to get funding for the EECS department, CS70 has decided to add a gambling feature to their website (www.eecs70.org/gamble). In this gambling game, you start off with 2 pigeons, and then at every turn, you flip a **fair coin**. If you flip heads, then you get one pigeon. If you flip tails, then you lose one pigeon. You win once you reach 4 pigeons, and lose if you reach 0 pigeons.

a. Draw this process as a Markov Chain. **[3 points]**

Answer:



b. Is the set of invariant distributions for this Markov Chain empty, non-empty finite, countably infinite, or uncountably infinite? Justify your answer. **[4 points]**

Answer: The Markov chain is not irreducible, so we cannot claim there is 1 unique invariant distribution. It turns out any distribution $[\ \alpha\ 0\ 0\ 0\ 1-\alpha]$ for $0 \leq \alpha \leq 1$ is an invariant distribution. Thus, any $\alpha \in \mathbb{R}[0,1]$ is a solution, which is an uncountably infinite set.

c. What is the probability of winning? Justify. (Hint: Try to **not** use first-step equations.) **[2 points]**

Answer: 0.5 by symmetry. Alternatively, the first step equations could be written and solved.

d. To monetize the game, CS70 staff decided that in order to flip the coin, you have to pay 1 dollar for every pigeon you currently have. What is the expected number of dollars you'll have to pay to play a single game to completion (either winning or losing)? Just set up the equations and state which value you're solving for; you **do not need to solve them**. **[4 points]**

Answer: Let's write out the equations. Let D_i be the expected number of dollars to complete the game when you currently have i pigeons. Thus, we start at D_2.

$$
\begin{aligned}
D_0 &= 0 \\
D_1 &= 1 + 0.5D_0 + 0.5D_2 \\
D_2 &= 2 + 0.5D_1 + 0.5D_3 \\
D_3 &= 3 + 0.5D_2 + 0.5D_4 \\
D_4 &= 0
\end{aligned}
$$

Solving gives $D_1 = 5$, $D_2 = 8$, $D_3 = 7$. Thus, the answer is 8.

12 Alphabetical Permutations

Given a uniformly random permutation of the first seven letters (ABCDEFG) of the English alphabet, let X be the number of two-letter substrings where the two letters are alphabetically consecutive.

As an example, for the string ABDCEFG, $X = 3$ since AB, EF and FG are valid, alphabetically consecutive substrings, but BD and DC are not.

a. What is the probability that AB appears in your permutation? Let this answer be p. **[3 points]**

Answer: The probability that AB appears in the permutation is $6 \cdot \frac{1}{7} \cdot \frac{1}{6} = \frac{1}{7}$ – since there are 6 positions you can start from, and the probability of getting the two letters A and B is $\frac{1}{7} \cdot \frac{1}{6}$.

b. What is $\mathbb{E}(X)$? You may express your answer in terms of p, if you wish. **[4 points]**

Answer: Let X_i be an indicator for two consecutive letters in the original string, $X_i = 1$ if the letters are consecutive in the permutation, and 0 otherwise. Then, $X = \sum_{i=1}^{6}(X_i)$, since there are 6 consecutive substrings in the original 7-letter string. The probability that X_i appears in the permutation is $6 \cdot \frac{1}{7} \cdot \frac{1}{6} = \frac{1}{7}$ – since there are 6 positions you can start from, and the probability of getting the two letters is $\frac{1}{7} \cdot \frac{1}{6}$. Thus, $\mathbb{E}(X) = \sum_{i=1}^{6} \mathbb{E}(X_i) = 6 \cdot \frac{1}{7} = \frac{6}{7}$. Using the subpart above, this is $6p$.

c. What is the probability that both AB and CD appear in our permutation? Let this answer be q. **[3 points]**

Answer: $\frac{5!}{7!} = \frac{1}{42}$. We group each pair as "one item" so we have $7 - 4 + 2 = 5$ items to order.

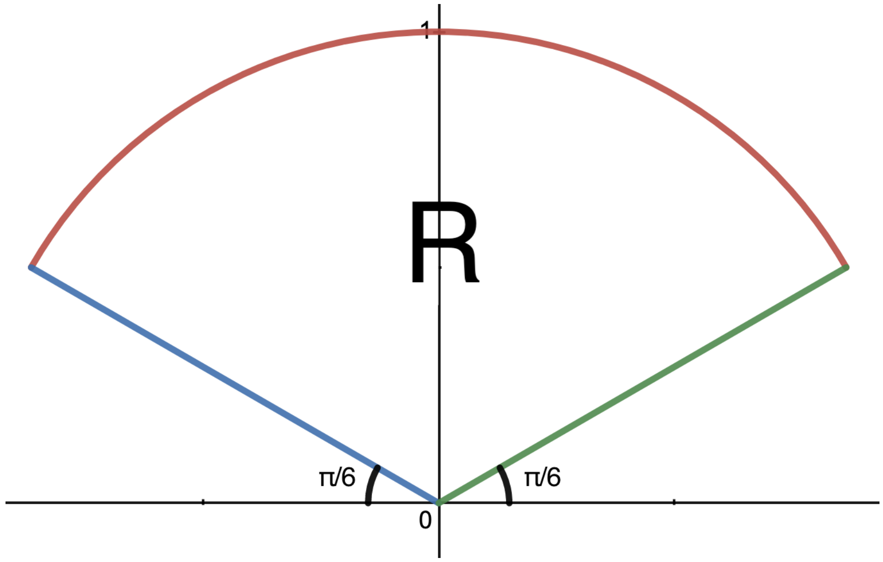
Alternatively, we can use a conditional probability approach. To find the chance that both AB and CD appear, $\mathbb{P}[AB \cap CD]$, we can find $\mathbb{P}[AB]\,\mathbb{P}[CD|AB]$. We already know, from part a, that $\mathbb{P}[AB] = \frac{1}{7}$, or p. To find $\mathbb{P}[CD|AB]$, we need to consider two cases. Case 1: C has 4 possible starting locations with an open slot after it (because AB has been placed at the start or end of the permutation), or Case 2: C has 3 possible starting locations with an open slot after it (because AB is taking up space in the middle). There are 6 possible locations for AB, and 2 are the start and end. Thus, given that AB appears, Case 1 has a $\frac{2}{6}$ chance of occurring, and Case 2 has a $\frac{4}{6}$ chance of occurring. The chance that CD appears in each of these cases can be found similarly to how we found p in part a: in Case 1, it is $4 \cdot \frac{1}{5} \cdot \frac{1}{4} = \frac{1}{5}$; in Case 2, it is $3 \cdot \frac{1}{5} \cdot \frac{1}{4} = \frac{3}{20}$. Thus, our final answer is $\mathbb{P}[AB]\,\mathbb{P}[CD|AB] = p(\frac{2}{6} \cdot \frac{1}{5} + \frac{4}{6} \cdot \frac{3}{20}) = p(\frac{8+12}{120}) = \frac{1}{6}p = \frac{1}{42}$.

d. What is $\text{Var}(X)$? You may express your answer in terms of p and/or q, if you wish. **[4 points]**

Answer:
We know $Var(X_1 + ... + X_6) = Cov(X_1 + ... + X_6, X_1 + ... + X_6) = 6Var(X_i) + 6 \cdot 5(X_i, X_j)$ for $i \neq j$. For $Var(X_i) = p(1-p) = \frac{6}{7}(1 - \frac{6}{7}) = \frac{36}{49}$. For the second term, there are two cases to consider – either X_i and X_j share a letter or they don't. If they share a letter, $\mathbb{E}[X_iX_j] = \mathbb{P}[X_i = 1 \cap X_j = 1]$ which is the probability the three letters appear in sequence, or $\frac{5!}{7!}$. Thus, $Cov(X_i, X_j) = \frac{1}{42} - \frac{1}{7} = \frac{-5}{42}$. If they don't share a letter, then $\mathbb{E}[X_iX_j] = \mathbb{P}[X_i = 1 \cap X_j = 1] = \frac{5!}{7!}$, since we can consider concatenating the two groups of letters (i.e. treat AB as a single letter and CD as a single letter), thus there are 5 "letters" left with 5! permutations. Then, $Cov(X_i, X_j) = \frac{1}{42} - \frac{1}{7} = \frac{-5}{42}$. So we have $6Var(X_i) + 6 \cdot 5 \cdot \frac{-5}{42} = \frac{6 \cdot 36}{49} + \frac{-150}{42} = \frac{1296 - 1050}{294} = \frac{246}{294}$. Using the previous subpart, this is equivalent to $6p(1-p) + 6 \cdot 5(q - p^2)$. To simplify, it's $6p + 30q - 36p^2$.

13 Topping Pizzas

Consider a pizza slice represented by the region R of the xy-plane below, consisting of an arc from a circle of radius 1, and two lines passing through the origin at angles of $\frac{\pi}{6}$ radians (30°).



Suppose Nate throws a pepperoni slice (which we represent as a single point) uniformly onto R, so the PDF for the location of the pepperoni is

$$f(x,y) = \begin{cases} c, & (x,y) \in R \\ 0, & (x,y) \notin R \end{cases}$$

a. Solve for the value of c. **[3 points]**

   Answer: The area of the unit semicircle is $\frac{\pi}{2}$. The region R is only $\frac{2}{3}$ of the semicircle, so the area is $\frac{\pi}{3}$. Thus, $c = \frac{3}{\pi}$.

b. Let X be a random variable representing the x-coordinate of the pepperoni's position. Find $\mathbb{E}[X]$. **[3 points]**

   Answer: This pizza slice is symmetric about the y-axis and the distribution is uniform, so $\mathbb{E}[X] = 0$.

c. In the region where $X = 0$, what is the variance of Y, where Y is the random variable representing the y-coordinate of the pepperoni? **[3 points]**

   Answer: Y is uniform along a line from 0 to 1, so $Var(Y) = Var(U(0,1)) = \frac{1}{12}$.

d. Let $A \sim \text{Exp}(\lambda)$ and $B \sim \text{Exp}(\lambda)$ be two independent random variables. Let $C = \min\{4A, 2B\}$. What distribution does C follow? **[4 points]**

   Answer: $\mathbb{P}[C > k] = \mathbb{P}[4A > k \cap 2B > k] = \mathbb{P}[4A > k]\,\mathbb{P}[2B > k] = \mathbb{P}\left[A > \frac{k}{4}\right]\mathbb{P}\left[B > \frac{k}{2}\right] = e^{-\lambda k/4 - \lambda k/2} = e^{-3\lambda k/4}$. So $C \sim Exp(\frac{3}{4}\lambda)$

14 Kiwi Basil Green Tea with Lychee Jelly and Extra Ice

Nikki wants to estimate how much boba she buys on average per week; we'll call this number λ. Assume that the number of drinks she buys is independent across weeks, and that it is distributed according to Poisson(λ).

a. Nikki guesses that $\lambda = 5$. If this is true, what's the probability that she buys exactly 7 drinks within **two** weeks? **[2 points]**

Answer: Since one week is distributed according to Poisson(5), two weeks are distributed according to Poisson(10). So this is the probability that a Poisson(10) variable equals 7, which is $e^{-10}\frac{10^7}{7!}$.

b. Nikki wants to make a better guess for λ. Let X_i be the number of drinks she buys on week i where each $X_i \sim Poisson(\lambda)$. She tracks the number of drinks she buys for n weeks to get n i.i.d. variables $X_1, \ldots X_n$. Let $\hat{\lambda} = \frac{X_1+\ldots+X_n}{n}$. What is E$[\hat{\lambda}]$? **[3 points]**

Answer: $\mathbb{E}[\hat{\lambda}] = \frac{1}{n}(\mathbb{E}[X_1] + \ldots + \mathbb{E}[X_n]) = \frac{1}{n}(n\mathbb{E}[X_1]) = \mathbb{E}[X_1] = \lambda$.

c. What is $Var(\hat{\lambda})$? **[3 points]**

Answer: $Var(\hat{\lambda}) = \frac{1}{n^2}Var(X_1 + \ldots + X_n)$. Since they're independent, this equals $\frac{1}{n^2}(Var(X_1) + \ldots + Var(X_n)) = \frac{1}{n^2}(nVar(X_1))$ since they're identically distributed. So this is $\frac{\lambda}{n}$.

d. What is the approximate probability that Nikki's estimate is at least 0.02 away from λ? For this subpart, assume n is large, and express your answer in terms of n, λ and Φ, the CDF of the standard normal distribution. **[5 points]**

Answer: We want to find $\mathbb{P}\left[|\hat{\lambda} - \lambda| \geq 0.02\right]$. Note that $\hat{\lambda} - \lambda \sim N(0, \frac{\lambda}{n})$ by CLT. Since normal distributions are symmetric, $\mathbb{P}\left[|\hat{\lambda} - \lambda| \geq 0.02\right] = 2\mathbb{P}\left[\hat{\lambda} - \lambda \geq 0.02\right] = 2 - 2\mathbb{P}\left[\hat{\lambda} - \lambda \leq 0.02\right] = 2 - 2\Phi(\frac{0.02-0}{\sqrt{\lambda/n}})$.

15 Goofy Gaussians

Let $X, Z \sim \mathcal{N}(0,1)$ be independent and identically distributed. Consider the new random variable $Y = \rho X + \sqrt{1-\rho^2}Z$ for $0 < \rho < 1$.

a. What is the distribution of Y? **[3 points]**

Answer: The sum of two independent normals is still normal, so

$$Y \sim \mathcal{N}(0,\rho^2) + \mathcal{N}(0,1-\rho^2) \sim \mathcal{N}(0,1).$$

b. What is $\text{Cov}(X,Y)$? **[3 points]**

Answer: By independence of X and Z, we have:

$$\text{Cov}(X,Y) = \text{Cov}(X,\rho X) + \text{Cov}(X, \sqrt{1-\rho^2}Z) = \rho \text{Var}(X) = \rho$$

c. What is the distribution of $\alpha X + \beta Y$? **[4 points]**

Answer:

$$\begin{aligned}\alpha X + \beta Y &= \alpha X + \beta \rho X + \beta \sqrt{1-\rho^2}Z = (\alpha + \beta\rho)X + \beta\sqrt{1-\rho^2}Z \\ &\sim \mathcal{N}(0,(\alpha+\beta\rho)^2) + \mathcal{N}(0,\beta^2(1-\rho^2)) = \mathcal{N}(0,\alpha^2 + 2\alpha\beta\rho + \beta^2)\end{aligned}$$

d. What is $\mathbb{P}(X > Y + 1)$? You may express your answer in terms of $\Phi(\cdot)$, the standard normal CDF. (Hint: Use the previous part). **[4 points]**

Answer: This is the same as $\mathbb{P}(X - Y > 1)$. Since $X - Y \sim \mathcal{N}(0, 2-2\rho)$ by the previous part, this is:

$$\mathbb{P}(\mathcal{N}(0,2-2\rho) > 1) = 1 - \mathbb{P}(\mathcal{N}(0,2-2\rho) \leq 1) = 1 - \mathbb{P}\left(\mathcal{N}(0,1) \leq \frac{1}{\sqrt{2-2\rho}}\right) = 1 - \Phi\left(\frac{1}{\sqrt{2-2\rho}}\right)$$

e. Draw your favorite video game! Are esports sports? Debate below. **[0 points]**

Answer: League of Legends. Yes.