CS 70 Discrete Mathematics and Probability Theory
Spring 2023 Ayazifar 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.

SID: ______

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. Other stuff.

1. False $\implies$ False.

   Answer: True. The implication is true, if the if condition is False.
2. False $\implies$ True.

   Answer: True. The implication is true, if the if condition is False.
3. $(\neg(\exists x \in S)(P(x))) \implies (\forall x \in S)(P(x) \implies Q(x))$

   Answer: Always True. P(x) is always false, so the implication is always true.
4. $(\forall x \in S)(\exists y \in S)(Q(x,y) \implies P(x)) \equiv (\forall x \in S)(\neg P(x) \implies \neg(\forall y \in S)(Q(x,y)))$

   Answer: Always True. $(\forall x \in S, \exists y \in S, Q(x,y) \implies P(x)) \equiv (\forall x \in S, \exists y \in S, \neg P(x) \implies \neg Q(x,y)) \equiv (\forall x \in S, (\neg P(x) \implies \exists y \in S, \neg Q(x,y)) \equiv (\forall x \in S, (\neg P(x) \implies \neg \forall y \in S, Q(x,y))$.
5. In any stable matching instance, there is no stable pairing where every job gets its least favorite candidate.

   Answer: False. Each job could have a different least favorite candidate whose favorite candidate is the corresponding job. Thus, the pairing that pairs each candidate with its favorite job is candidate optimal and stable.
6. If the propose-and-reject algorithm with jobs proposing terminates on the first day, then the matching is both job-optimal and candidate-optimal.

   Answer: False. If the favorite candidates of the jobs are all separate, then the candidates preference list is not considered at all and thus one can make preferences lists for the candidate that give a different candidate optimal solution.

3. Quick proofs.

1. Prove or disprove: For $x, y \in \mathbb{N}$ where $x \leq y$, then either $x \leq y/2$ or $y - x \leq y/2$.

   Answer: True. For the sake of contradiction assume that $x > y/2$ and $y - x > y/2$ then $x + y - x = y > y$ which is a contradiction.

2. Consider the equation $b^2 = 3a^2$.

   (a) What is the number of solutions for the equation for $a, b \in \mathbb{N}$? (Your answer should possibly be a natural number or infinity.)

   Answer: 1.

   (b) (5 points) Give a proof for your answer.

   Answer: a^2 and b^2 can only have even prime powers in their factorization, yet b^2 must have an odd power of 3 in its factorization. Hence, the only solution is when $a = b = 0$.
   Another solution: A solution must have the property that $\frac{a}{b} = \sqrt{3}$, which contradicts that $\sqrt{3}$ is irrational. Similarly, when $b = a\sqrt{3}$ where a is a natural number implies that b is not a natural number.

4. Mods.

1. If $f(x) = ax \pmod{p}$ where $a \neq 0$, p is prime, and $\gcd(a, p) = 1$, what is the size of the following set:

$$\{f(0) \bmod p, \ldots, f(p-1) \bmod p\}$$

   (Answer could be in terms of p.)

   Answer: p. The function is a bijection since a has a multiplicative inverse modulo p.

2. If $f(x) = ax \pmod{N}$ where $N > 1$ and $\gcd(a, N) = d$, what is the size of the following set:

$$\{f(0) \bmod N, \ldots, f(N-1) \bmod N\}$$

   (Answer could be in terms of a, d and N.)

   Answer: N/d. It is all possible multiples of d in $\{0, \ldots, N-1\}$.

3. What is the product of the numbers 1^2, 2^2, $\ldots$, $(p-1)^2$ modulo p for a prime p? (Answer could be a number or an expression involving p.)

   Answer: 1. Since each $a \neq 0 \pmod{p}$ has a multiplicative inverse, each number and its inverse appears once in the product and thus the product is $1 \pmod{p}$.

5. Quick Bayes.

(5 points) I pick one of two dice with equal probability: one die is a tetrahedron with 1, 2, 3, 4 on the four sides, and one die is a six sided die with 1, 2, 3, 4, 5, 6 on the sides. When rolling a die, all sides are equally likely.

Suppose I roll a 3. What is the probability that I rolled the tetrahedron? (Show work if desired, but clearly indicate final answer.)

Answer: From ChatGPTPlus:

Let A be the event that we choose the tetrahedron die, and let B be the event that we roll a 3. We want to find P(A|B), the probability that we chose the tetrahedron die given that we rolled a 3. By Bayes' Theorem, we have:

P(A|B) = P(B|A) * P(A) / P(B)

where P(B|A) is the probability of rolling a 3 given that we chose the tetrahedron die, P(A) is the prior probability of choosing the tetrahedron die, and P(B) is the probability of rolling a 3 regardless of which die we chose.

We can calculate these probabilities as follows:

P(B|A) = 1/4, since the tetrahedron die has one face with a 3.

P(A) = 1/2, since we choose between the two dice with equal probability.

P(B) = P(B|A) * P(A) + P(B|A') * P(A') = (1/4) * (1/2) + (1/6) * (1/2) = 5/24, where A' is the complement of A (i.e., the event that we chose the six-sided die).

Therefore, we have:

P(A|B) = P(B|A) * P(A) / P(B) = (1/4) * (1/2) / (5/24) = 3/5.

Thus, the probability that we chose the tetrahedron die given that we rolled a 3 is 3/5 or 0.6.

6. Graphs with friends.

Consider a process with $n \geq 7$ people where each person i arrives and chooses $\min(i,7)$ different people from $\{0,\dots,i-1\}$ to be friends with (and everyone accepts those friendships). Consider the resulting (undirected) graph G, where vertices correspond to the n people, and edges correspond to the friendships among these n people.

1. What is the total number of edges in the graph? (Answer could be a number or an expression involving n.)

   Answer: $7(n-7)+\binom{7}{2}$. The first 7 form a complete graph. After that, each person introduces 7 new edges.

2. What is the maximum possible degree of any vertex in any such graph G? (Answer could be a number or an expression involving n.)

   Answer: $n-1$. Everyone could choose the first person to be friends with.

3. (a) What is the minimum number of colors required to vertex color any n vertex graph that results from this process? (Answer could be a number or an expression involving n.)

   Answer: 8. See next answer

   (b) (5 points) Give a proof for your answer to the previous part.

   Answer: For tightness, we argue that we cannot have less than 8 colors. This is seen via contradiction since person 0 must be connected to person 1, 2, ..., 7 and so less than 8 colors does not suffice. Assume the graph on the first i people is 8 colorable, then when adding the ith person, one has at most 7 neighbors which use 7 colors, and one of the 8 colors is available to color this person with in a manner that ensures all edges incident to previous people are multicolored. The base case is person 0 can be colored.

4. If each person i chooses a random subset of $\min(7,i)$ previous people to be friends with, what is the expected degree of person 0? (Answer is in terms of n, and uses a summation.)

   Answer: $\sum_{i=1}^{n} \frac{\min(i,7)}{i}$. The probability that person i chooses person 0 as a neighbor is $\min(i,7)/i$ since it chooses 7 out of the i possibilities. For what it's worth, this is $O(\ln n)$.

7. More Graphs

1. A hypercube of dimension 3 has more edges than K_4.

   Answer: True. The hypercube has $3\times 2^2 = 12$ edges and the K_4 has 6 edges.

2. What is the maximum number of vertices in a graph with e edges and c connected components? (Answer could be in terms of e and/or c.)

   Answer: $e+c$. Consider adding the e edges to n disconnected vertices. Each edge reduces the number of connected components by at most 1, thus $c \geq n-e$ and $n \leq e+c$. One can achieve this bound by always introducing the edges to decrease the number of components by 1 in each step.

3. The number of odd degree vertices in any graph is odd.

   Answer: False. The sum of the degrees of the vertices in a graph is $2e$ and is thus even and thus only an even number of them can be odd.

4. What is the average degree of an n vertex tree? (The answer should be in terms of n and should be exact.)

   Answer: $2-2/n$. The number of edges is $n-1$, and thus the sum of degrees is $2n-2$ and divide by n.

5. (5 points) For a planar graph on v vertices with a planar drawing where the average number of sides on each face is s, what is the total number of edges? (Recall that Euler's formula is $v+f=e+2$. Answer should be in terms of only v and s and not use f.)

   Answer: $\frac{s}{s-2}(v-2)$. The sum of the face sizes is fs and is also $2e$ as each edge is incident to two faces. Thus, we have $v+2e/s=e+2$ or that $e=(v-2)/(1-2/s)=\frac{s}{s-2}(v-2)$.

8. Polynomials.

1. Consider a polynomial $P(x)$ of degree 2 under arithmetic modulo 7, where

$$\begin{aligned} P(0) &\equiv 0 \pmod 7 \\ P(1) &\equiv 2 \pmod 7 \\ P(6) \equiv P(-1) &\equiv 2 \pmod 7 \end{aligned}$$

   (a) If $P(x)=(x-6)Q(x)+r$, what is r?
   Answer: $r=2$, $P(6)=(6-6)Q(x)+r=2$ and thus $r=2$.

   (b) What is $P(x)$?
   Answer: $2x^2 \pmod 7$. First point tells you constant coefficent is 0, and has form ax^2+bx. The second says $a+b=2 \pmod 7$, the last says $a-b=2 \pmod 7$, thus $a=2$ and $b=0$.

   (c) Suppose a degree 0 polynomial, $Q(x)$, differs from $P(x)$ on one of the given points. What is $Q(x)$?
   Answer: $Q(x)=2$. The original polynomial is a constant and thus at least two of the points have the value of the polynomial.

2. Given a polynomial $P(x)$ of degree 2 with $P(0)=0$, $P(1)=P(-1)=A$, what is $P(x)$ in terms of A?

   Answer: Ax^2. The constant coefficient is 0, and thus it has form $a_2x^2+a_1x$, and $a_2+a_1=A$ and $a_2-a_1=A$, and thus $a_2=A$ and $a_1=0$.

3. Consider a polynomial $P(x)$ of degree d modulo a prime p. How many polynomials $Q(x)$ of degree at most d satisfy $P(1)=Q(1)$, $P(2)=Q(2)$, $\ldots$, $P(k)=Q(k)$? (You may assume $k \leq d$ and that $p \geq d+1$ and $P(x)$ itself should be counted.)

   Answer: p^{d+1-k}. $d+1$ points specify a polynomial, so you need to choose values for $d+1-k$ points.

9. Countability/Computability.

1. For any two distinct real numbers in $[0,1]$, there is a rational in between them.

   Answer: True. The real numbers must differ at some bit in the representation. The rational number can agree with the smaller one up to that point and then differ while still being smaller than the bigger one.

2. The cardinality of the rational numbers is equal to the cardinality of the set of pairs of real numbers in $[0,1]$.

   Answer: False. The real numbers and pairs of real numbers are uncountable, the rationals are countable.

3. The number of computer programs is countable.

   Answer: True. A computer program is a finite text string. The set of finite text strings are countable.

4. The number of outputs for any computer program on any finite length input is countable. (Note that the output of a computer program could be an infinite sequence of digits, for example, a square root program could run forever while printing the digits of $\sqrt{2}$.)

   Answer: True. The output for a computer program and an input can be specified by the program and the input, and thus the number of such outputs is countable.

5. There is a program P such that on inputs Q and x, $P(Q,x)$ halts if and only if the program Q does not halt on input x.

   Answer: False. This is the program Turing that we discussed in class. It is different from every program as running Turing on Turing would yield a contradiction.

10. Counting: Warmup

1. What is the number of 2 element subsets of a set of n distinct items?

   Answer: $\binom{n}{2}$. There are $n(n-1)$ ordered pairs by the first rule of counting and then you divide out the 2 possible pairs that are equivalent up to ordering.

2. What is the number of non-empty subsets of a set of n distinct items?

   Answer: $2^n - 1$. There are 2^n subsets of n items and one empty subset.

3. What is the number of orderings of k distinct items?

   Answer: $k!$. The first rule of counting proceeds by noting there are k choices for the first item and $k-1$ choices for the second, and so on.

4. How many Hamiltonian paths are there on K_n?

   Answer: $n!$. We have n choices for the vertex, $n-1$ choices for the second vertex, etc.

5. How many Hamiltonian cycles are there on K_n?

   Answer: $(n-1)!$. Same as previous part, but divide by n to account for different starting points yielding the same Hamiltonian cycle.

11. Stirling numbers.

Stirling numbers are numbers $S(n,k)$ which denote the number of partitions of an n-element set into k non-empty subsets. For example, $S(4,2) = 7$, because there are 7 different partitions of a 4-element set $\{1,2,3,4\}$ into 2 subsets:

$$\{1\}\{2,3,4\} \quad \{1,2\}\{3,4\} \quad \{4\}\{1,2,3\} \quad \{1,4\}\{2,3\} \quad \{2\}\{1,3,4\} \quad \{1,3\}\{2,4\} \quad \{3\}\{1,2,4\}$$

Note that the partition order does not matter; that is, $\{1,2\}\{3,4\}$ is equivalent to $\{3,4\}\{1,2\}$.

1. Find a formula for $S(n,n-1)$.

   Answer: If we partition a set of n elements into $n-1$ non-empty subsets, there must be exactly one subset with exactly 2 elements, and the rest of the subsets must only have 1 element. There are $\binom{n}{2}$ ways to choose the 2-element subset, and the rest of the subsets are fixed; this means that $S(n,n-1) = \boxed{\binom{n}{2}}$.

2. Find a formula for $S(n,2)$.

   Answer: Here, we're partitioning a set of n elements into 2 subsets. Suppose we label the subsets as subset A and subset B. Each element in the set has two choices: it can go in A or in B. This gives us 2^n possibilities.

   However, there are a few cases we must also consider. First, we must exclude the partitions that make either A or B empty, as we must not have any empty subsets. There are only two cases; either A is empty, or B is empty. This leaves $2^n - 2$ possibilities.

   Secondly, the subsets are actually supposed to be indistinguishable; this means that the partition $\{1,2\}\{3,4\}$ is exactly the same as the partition $\{3,4\}\{1,2\}$. To account for this, we can just divide by 2, accounting for swapping A and B. This leaves $2^{n-1} - 1$ possibilities.

   As such, we have that $S(n,2) = \boxed{2^{n-1} - 1}$.

3. (10 points) Prove that $S(n+1,k) = k \cdot S(n,k) + S(n,k-1)$ using a combinatorial proof.

   Answer: Looking at the LHS, we are partitioning a set of $n+1$ elements into k non-empty subsets.

   Suppose we look at the first element of the set; where does it go? There are two cases; it goes in a singleton set by itself, or it goes in a larger subset.

   If the first element goes into a subset by itself, then the number of partitions of the remaining elements is $S(n,k-1)$.

   If the first element goes into a larger subset, there are $S(n,k)$ partitions of the remaining elements, and we can choose any one of these k subsets to put the first element into, giving us $k \cdot S(n,k)$ possibilities.

   Adding these cases up, we have $S(n+1,k) = k \cdot S(n,k) + S(n,k-1)$, as desired.

4. (5 points) How many ways can you select n numbers *without order* and with repetition from the set $\{1,2,\ldots,m\}$, such that there are exactly k distinct values among the numbers selected? *Hint: this does not involve Stirling numbers.*

   Answer: Since we are selecting without order, suppose WLOG that the numbers are sorted. This means that we can group the numbers by value through stars and bars.

   We must have *exactly* k distinct values, so none of the groups can be empty. This means that out of our n numbers, k of them are already fixed in value, leaving $n-k$ numbers to group by value. As such, we have $k-1$ bars, separating our $n-k$ stars. This means that there are a total of $\binom{n-1}{k-1} = \binom{n-1}{n-k}$ different ways to group the n numbers by value.

   We must also choose which k values we have, for a total of $\binom{m}{k}$ possibilities. Putting this together, we have a total of $\boxed{\binom{n-1}{k-1}\binom{m}{k} = \binom{n-1}{n-k}\binom{m}{k}}$ different ways to select n numbers without order and with repetition such that there are exactly k distinct values.

5. (5 points) How many ways can you select n numbers *with order* and with repetition from the set $\{1,2,\ldots,m\}$, such that there are exactly k distinct values among the numbers selected? Express your answer in terms of the Stirling numbers.

   Answer: Firstly, we must choose what the k distinct values should be; there are $\binom{m}{k}$ possible choices here, from m total values.

Next, we can group these distinct values through their indices; that is, we can partition the indices such that each subset of indices points to numbers with the same value. Notice that this is equivalent to creating a partition of the n indices into k different subsets of indices.

For example, if our selected numbers are $(1,3,2,1,1,3)$, then we can group by value through a partition of the indices (starting at 1) as $\{1,4,5\}\{3\}\{2,6\}$. Specifically, the 1st, 4th, and 5th numbers are all 1, the 3rd number is a 2, and the 2nd and 6th numbers are both 3.

We know that the total number of partitions of n elements into k non-empty subsets is equal to $S(n,k)$. However, in this scenario, the subsets are distinguishable; they each represent a single value in $\{1,2,\ldots,m\}$. To account for this, we need to multiply by $k!$ to allow for order in the partitioned subsets.

Putting this all together, we have $\boxed{S(n,k)\cdot k!\cdot\binom{m}{k} = S(n,k)\cdot\frac{m!}{(m-k)!}}$ different ways to select n numbers with order and with repetition such that there are exactly k distinct values.

12. Tails, tails, and tails.

Suppose Shreyas stores some number S initialized to 0. Every time he flips a fair coin, if it lands heads he increments S by 1 and if it lands tails he decrements S by 1. He wishes to calculate $\mathbb{P}[S \geq 20]$ after flipping the coin 100 times.

1. (5 points) Provide an exact answer using a summation. (Hint: $S \geq 20$ is equivalent to saying that Shreyas flips 20 more Heads than Tails.)

   Answer: $\sum_{k=60}^{100}\binom{100}{k}\left(\frac{1}{2}\right)^{100}$. We see that for S to be at least 20, Shreyas must flip at least 60 heads. We can get the value of 60 algebraically. Suppose H is the number of heads and T is the number of tails. Solving the equations $H+T=100$ and $H-T=20$, then we get $H=60$. Now, using $\mathbb{P}[X=i]=\binom{n}{i}p^i(1-p)^{n-i}$ for binomial variable (in this case $n=100$ and $p=1/2$), we get the answer.

2. (5 points) Provide an upper bound using Markov's inequality.

   Answer: $\frac{5}{6}$. Define $Y=100+S$. Since for 100 flips S is at least -100, Y is a nonnegative random variable. Recall that Markov's inequality requires the random variable to be nonnegative. We now apply Markov,

$$\mathbb{P}[S \geq 20] = \mathbb{P}[Y \geq 120] \leq \frac{\mathbb{E}[Y]}{120} = \frac{100}{120} = \frac{5}{6}.$$

   Another solution: We use the Hint from part A since H is guaranteed to be nonnegative,

$$\mathbb{P}[S \geq 20] = \mathbb{P}[H \geq 60] \leq \frac{\mathbb{E}[H]}{60} = \frac{5}{6}.$$

3. (5 points) Provide an upper bound using Chebyshev's inequality.

   Answer: $\frac{1}{4}$. Decompose $S = X_1 + X_2 + \ldots + X_{100}$ where each $X_i = 1$ if heads and -1 if tails. We see that $\mathbb{E}S = 0$ by linearity of expectation and $\mathrm{Var}(S) = 100\cdot 1 = 100$ since the variance of the sum of independent variables is the sum of their variances. Hence,

$$\mathbb{P}[S \geq 20] \leq \mathbb{P}[|S-0| \geq 20] \leq \frac{\mathrm{Var}(S)}{20^2} = \frac{100}{400} = \frac{1}{4}.$$

   Remark: We also provided full credit for the tighter solution $\frac{1}{8}$ derived by capitalizing on the symmetry of the binomial distribution. This was stronger than what we were looking for.

4. (5 points) Provide an upper bound using the Central Limit Theorem. (Leave your answer in terms of Φ, where Φ is the standard Normal CDF, if you deem necessary.)

   Answer: $1-\Phi(2)$. Use the same decomposition and expectation/variance results from the previous part. If the variance is 100 then the standard deviation is 10. Central Limit Theorem tells us that $Z = \frac{S-0}{10}$ is unit normal. Hence,

$$\mathbb{P}[S \geq 20] = \mathbb{P}\left[\frac{S-0}{10} \geq \frac{20-0}{10}\right] = \mathbb{P}[Z \geq 2] = 1-\Phi(2).$$

13. Uniform parameter for a geometric distribution.

Suppose Jonathan picks a real number R uniformly at random from the range $[0.25, 0.75]$. Then, he takes a coin that yields heads with probability R and tails otherwise and flips it until it yields heads. Let J denote the number of flips (including the last flip that yields heads).

1. What is $\mathbb{E}[R]$?

   Answer: 0.5. Its a uniform random variable, or one can do the integral.

2. (5 points) Compute $\mathbb{E}[J]$. Show your work and clearly indicate your final answer.

   Answer: $2\ln 3$. We know that $J \sim Geo(r)$, which means that $\mathbb{E}[J \mid R = r] = \frac{1}{r}$. Note that the PDF of R is 2 from 0.25 to 0.75. Then, by iterated expectation and LOTUS, we have:

$$\mathbb{E}[J] = \mathbb{E}[\mathbb{E}[J \mid R]] = \mathbb{E}\left[\frac{1}{R}\right] = \int_{0.25}^{0.75} \frac{1}{r} \cdot 2\,\mathrm{d}r = 2\ln r\Big|_{0.25}^{0.75} = 2(\ln 0.75 - \ln 0.25) = 2\ln 3$$

3. Compute $\mathbb{P}[J > 70 \mid R = r]$.

   Answer: Since $J \sim \text{Geo}(r)$, we have $\mathbb{P}[J > 70 \mid R = r] = (1-r)^{70}$.

4. (5 points) Compute $\mathbb{P}[J > 70]$. Show your work and clearly indicate your final answer. (The answer is a bit messy.)

   Answer: We know that $J \sim \text{Geo}(r)$, which means that $\mathbb{P}[J > 70 \mid R = r] = (1-r)^{70}$. Note that the PDF of R is 2 from 0.25 to 0.75. Then, by total probability and LOTUS, we have:

$$\mathbb{P}[J > 70] = \int_{0.25}^{0.75} \mathbb{P}[J > 70 \mid R = r] f_R(r)\,\mathrm{d}r = \int_{0.25}^{0.75} 2(1-r)^{70}\,\mathrm{d}r$$
$$= -\frac{2}{71}(1-r)^{71}\Big|_{0.25}^{0.75} = \frac{2}{71}(0.75)^{71} - \frac{2}{71}(0.25)^{71}$$

14. How many? And for how long?

(6 points) You purchase a box of light bulbs where the number of light bulbs is a Poisson random variable, $N \sim \text{Poisson}(\mu)$.

Let the random variable X_i denote the lifetime of the i^{th} bulb. The lifetimes of the bulbs are independent.

Moreover, each X_i obeys the exponential distribution

$$f_{X_i}(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0, \end{cases}$$

where $\lambda > 0$.

Determine a reasonably-simple expression for $\mathbb{E}[T]$, the expected total life that you can get from the light bulbs in your box. Your expression must be in closed form, and involve nothing other than a subset of the parameters λ and μ.

Note that

$$T = X_1 + \cdots + X_N,$$

where N is the random variable for the number of light bulbs in the box.

Answer: $\frac{\mu}{\lambda}$. $\mathbb{E}[T \mid N] = N\frac{1}{\lambda}$ by linearity of expectation and the expectation of an exponential random variable. Then, we have

$$\mathbb{E}[T] = \mathbb{E}[\mathbb{E}[T \mid N]] = \mathbb{E}\left[\frac{N}{\lambda}\right] = \frac{\mu}{\lambda}.$$

Equivalently, we have

$$\mathbb{E}[T] = \sum_{i=0}^{n} \mathbb{P}[N = i]\mathbb{E}[T \mid N = i] = \sum_{i=0}^{n} \mathbb{P}[N = i]\frac{i}{\lambda} = \frac{1}{\lambda}\sum_{i=0}^{n} i \cdot \mathbb{P}[N = i] = \frac{1}{\lambda}\mathbb{E}[N] = \frac{\mu}{\lambda}.$$

15. Just a moment or three.

Let $X \sim \mathcal{N}(\mu, \sigma^2)$.

1. Compute $\mathbb{E}[X]$.

   Answer: $\mathbb{E}[X] = \mu$. By definition of Normal random variable parameters.

2. Compute $\mathbb{E}[X^2]$.

   Answer: We can write $X = Y + \mu$, where $Y \sim \mathcal{N}(0, \sigma^2)$. Therefore, $\mu^2 + \sigma^2$. $\mathbb{E}X^2 = \text{Var}(X) + \mathbb{E}X^2 = \mu^2 + \sigma^2$.

3. (5 points) Compute $\mathbb{E}[X^3]$. *Hint:* $(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$.

   Answer: $3\sigma^2\mu + \mu^3$. We can write $X = Y + \mu$, where $Y \sim \mathcal{N}(0, \sigma^2)$. Then, we know that $\mathbb{E}[Y] = 0$, $\mathbb{E}[Y^2] = \sigma^2$, and $\mathbb{E}[Y^3] = 0$ because Y^3 is symmetric around 0.

$$\begin{aligned} \mathbb{E}[X^3] &= \mathbb{E}[Y+\mu)^3] = \mathbb{E}[Y^3 + 3Y^2\mu + 3Y\mu^2 + \mu^3] \\ &= \mathbb{E}[Y^3)] + \mathbb{E}[3Y^2\mu] + \mathbb{E}[3Y\mu^2] + \mathbb{E}[\mu^3] \\ &= 0 + 3\sigma^2\mu + 0 + \mu^3 = 3\sigma^2\mu + \mu^3 \end{aligned}$$

16. Joint Pizza Social.

Jonathan is taking some CS 70 staff members to a pizza social. However, he is unsure of how many staff members will show up! Let S be the number of staff members who show up and let P be the number of pizzas Jonathan orders. The joint distribution of S and P is given by the following table:

	$S=1$	$S=2$	$S=3$
$P=1$	0.05	0.1	0
$P=2$	0.1	0.05	0.15
$P=3$	0.05	0.1	0.4

1. Compute the marginal distributions of S and P.
   Answer: S: $0.2, 0.25, 0.55$ P: $0.15, 0.3, 0.55$.
2. Compute $\mathbb{E}[S]$ and $\mathbb{E}[P]$.
   Answer: We have

$$\mathbb{E}[S] = 0.2\cdot 1 + 0.25\cdot 2 + 0.55\cdot 3 = 0.2 + 0.5 + 1.65 = 2.35$$
$$\mathbb{E}[P] = 0.15\cdot 1 + 0.3\cdot 2 + 0.55\cdot 3 = 0.15 + 0.6 + 1.65 = 2.4$$

3. Compute $\mathbb{P}[S = P]$.
   Answer: By law of total probability, it is $0.05 + 0.05 + 0.4 = 0.5$.
4. Compute $\mathbb{P}[S = 3 \mid S \geq P]$.
   Answer: $\mathbb{P}[S \geq P] = 0.05+0.1+0+0.05+0.15+0.4 = 0.75$ (summing up the corresponding entries). Further, $\mathbb{P}[S = 3 \cap S \geq P] = \mathbb{P}[S = 3] = 0.55$. Therefore, the conditional probability is

$$\mathbb{P}[S = 3 \mid S \geq P] = \frac{\mathbb{P}[S = 3 \cap S \geq P]}{\mathbb{P}[S \geq P]} = \frac{0.55}{0.75} = \frac{11}{15} \approx 0.7333.$$

17. Indicators to an inequality.

A Bernoulli random variable I_A—called the *indicator* of event A—maps each sample point ω in the sample space Ω to either 0 or 1 according to the following rule:

$$I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A. \end{cases}$$

To avoid clutter, we often omit the dependence on ω, and write $I_A = 1$ if "A happens (or A is true)," and $I_A = 0$ if "A does not happen (or A is not true)." Clearly, $\mathbb{P}[\{I_A = 1\}] = \mathbb{P}[A]$.

Indicators for other events are defined in a similar fashion.

1. Show that $\mathbb{E}[I_A] = \mathbb{P}[A]$.
   Answer: $\mathbb{E}[I_A] = \mathbb{P}[I_A = 1] \times 1 + \mathbb{P}[I_A = 0] \times 0 = \mathbb{P}[I_A = 1] = \mathbb{P}[A]$.
2. Consider two events A and B and their corresponding indicators I_A and I_B.
   (a) Show that

$$I_A I_B = I_{A\cap B} = \min(I_A, I_B).$$

   Answer: One can observe that $I_A \times I_B = 1$ if and only if $I_A = 1$ and $I_B = 1$. This means that $\omega \in A \cap B$, and thus $I_A I_B = I_{A\cap B}$. Similarly, since both I_A and I_B must be 1, this can only happen when $\min(I_A, I_b) = 1$, giving us the two equalities $I_A I_B = I_{A\cap B} = \min(I_A, I_B)$.

(b) Show that

$$I_{A\cup B} = I_A + I_B - I_{A\cap B} = \max(I_A, I_B).$$

Answer: $I_{A\cup B}$ is 1 for $\omega \in A\cup B$. If $\omega \in A$ and $\omega \notin B$ then the expression $I_A + I_B - I_{A\cap B}$ is $1 = 1+0-0 = \max(1,0)$. Its symmetric, switching A and B. If $\omega \in A\cap B$, then we have $1 = 1+1-1 = 1 = \max(1,1)$ as desired. If $\omega \notin A\cup B$, then $I_{A\cup B} = 0$, and $0 = 0+0-0 = \max(0,0) = 0$. In each case, $I_{A\cup B} = 1$ when at least one of the indicator variables is 1, which occurs only when $\max(I_A, I_B) = 1$.

(c) (5 points) Show that the indicators I_A and I_B are uncorrelated if and only if the events A and B are independent.

Answer: $\text{Cov}(I_A, I_B) = \mathbb{E}[I_A I_B] - \mathbb{E}[I_A]\mathbb{E}[I_B] = \mathbb{E}[I_{A\cap B}] - \mathbb{E}[I_A]\mathbb{E}[I_B] = \mathbb{P}[A\cap B] - \mathbb{P}[A]\mathbb{P}[B]$. This is 0 if and only if $\mathbb{P}[A\cap B] = \mathbb{P}[A]\mathbb{P}[B]$, in which case A and B are independent.

3. (8 points) Let $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ be arbitrary (possibly negative) constants, and let $A_1, \ldots, A_n$ denote arbitrary events in a sample space Ω.

Show that

$$\sum_{k=1}^{n}\sum_{\ell=1}^{n} \alpha_k \alpha_\ell \mathbb{P}[A_k \cap A_\ell] \geq 0. \tag{1}$$

Hint: Let $I_{A_1}, \ldots, I_{A_n}$ denote the indicators for $A_1, \ldots, A_n$, respectively. Then, take advantage of the fact that

$$0 \leq \left(\sum_{k=1}^{n} \alpha_k I_{A_k}\right)^2$$

to progress toward the unintuitive result of Equation (1).

Answer: Using the hint, we have

$$\begin{aligned}
0 &\leq \left(\sum_{k=1}^{n} \alpha_k I_{A_k}\right)^2 \\
&= \left(\sum_{k=1}^{n} \alpha_k I_{A_k}\right)\left(\sum_{\ell=1}^{n} \alpha_\ell I_{A_\ell}\right) \\
&= \left(\sum_{k=1}^{n}\sum_{\ell=1}^{n} \alpha_k \alpha_\ell I_{A_k} I_{A_\ell}\right)
\end{aligned}$$

Taking expectations of both sides, we obtain:

$$0 \leq \left(\sum_{k=1}^{n}\sum_{\ell=1}^{n} \alpha_k \alpha_\ell \mathbb{E}[I_{A_k} I_{A_\ell}]\right) = \left(\sum_{k=1}^{n}\sum_{\ell=1}^{n} \alpha_k \alpha_\ell \mathbb{P}[A_k \cap A_\ell]\right),$$

where the final step follows from the fact that $I_{A_k} \times I_{A_\ell}$ is 1 if and only if both A_k and A_ℓ hold.

18. Probability Mass Function.

The probability mass function of a discrete random variable M is

$$p_M(m) = \begin{cases} \alpha m & \text{if } m \in 1, \ldots, n \\ 0 & \text{otherwise.} \end{cases}$$

Here, α is a constant.

In what follows, you may or may not find the following facts useful:

Sum of the First n Positive Integers:

$$\sum_{\ell=1}^{n} \ell = \frac{n(n+1)}{2}.$$

Sum of the *Squares* of the First n Positive Integers:

$$\sum_{\ell=1}^{n} \ell^2 = \frac{n(n+1)(2n+1)}{6}.$$

1. Determine α.

   Your answer must be a reasonably simple expression in terms of n.

   Answer: $\frac{2}{n(n+1)}$. The probabilities must sum to 1 and the sum of the first positive n integers is $\frac{n(n+1)}{2}$ which should be scaled down $\frac{2}{n(n+1)}$ to get 1.

2. Determine $\mathbb{E}[M]$. Leave your result here in terms of α.

   Answer: $\frac{(2n+1)}{3}$ or $\alpha\frac{n(n+1)(2n+1)}{6}$. This is $\sum_{i=1}^{n} m \times \alpha m = \alpha \sum_{i=1}^{n} m^2 = \alpha\frac{n(n+1)(2n+1)}{6} = \frac{2n+1}{3}$.

3. Determine the cumulative distribution function of M, defined by

$$\forall m \in \mathbb{Z}, \qquad F_M(m) = \mathbb{P}[M \leq m].$$

   Leave your result here in terms of α.

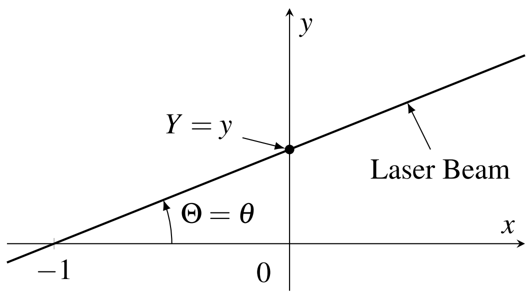
   Answer: We take the sum of the pmf up to m. The result is piecewise

$$F_M(m) = \begin{cases} 0 & \text{if } m < 1 \\ \dfrac{m(m+1)}{n(n+1)} \text{ or } \dfrac{\alpha m(m+1)}{2} & \text{if } m \in 1,\ldots,n \, . \\ 1 & \text{if } m > n \end{cases}$$

19. Spinning Laser. (Watch your eyes!)

A rotatable bidirectional laser is anchored at the point $(-1,0)$ in the xy-plane. We spin the laser and let it come to rest at a uniformly-random angle Θ, where the angle is measured in reference to the positive direction of the x-axis, as shown in the diagram below. Accordingly, Θ is distributed uniformly between $-\pi/2$ and $\pi/2$.

The random variable Y denotes the corresponding point of incidence of the laser beam on the vertical axis.



1. The random variable Y is a function of the random variable Θ. So, first, determine a simple expression for Y in terms of Θ.

   Answer: $Y(\theta) = \tan(\theta)$. This is due to the fact that for the right triangle formed by the x and y axis and the laser beam, we have $\tan(\theta) = y/1$.

2. Determine a reasonably simple expression for the cumulative distribution function (CDF) of Y, defined by

$$F_Y(y) = \mathbb{P}[Y \le y].$$

   Answer: $\frac{\arctan y}{\pi} + \frac{1}{2}$. The event $Y \le y$, is the same event as $\theta \le \arctan y$, and θ has pdf, $f_\Theta(\theta) = \frac{1}{\pi}$ over the range $[-\pi/2, \pi/2]$. Thus,

$$\mathbb{P}[\theta \le \arctan y] = \frac{1}{\pi}\int_{-\pi/2}^{\arctan y} \mathrm{d}\theta = \frac{\arctan y}{\pi} + 1/2.$$

3. Determine a reasonably simple expression for the PDF $f_Y(y)$. In doing so, and depending on how you tackle the problem, you may or may not find it useful to know that

$$\frac{d}{dy}\arctan y = \frac{1}{1+y^2}.$$

   Answer: $\frac{1}{\pi}\frac{1}{1+y^2}$. This is the derivative of the previous part.

20. Confused Concierge.

Four guests $\mathsf{G}_1, \dots, \mathsf{G}_4$ leave their umbrellas with Babak, the concierge at Hotel Satish. Babak is absent-minded and mixes up the umbrellas uniformly at random when handing them back to the guests.

We will determine $\mathbb{P}[C]$, the probability that Babak returns at least one umbrella to its rightful owner.

Let C_ℓ denote the event that Babak returns the correct umbrella to guest G_ℓ, where $\ell \in \{1,2,3,4\}$.

Recall the **Inclusion-Exclusion Principle** for n events $C_1, \dots, C_n$:

$$\mathbb{P}\left[\bigcup_{j=1}^{n} C_j\right] = \sum_k \mathbb{P}[C_k] - \sum_{k<\ell} \mathbb{P}[C_k \cap C_\ell] + \sum_{k<\ell<m} \mathbb{P}[(C_k \cap C_\ell \cap C_m)] - \cdots + (-1)^{n-1}\mathbb{P}(C_1 \cap \cdots \cap C_n).$$

For $n = 2$, the inclusion-exclusion principle translates to

$$\mathbb{P}[C_1 \cup C_2] = \mathbb{P}[C_1] + \mathbb{P}[C_2] - \mathbb{P}[C_1 \cap C_2].$$

For $n = 3$, the inclusion-exclusion principle translates to

$$\begin{aligned}\mathbb{P}[C_1 \cup C_2 \cup C_3] &= \mathbb{P}[C_1] + \mathbb{P}[C_2] + \mathbb{P}[C_3] \\ &\quad - \mathbb{P}[C_1 \cap C_2] - \mathbb{P}[C_1 \cap C_3] - \mathbb{P}[C_2 \cap C_3] \\ &\quad + \mathbb{P}[C_1 \cap C_2 \cap C_3].\end{aligned}$$

You need *not* write all the terms for the case $n = 4$, because there are too many. Instead, use the structures of the $n = 2$ and $n = 3$ cases to construct in your mind the general structure of the types of terms involved in the case $n = 4$.

1. What is $\mathbb{P}[C_1]$?

   Answer: $\frac{1}{4}$. One way to do this is computing the fraction of permutations that return umbrella 1 to Guest G_1 over the total number of permutations: $\dfrac{3!}{4!}$.

2. Using a single word, explain why $\mathbb{P}[C_i] = \mathbb{P}[C_j]$ for all $i, j \in \{1,2,3,4\}$.

   Answer: Symmetry. The calculation above applies to each C_j as the guest G_j gets their umbrella back in 3! ways out of a total of 4! permutations.

3. Using inclusion/exclusion, how many terms of the type $\mathbb{P}[C_k \cap C_\ell]$ are there?

   Answer: $\binom{4}{2}$. This follows from choosing 2 values of k and ℓ from 4.

4. (5 points) Compute the probability that at least one guest gets their umbrella back. Show your work, but clearly indicate your final answer.

   Answer: $15/24$. The desired probability is $\mathbb{P}(C_1 \cup C_2 \cup C_3 \cup C_4)$, which is obtained from the Inclusion-Exlusion Principle as follows:

$$\begin{aligned}\mathbb{P}[C_1 \cup C_2 \cup C_3 \cup C_4] &= \binom{4}{1} \times \frac{3!}{4!} - \binom{4}{2} \times \frac{2!}{4!} + \binom{4}{3} \times \frac{1!}{4!} - \binom{4}{4} \times \frac{1!}{4!} \\ &= 4 \times \frac{1}{4} - 6\frac{1}{12} + 4 \times \frac{1}{24} - 1 \times \frac{1}{24} \\ &= 1 - \frac{1}{2} + \frac{1}{6} - \frac{1}{24} = 15/24\end{aligned}$$

   We plug into the inclusion/exclusion formula. Each summand corresponds to the number of an intersection of a certain size, times the probability of those intersections. You could, of course, have used the brute-force method of writing all the 4! permutations and identifying the arrangements that had no match (9 out of 24) and subtracted it from 24 to get 15.

21. Balls and Bins.

Leanne throws n distinct balls into n distinct bins, labeled $1,2,\ldots,n$. Let X represent the number of ordered pairs $(i,i+1)$ of bins, for $1 \le i \le n-1$, such that bin i and bin $i+1$ are both empty.

1. (5 points) What is $\mathbb{E}[X]$? Show your work, but clearly indicate your final answer.
   ]2in
   Answer: Let I_k be an indicator variable for the event that both bin k and $k+1$ are empty. Each of the n balls needs to land in the other $n-2$ bins, so the probability that both bins are empty is $\left(\frac{n-2}{n}\right)^n$, and thus $\mathbb{E}[I_k] = \left(\frac{n-2}{n}\right)^n$. By Linearity of Expectation, we have

$$\mathbb{E}[X] = \sum_{k=1}^{n-1} \mathbb{E}[I_k] = (n-1)\left(\frac{n-2}{n}\right)^n.$$

2. (10 points) What is $\mathrm{Var}(X)$? Show your work, but clearly indicate your final answer.
   Answer: We have that

$$\begin{aligned}\mathbb{E}[X^2] &= \mathbb{E}[(I_1 + I_2 + \cdots + I_n)^2] \\ &= \sum_{k=1}^{n} \mathbb{E}[I_k^2] + \sum_{i \neq j, |i-j| \neq 1} \mathbb{E}[I_i I_j] + \sum_{i \neq j, |i-j| = 1} \mathbb{E}[I_i I_j].\end{aligned}$$

Since each I_k is an indicator, $\mathbb{E}[I_k^2] = \mathbb{E}[I_k]$. For $|i-j| \neq 1$, $\mathbb{E}[I_i I_j]$ corresponds to the four bins $i, i+1, j$, and $j+1$ all being empty, which occurs with probability $\left(\frac{n-4}{n}\right)^n$; there are $(n-1)(n-2) - 2(n-2) = (n-2)(n-3)$ such pairs. For $|i-j| = 1$, $\mathbb{E}[I_i I_j]$ corresponds to three bins being empty, which occurs with probability $\left(\frac{n-3}{n}\right)^n$; there are $2(n-2)$ such pairs. Thus, we get

$$\mathbb{E}[X^2] = (n-1)\left(\frac{n-2}{n}\right)^n + (n-2)(n-3)\left(\frac{n-4}{n}\right)^n + 2(n-2)\left(\frac{n-3}{n}\right)^n,$$

so

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = (n-1)\left(\frac{n-2}{n}\right)^n + (n-2)(n-3)\left(\frac{n-4}{n}\right)^n + 2(n-2)\left(\frac{n-3}{n}\right)^n - (n-1)^2\left(\frac{n-2}{n}\right)^{2n}.$$