This document lists a bunch of problems, that we read, solve, and discuss in class. More problems are added, as the course progresses.

1 Basics

1. In each part of this problem, state the sample space and give two examples of events. The first part is done for you.

A. Tossing a *coin* is a random experiment with two possible outcomes: H and T. The sample space S is $\{T, H\}$. There are four possible events: $\emptyset = \{\}$, $\{T\}$, $\{H\}$, and $S = \{T, H\}$. Notice that H is not an event, but $\{H\}$ is.

B. Rolling a *die* has six possible outcomes: 1, 2, 3, 4, 5, or 6.

C. Drawing a *card* from a standard card deck has 52 possible outcomes. The cards are organized into four *suits* (spades, hearts, diamonds, clubs) and 13 *kinds* (king, queen, jack, 10, 9, 8, 7, 6, 5, 4, 3, 2, ace), and there is exactly one card of each suit-kind combination. Additionally, spades and clubs are categorized as *black*. Hearts and diamonds are categorized as *red*. Kings, queens, and jacks are categorized as *face cards*.

D. Choosing a DNA *nucleotide* has four possible outcomes: adenine, cytosine, guanine, thymine.

2. Fix a day of the year (such as April 8). What is the probability that a randomly chosen person has that birthday?

3. You draw one card from a standard deck. What is the probability that it's a 4? That it's a face card? That it's a face card or it's red?

2 Counting

4. A *codon* is a sequence of nucleotides that encodes either an amino acid or a special "stop" instruction. The order of nucleotides within the codon matters. (It can affect the amino acid produced.) There are 20 amino acids, so 21 distinct codons are needed. Assuming that all codons are of the same length, how long must they be?

5. Anna, Babatope, and Chiara race. Ties can't happen, so one person ends up first, one ends up second, and one ends up third. How many ways are there for the race to end up?

6. A *hand* is a set of five cards drawn from a single deck (which forces the five cards to be distinct). How many possible hands are there? The order of the cards within the hand does not matter, but maybe first solve the problem assuming that the order does matter.

7. You draw a hand. What is the probability that it contains four aces? How would your answer change, if there were six cards in a hand?

8. A *royal flush* is a hand consisting of an A, K, Q, J, 10, all of a single suit. What is the probability that a randomly dealt hand is a royal flush?

9. You roll five dice. What is the probability that you get exactly three 4s?

10. A *full house* is a hand in which three cards are of the same kind as each other, and the other two cards are of the same kind as each other. For example, {KS, KH, KD, 8H, 8C} is a full house. What is the probability of a full house? It's the number of full-house hands divided by $\binom{52}{5}$. But what is the number of full-house hands?

Here are four answers. Which ones are correct?

A. First pick two kinds ($\binom{13}{2}$ options). Then pick three cards from the first kind ($\binom{4}{3}$ options) and two cards from the second kind ($\binom{4}{2}$ options). So the answer is

$$\binom{13}{2} \cdot \binom{4}{3} \cdot \binom{4}{2}.$$

B. First pick the kind that will have three cards (13 options). Within that kind, pick three cards ($\binom{4}{3}$ options). Then pick the kind that will have two cards (12 options). Pick two cards from that kind ($\binom{4}{2}$). The answer is

$$13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2}.$$

C. First pick a card (52). Pick another card from the same kind (3) and another (2). Then pick a card of a different kind (48), and another card from that kind (3). Finally, divide by 3! to allow the first three cards to permute, and divide by 2! to allow the last two cards to permute. The answer is

$$\frac{52 \cdot 3 \cdot 2 \cdot 48 \cdot 3}{3! \cdot 2!}.$$

D. Follow the same strategy as C. However, instead of dividing by $3! \cdot 2!$, just divide by 5!, to allow all five cards to permute.

11. How many 13-letter words (valid or invalid) can you make by rearranging the 13-letter word "probabilistic"? Keep in mind that "probabilistic" has only 10 distinct letters. For example, switching the two "b"s does not produce a new word.

12. A *monomial* is a product of variables raised to some powers. The *degree* of a monomial is the sum of the powers. For example, $x^2y^3z^2$ is a monomial of degree 7, xz^9 is a monomial of degree 10, and y is a monomial of degree 1. So here's my question: How many degree-5 monomials are there, in the variables x, y, z?

13. How many degree-d monomials are there, in n variables? (By the way, this kind of problem is sometimes called Bose-Einstein, because it came up in Bose's and Einstein's studies of arrangements of bosons.)

3 Birthday problem

14. Police use DNA to investigate crimes. In Arizona as of 2001, nine loci (chunks) of DNA are recorded to make a profile for a person. Without going into the technicalities, there are 754,000,000 possible profiles that can arise from these nine loci. For simplicity, assume that the profiles are all equally probable.

A. DNA (one profile's worth) is found at a crime scene. What is the probability that a randomly chosen person's profile will match it? So if a defendant's DNA matches that found at the crime scene, then do you, as a jury member, consider it strong or weak evidence that the defendant was there?

B. The defense attorney, whose job it is to cast doubt on such evidence, makes a shocking revelation. After studying the database, which has 65,493 profiles in it, the attorney has found two people with identical profiles. So maybe the database is flawed. For example, maybe some of the data are damaged, or maybe not all profiles are equally probable. What do you think?

4 Conditional probability

15. In a hypothetical city of $100,000$ voters, 52% identify as Black, 48% as not Black, 60% as Democratic, and 40% as not Democratic. Also, three quarters of Black voters are Democrats. Let B be the event that a randomly chosen voter is Black and D the event that they are Democratic. Fill in the contingency table below.

	D	D^c	
B	$P(BD) =$	$P(BD^c) =$	$P(B) =$
B^c	$P(B^cD) =$	$P(B^cD^c) =$	$P(B^c) =$
	$P(D) =$	$P(D^c) =$	

Then compute the probability that a randomly chosen Democrat is Black.

16. This example is about medical screening for a particular disease. Suppose that 0.1% of the population has the disease. Let D be the event that a randomly chosen person has the disease, and let T be the event that they test positive. The test has a "sensitivity" of 99%, which means that $P(T|D) = 0.99$, and which implies that the "false negative rate" is $P(T^c|D) = 0.01$. The test has a "specificity" of 99%, which means that $P(T^c|D^c) = 0.99$, and which implies that the "false positive rate" is $P(T|D^c) = 0.01$. Fill in the contingency table below.

	D	D^c	
T			
T^c			

Unfortunately, you've just been told that you've tested positive. I'm sorry. What is the probability that you have the disease?

17. You are an epidemiologist studying drug use. If you simply ask people, "Have you ever injected illegal drugs?", then you might not get honest answers. So instead you adopt the following protocol. You tell each participant:

1. Flip a coin twice. Don't show me or tell me the results.
2. If the first flip was H, then answer the question, "Was the second flip H?"
3. If the first flip was T, then answer the question, "Have you ever injected illegal drugs?"

Let F be the event that the first flip was H, and let Y be the event that they answered "Yes". In this notation, what do you want to know about the population? What does your questioning of a bunch of people tell you? How can you compute what you want from what you know?

18. On Monty Hall's game show, there are three closed doors. Behind one door is a car, and behind each of the other doors is a goat. You're trying to pick the door that hides the car. The game begins when you pick a door. Then Monty opens one of the other doors that hides a goat. Then you can either stay with your choice of door, or switch to the other closed door.

A. What's your probability of winning the car, if you employ the strategy of always staying with your initial pick of door?

B. What's your probability of winning, if you always switch to the other closed door?

19. Ever since we were kids, my brother has been interested in rigging dice so that they roll 6 more than 1/6 of the time. (There are various approaches: sand off edges, insert a lead weight, etc.) While visiting him one day, I see ten dice on his coffee table. I pick one up and immediately roll three 6s. So I ask him, "Is this die rigged?" Without looking up from the lentils that he's rinsing, he says, "One of those dice is rigged so that 28% of the time it rolls 6. The other nine dice are fair." What is the probability that I rolled the rigged die?

20. Last I checked, studies of the Internet estimated that about 80% of e-mail was junk mail. However, your e-mail inbox doesn't look that bad (I hope), because your e-mail provider and your e-mail program filter out most junk mail before you ever see it. Junk mail filtering is partly based on the words within messages. For example, if you work in the wristwatch industry, then you might receive a lot of legitimate e-mail about Rolex watches. But if you don't work in that industry, then a message containing the word "Rolex" is probably junk. More precisely, the presence of "Rolex" increases the probability that a given message is junk, compared to otherwise-similar messages that do not contain "Rolex". Use this idea to design a junk-mail filter based on a single word such as "Rolex".

21. In an experiment, a scientist proposes a hypothesis and gathers some data. Let H be the event that the hypothesis is true. Let D be the event that, in any repetition of the experiment, data like the scientist's, or data more "extreme" than that, are collected. Using standard statistical techniques, the scientist computes a *p-value*, which is $P(D|H)$. They find it to be 0.02. That is, if the hypothesis were true, then such data would probably not have shown up. They conclude, "The hypothesis is probably false." Discuss.

22. Draw two cards. Let A be the event that you draw an ace on the first draw, and let B be the event that you draw an ace on the second draw. Are A and B independent?

23. Recall our hypothetical city of 100,000 voters, with B the event that a randomly chosen voter is Black and D the event that the voter is a Democrat. See below for the contingency table. Are B and D independent?

	D	D^c	
B	0.39	0.13	0.52
B^c	0.21	0.27	0.48
	0.60	0.40	1

24. Flip two coins. Let A be the event that the first flip was H, B the event that the second flip was H, and C the event that the two flips gave the same result (either HH or TT). Are A, B, and C independent?

25. A giant tree is the only member of its species in its forest. In any given year, it has a 0.1% chance of dying in early-summer lightning storms and a 1% chance of reproducing in the late-summer mating season (via pollen drifting great distances on the wind).

A. What is its probability of reproducing ever?

B. When writing this problem, why did I specify "early-summer" and "late-summer"?

5 Six discrete distributions

26. To finish the board game that I'm playing with my daughter, I need to roll a 3. I try again and again until I get that 3. Let X be the number of failures before my first success.

A. What's the support of X?

B. What's the probability that I have four or more failures? Five or more? Exactly four?

C. What's the PMF of X?

27. I'm an airport security screener. Suppose that 1/30 of travelers carry some kind of contraband: bombs, big shampoo, etc. Let $X \sim \text{Geom}(1/30)$.

A. What is the meaning of X, in the context of my airport work?

B. What does $P(X \geq 20)$ mean, and what is its value?

C. What is the value of $P(X \geq 20 | X \geq 15)$? Discuss.

28. In quantum computing, Shor's algorithm is used to do certain tasks, such as factoring large integers, that are relevant to cryptography. The crux of the algorithm is a subroutine that, each time it's invoked, has probability $4/\pi^2$ of solving the problem. How many invocations of the subroutine are needed to solve the problem? Answer in terms of a geometrically distributed random variable.

29. Desirée is a young scientist with many ideas for projects. For each idea, she writes up a grant proposal for the National Science Foundation. Each proposal has a 20% of being funded, independently of the others. If an idea is not funded, then she forgets about it and moves on to her next idea. To earn tenure, she needs two projects to be funded. How many projects should she propose, so that she has at least a 50% chance of earning tenure?

30. Desirée's colleague Jimmy is in the same boat, but he has only two project ideas. If a proposal isn't funded in one year, then he edits it and re-submits it the next year. He can propose each idea only once per year, so he has five chances per idea, before his tenure decision is made. What's the probability that Jimmy gets tenure?

31. You have five grandchildren, all of whom love the Captain Punchalot cartoon. The McDonald's restaurant chain sells Happy Meals, each of which contains one figurine — of either Captain Punchalot, her hilarious sidekick Plocky-Dee, her pet Komodo dragon Mabel, or the evil Dr. Davis — uniformly randomly chosen.

A. How many Happy Meals must you buy, to obtain a Plocky-Dee figurine for each of your five grandchildren? Answer using a random variable from a specific distribution.

B. What is the probability that you will achieve your goal in 10 or fewer purchases?

32. Unlike you, I have only one grandchild, and I want to collect all four Captain Punchalot figurines for her. I want to know how many purchases are needed.

A. Explain how the answer can be expressed using a sum of geometric random variables.

B. Can you re-phrase your answer using negative-binomial random variable? If so, do it. If not, explain why not.

6 Expectation, variance

33. In each part, compute the expectation $E(X)$.

A. Let X be a die roll.

B. What if the sides were numbered $1, 2, 4, 8, 16, 32$?

C. What if the sides were numbered in the ordinary way, but the die was biased to produce 6 half the time, with the other outcomes equally probable?

34. An insurance company insures homes, and the contents of those homes, against fire. Based on their data, the company identifies the three groups of policies below. What is the expected payout by the insurance company to a randomly selected policy holder?

	probability	payout
no fire	0.9989	0
one minor fire	0.001	100,000
one major fire	0.0001	1,000,000

35. Let $X \sim \text{Binom}(8, p)$. Let $Y = (X - 3)^2$.

A. What is the support of X?

B. What is the support of Y?

C. What is $P(Y = 0)$? What is $P(Y = 1)$? In general, what is the PMF of Y?

36. The contingency table below describes a new hypothetical city. For example, 31% of voters are female Democrats. Pick a voter randomly. Let X be their party affiliation (1, 2, or 3) and Y their gender (1, 2, or 3).

	1. Democrat	2. Independent	3. Republican	
1. Female	31%	1%	16%	
2. Male	15%	3%	29%	
3. Other	3%	1%	1%	

A. What probability distribution does X have?

B. What probability distribution does Y have?

C. Are X and Y independent?

37. Pick two people at random. Let $X, Y \sim \text{DUnif}(\{1, 2, 3, \ldots, 365\})$ be their birthdays. Assume that X and Y are independent. Are $X + Y$ and $X - Y$ independent?

38. Roll two dice. Let X be the first and Y the second. Let $Z = \max(X, Y)$.

A. Compute $E(Z)$ by finding the support, the PMF, and the expectation from those.

B. Compute $E(Z)$ using the "shortcut" (Equation 4.3).

39. The definition of standard deviation requires that $V(X) \geq 0$ always. How do we know that that's true? And under what conditions is $V(X) = 0$?

40. Let X be the high temperature in Northfield on an October day, measured in centigrade. Then what is $Y = \frac{9}{5}X + 32$ in English? What are $E(Y)$, $V(Y)$, and $SD(Y)$, in terms of $E(X)$, $V(X)$, $SD(X)$?

7 Joint distributions

41. The *standard 2D normal distribution* has PDF

$$f_{X,Y}(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}.$$

If X and Y have this joint distribution, then are they independent?

42. Suppose that X and Y are jointly distributed according to the standard 2D normal distribution described in the preceding problem. What are the meaning and the value of $E(\sqrt{X^2+Y^2})$?

43. Suppose that the radius of a sphere is $R \sim \text{Unif}(0,4)$. Then what is the volume V? To answer that question, find the support and the PDF on that support.

44. A random variable Y is *log-normally distributed* if $\log Y$ is normally distributed (with any mean μ and variance σ^2). Find the support and PDF of Y.

45. Consider a Poisson process with rate λ. Let T be the second arrival time.

A. Explain how T is the sum of two IID (independent and identically distributed) random variables X, Y. What is their distribution?

B. Compute the PDF of T using convolution.

46: Xiuxiong picks a number X uniformly on $[0,1]$. Then Yolante picks Y uniformly on $[0,X]$. What is the PDF of Y?

47: Marjane is an Iranian girl. Let Y be Marjane's life span in years. Based on what you know of how people live and die, how do the following life expectancies compare?

A. $E(Y)$

B. $E(Y|Y \geq 0)$

C. $E(Y|Y \geq 2)$

D. $E(Y|Y \geq 4)$

48. Recall Problem 46 above: Xiuxiong picks a number X uniformly on $[0,1]$, and then Yolante picks Y uniformly on $[0,X]$.

A. What is $E(Y|X=x)$? Does the answer make intuitive sense?

B. What is $E(Y|X)$, symbolically? What exactly is the distribution of this random variable?

C. Compute the expectation of the random variable $E(Y|X)$. Also compute $E(Y)$ using the PDF that you found in Problem 46.

49. Ezinma works as a waitress. On her ith day of work, she makes X_i dollars in tips. Assume for simplicity that the X_i are IID (independent and identically distributed), each with expectation μ. Let $S_n = X_1 + X_2 + \cdots + X_n$ be her tip total after n days.

A. What is $E(S_m|S_n)$, where $m = n$?

B. What is $E(S_m|S_n)$, where $m > n$?

C. Can you handle the other case, where $m < n$, in the same way?

50. Return once more to the setting of problems 46 and 48, where Xiuxiong picks a number X uniformly on $[0, 1]$, and then Yolante picks Y uniformly on $[0, X]$. What is the conditional variance of Y given $X = x$?

8 Limit theorems

51. Which of our named distributions are approximately normal, because of the central limit theorem?

52. Let \$$X$ be a casino's earnings from a single play of roulette. Recall from Exam B that $E(X) = 1/19$ and $V(X) = 360/361$. Assume that the casino has $n = 10{,}000$ independent roulette plays per day.

A. What are the expectation and variance of the casino's daily roulette earnings?

B. Using the central limit theorem, describe the casino's daily roulette earnings and the uncertainty in those earnings. For example, there is a 95% chance that the earnings are between what and what?

C. If the casino could increase n, would that be good for its business or not? Discuss.

9 Review

53. Parts A and B describe two "games". For each game, compute your expected winnings and the variance in those winnings. Then do part C.

A. Flip a coin. If it's heads, then you win \$100. If it's tails, then you lose \$100.

B. Flip 100 coins. For each one, you win \$1 (heads) or lose \$1 (tails).

C. Does the central limit theorem have anything to say about any of this?

54. In one sector of the economy, there are n companies. Let $X_1, X_2, \ldots, X_n$ be their profits next year, respectively. Let $S = X_1 + \ldots + X_n$ be the total profit of that sector next year. For simplicity, assume that the profits are identically distributed, each with expectation μ and variance σ^2. (This assumption is highly unrealistic, but adding more realism doesn't improve this problem.)

A. Why should policy makers care about $E(S)$ and $V(S)$? Do policy makers want these two numbers to be great or small?

B. Assuming that the profits are independent, what are $E(S)$ and $V(S)$?

C. No longer assuming that profits are independent, how bad could $V(S)$ be? (Hint: Using what you know about $\text{Corr}(X_i, X_j)$, put a bound on $\text{Cov}(X_i, X_j)$.)

D. Should policy makers encourage independence or dependence? Discuss.

55. In class, we've glimpsed how regression relates to conditional expectation. This problem fills in slightly more detail. Suppose that X and Y are jointly distributed continuous random variables. You plot some data points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, and the point cloud looks vaguely like a parabola opening up. So you suspect that Y depends quadratically on X. That is, $Y = aX^2 + bX + c + \mathcal{E}$, where a, b, and c are unknown constants and $\mathcal{E}$ is a continuous random variable capturing the "noise" or "error" in the quadratic relationship.

A. Assume that $Y = aX^2 + bX + c + \mathcal{E}$. Assume that $\mathcal{E}$ has expectation 0 and is independent of X. Using basic facts about conditional expectation, compute $E(Y|X)$, simplifying your answer as much possible.

B. (My fall 2025 Math 240 students are not expected to be able to do this part.) Impose the same assumptions. Also assume that the variance of $\mathcal{E}$ is independent of X (a condition called homoscedasticity). Use basic facts about conditional variance to compute $V(Y|X)$, again simplifying.

56. Recall that a Rayleigh-distributed X has PDF $f_X(x) = xe^{-x^2/2}$ on support $(0, \infty)$. Suppose that I choose a Rayleigh X and then you choose $Y \sim \text{Norm}(X, 1)$.

A. What is the joint PDF of X and Y?

B. What is the marginal PDF of Y? You may leave your answer in non-closed form: a sum, series, or integral.

C. Are X and Y independent? Explain.

57. Let X have PDF $f_X(x) = x^{-2}$ on support $[1, \infty)$. Let Y be independent with the same distribution. Let $Z = XY$.

A. What is the support of Z?

B. What is the PDF of Z on that support?

58. You work for an online retailer that individually tracks its customers. Each customer is assigned a number λ, such that the occasions of her or his purchases follow a Poisson process with rate λ, where the time unit is the year. For example, one of your customers is a misunderstood wizard named Voldemort. He hasn't got much of a nose, so he can't smell food, but he buys a lot of candy, because he loves the sweet taste. He averages 5 purchases per month.

A. What is Voldemort's λ?

B. What is the expected time until Voldemort's next purchase?

C. What is the probability that Voldemort will make 3 or more purchases in the next 7 days?

59. Let $X \sim \text{Expo}(\lambda)$ and $Y = X^4$.

A. We can compute $E(X^4)$ using the law of the unconscious statistician. Set up the integral in detail, but do not compute its numerical value.

B. We can compute $E(Y)$ using the definition of expectation. Figure out the PDF of Y, and set up the integral for $E(Y)$ in detail, but do not compute the integral's numerical value.

C. Are you confident that the integrals would give the same answer, if we computed them? (Hint: What is the formal relationship between dy and dx?)

60. I often give you problems of the form, "Ask a question, whose answer is a random variable of the ... distribution." And questions of the form, "What is the probability that ...?" have never been correct. But it might be theoretically possible for such a question to be correct. This problem tries to guide you through the mental contortions necessary.

A. If a question of the form "What is the probability that ...?" has a random variable X as its answer, then what is the support of X, apparently?

B. Of the named distributions that we have studied, which distribution (or distributions) can have that support? With what parameter values?

C. Give a question of the form "What is the probability that ...?", whose answer is a random variable of that distribution with those parameter values. (Disclaimer: I tried to come up with such a question for half an hour, but I couldn't nail it down. I gave up.)

61. Let $X \sim \text{Norm}(\mu, \sigma^2)$. What is $P(\mu - \sigma \leq X \leq \mu + 2\sigma)$?

62. Here's a glimpse of how conditional expectation connects to a certain kind of regression. Let X and Y be continuous, and suppose that we have a data set $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$. What makes this kind of regression different from other kinds is how the abstract ideal $E(Y|X)$ is approximated. Let h and k be any two PDFs, such that k has mean zero. The approximation arises from declaring that

$$\begin{aligned} f_X(x) &= \frac{1}{n} \sum_{i=1}^{n} h(x - x_i), \\ f_{X,Y}(x, y) &= \frac{1}{n} \sum_{i=1}^{n} h(x - x_i) \cdot k(y - y_i). \end{aligned}$$

Show that, once we accept these declarations,

$$E(Y|X = x) = \sum_{i=1}^{n} \left(y_i \cdot \frac{h(x - x_i)}{\sum_{j=1}^{n} h(x - x_j)} \right).$$

63. In plain English, what does the strong law of large numbers say? How did we use it in class?

64. Let X have one of our named distributions. Compute the MGF of X from scratch.

65. Let $X \sim \mathrm{NBin}(r, p)$. For large r, approximately 95% of the probability mass is between what and what?