1 Initial concepts

In each problem, state the sample space, two examples of events, and two non-examples of events. The first problem is done for you.

1. Tossing a *coin* is a random experiment with two possible outcomes: 0 (also called heads) or 1 (also called tails). The sample space S is $\{0, 1\}$. There are four possible events: $\emptyset = \{\}$, $\{0\}$, $\{1\}$, and $S = \{0, 1\}$. Notice that 1 is not an event. Neither is "my dog ate the coin".

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

3. Drawing a *card* from a standard 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*.

4. Choosing a DNA *nucelotide* has four possible outcomes: adenine, cytosine, guanine, thymine.

2 Counting

5. Let F be the event that a chosen card is a face card, and let R be the event that a chosen card is red. In English, describe $F \cap R$, $F \cup R$, and FR. Also, how many outcomes are in each of these five events?

6. 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?

7. A *hand* is a set of five cards drawn from a single deck, which forces the five cards to be distinct. The order of the cards within the hand does not matter. How many possible hands are there?

3 Probability

8. You draw four cards from a single deck. What is the probability that they are all aces?

9. You draw a hand. What is the probability that it contains four aces?

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 is 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.

4 Birthday problem

11. How does the probability of distinct birthdays change, if we account for the fact that there are 366 possible birthdays? (Assume for simplicity that the extra day, February 29th, is equally probable as the other days, even though this is highly doubtful.)

12. Police use DNA to catch criminals. In a certain state's DNA database, 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 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 state's DNA 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 not all profiles are equally probable. What do you think?

5 Bayes' theorem

13. 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 ways to do it: 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 beer that he's brewing, he responds, "One of the dice on that table 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?

14. Studies of the Internet estimate that about 80% of e-mail is junk mail (unwanted commercial solicitations). 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".

15. 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 the *p-value*, which is $P(D|H)$. She finds it to be 0.02. That is, if the hypothesis were true, then she probably would not have collected such data. She concludes, "The hypothesis is probably false." Discuss.

6 Independence

16. Recall our hypothetical USA city of 100,000 people, with A the event that a randomly chosen person is African-American and D the event that the person is a Democrat. See below for the probability table. Are A and D independent?

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

17. You are playing a card game. The 7 of diamonds is in your hand. You are close to winning the game; as soon as you draw another 7 or another diamond, you win. Ignoring the rest of your hand and the other players' hands (or assuming that you're the only player and the 7 of diamonds is your only card), which do you expect to draw first: 7 or diamond? What's

the probability of drawing a 7 first? A diamond first? (Warning: After writing this problem, I realized that the draws are not independent. Why? And so how do you answer the problem?)

18. 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). What is its probability of reproducing ever?

7 Bernoulli, geometric

19. I'm playing a board game with my daughter. On each turn, each player rolls a die and takes some action. I'm about to win. I win the next time I roll a 1. My daughter is so far from winning that let's just ignore her. Let X be a random variable representing how many tries I need.

A. What is the support of X? What is the PMF of X?

B. What's the probability that I need 4 or more tries? What about 5 or more tries?

20. I'm an airport security screener. Suppose that 1/30 of travelers carry some kind of contraband: bombs, hidden 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.

21. In quantum computing, Shor's algorithm is used to do number-theoretic tasks such as factoring large integers. 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.

8 Binomial, hypergeometric, negative binomial

22. Desirée is a young scientist with many ideas for projects. For each project, she applies once to the National Science Foundation for funding. Each application has a 20% of being funded, independently of the others. To earn tenure, she needs two projects to be funded. How many times should she apply for funding, so that she has at least a 50% chance of earning tenure?

23. Desirée's colleague Jimmy is in the same boat, but he has only two project ideas. His strategy is to apply for funding repeatedly until both of those projects are funded. He can apply for each project only once per year, so he has only five chances per project, before his tenure decision is made. What's the probability that Jimmy gets tenure?

24. On a campus of 2,000 people, you randomly ask 200 of them the question, "Do you support the legalization of marijuana?" You get 90 "yes"s and 110 "no"s.

A. Explain how the number of "yes"s can be viewed as a hypergeometric random variable X. In particular, explain the parameter values; one of them should be unknown.

B. In a statistical analysis, the next step might be to compute the *maximum likelihood estimate* (MLE): the value of the parameters that maximizes the probability of seeing the data. State the marijuana legalization MLE problem explicitly, in mathematical notation, so that you could in principle find the answer, given a computer and enough time. (You are not expected to compute the answer by hand.)

25. 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?

26. 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. Explain how your answer compares to a negative-binomial random variable.

9 Functions of random variables, independence

27. 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?

28. Roll two dice. Let X be the first and Y the second. So $X, Y \sim \text{DUnif}(1, 2, 3, 4, 5, 6)$. Let $Z = \max(X, Y)$. Find the PMF of Z, including the support.

29. Let $X, Y \sim \text{DUnif}(1, 2, \ldots, 365)$ be the birthdays of two siblings. Assume that X and Y are independent. Let A be the event that $X > Y$ — that is, the first sibling is born later in the year than the second sibling is. Is X independent of the indicator random variable I_A?

30. Suppose that $X_1, \ldots, X_n$ are IID geometric random variables. Then what can you say about $X = X_1 + \cdots + X_n$?

10 Expectation, variance

31. 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

32. Use linearity of expectation to compute $E(X)$, where $X \sim \text{NBinom}(r, p)$.

33. Notice that $V(X) \geq 0$ always. Why? Under what conditions is $V(X) = 0$?

34. 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 the same functions of X?

35. This is a question about two of our named discrete distributions. The two parts are related only by the method used to analyze them.

A. If $X \sim \text{Binom}(n, p)$, then what is $V(X)$?

B. For what other distribution can you compute the variance using nearly the same argument?

11 Continuous random variables

36. For some constant $k > 0$, define a PDF f by

$$f(x) = \begin{cases} k(\cos(4\pi x) + 1) & \text{if } -1/4 \leq x \leq 1/4, \\ 0 & \text{otherwise.} \end{cases}$$

A. Roughly sketch the graph of f and its CDF F.

B. Find k.

C. Find F.

37. Let $X \sim \text{Expo}(\lambda)$. Compute $E(X)$. Then set up the computation for $V(X)$, and complete it if you have time.

38. You run the emergency room in a hospital in a big city. You get 10 patients per hour, on average. Patient arrivals can be modeled as a Poisson process.

A. Find λ, and pose questions answered by $X \sim \text{Pois}(\lambda)$ and $Y \sim \text{Expo}(\lambda)$.

B. What's the probability that more than 10 patients arrive in the next hour?

C. What's the probability that no patients arrive in the next hour? Compute it in two ways: using a Poisson random variable and an exponential random variable (perhaps X and Y).

39. A *hard disk drive* is a mechanical device that rotates thousands of times per second, often for years on end. Consequently it lasts only a few years before breaking. A data center has 100,000 drives. Approximately every 15 minutes, one of the drives fails and needs to be replaced. Repeat the preceding exercise in this new context. In fact, do part A twice, using two different units of time.

40. It turns out that if $X \sim \text{Norm}(\mu, \sigma^2)$, then $aX + b$ is also normal.

A. What are the parameters of the distribution of $aX + b$, in terms of μ, σ^2, a, b?

B. How would you scale and translate X to make a random variable $Z \sim \text{Norm}(0, 1)$?

41. Let $X \sim \text{NBinom}(r, p)$. For large r, X and X/r are approximately normal. Why? With what means and variances?

12 Joint distributions

42. Let $f(x, y) = c(1 - x^2 - y^2)$ on the unit disk $\{(x, y) : x^2 + y^2 \leq 1\}$.

A. Sketch the PDF.

B. Compute c.

C. Compute $P(\sqrt{X^2 + Y^2} \leq 1/2)$.

D. Compute $P(X \leq 0, Y \leq 0)$ in two ways: avoiding integration, and not avoiding it.

E. Compute $P(X \leq Y)$ in two ways.

43. Recall our hypothetical USA city of 100,000 people, with A the event that a randomly chosen person is African-American and D the event that the person is a Democrat. See below for the probability table. Let $X = I_A$ and $Y = I_D$. What is the joint PMF $p_{X,Y}$? What are the marginal distributions of X and Y?

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

44. For X, Y continuous, show that the conditional density $f_{Y|X}(y|x)$ integrates to 1.

45. Continuing problem 43 above, what is the conditional PMF $P(Y|X = 1)$?

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

47. 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 or not?

48. Let $X \sim \text{Unif}(-1, 1)$ and $Y = X^2$. Explain intuitively why X and Y are not independent. Compute their covariance. In English, why have I assigned you this exercise?

49. Let $X \sim \text{Expo}(\lambda)$ and $Y = 7X + 3$. Find the PDF of Y, including its support.

50. Let $X \sim \text{Norm}(\mu, \sigma^2)$ and $Y = e^X$. We say that Y is *log-normally distributed*. What is the support and PDF of Y?

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

A. Explain how T is the sum of two IID random variables X, Y. What is their distribution?

B. Compute the PDF of T using convolution.

52. Let $X, Y \sim \text{Expo}(\lambda)$ be IID. Find the PDF of $Z = Y/X$.

13 Conditional expectation

53. Recall Problem 46 above: Wyclef picks a number X uniformly on $[0, 1]$. Then Xiuxiong picks Y uniformly on $[0, X]$.

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

54. 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, each with expectation μ. Let S_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. Why am I not asking you about the other case: $E(S_m|S_n)$, where $m < n$?

14 Moment generating functions

55. Compute the moment generating function of $X \sim \text{Expo}(\lambda)$.

56. Prove that if X and Y are independent random variables, then $m_{X+Y}(t) = m_X(t) \cdot m_Y(t)$.

57. Let $Z \sim \text{Norm}(0, 1)$. Let σ be a positive constant and μ any constant.

A. Compute the MGF of σZ.

B. Compute the MGF of $\mu + \sigma Z$. (Hint: Regard μ as a random variable that is independent of all other random variables.)

C. What is the MGF of $X \sim \text{Norm}(\mu, \sigma^2)$?

15 Inequalities and limit theorems

58. What does Markov's inequality say, in the case where $\epsilon \approx 0$? Does the answer make sense? (If you can't evaluate whether it makes sense, then try to manipulate it further, until you can.)