1. **Waiting**

   A pair of dice is rolled until either a 4 is rolled (the numbers on the two dice add up to 4) or a 7 is rolled. What is the expected number of rolls needed?

2. **Poisson**

   The number of accidents (per month) at a certain factory has a Poisson distribution. If the probability that there is at least one accident is 1/2, what is the probability that there are exactly two accidents?

3. **Savings**

   Alice buys a piggy bank. Every day she picks a number X uniformly at random from $\{0,1,2\}$. If X is nonzero, she puts X dollars into her piggy bank for that day. If X is zero, she breaks the piggy bank and takes away all the money saved. What is the expected amount of money Alice gets when she breaks the bank?

4. **Quadruply-repeated ones**

   We say that a string of bits has k *quadruply-repeated ones* if there are k positions where four consecutive 1's appear in a row. For example, the string 0100111110 has two quadruply-repeated ones.

   What is the expected number of quadruply-repeated ones in a random n-bit string, when $n \geq 3$ and all n-bit strings are equally likely? Justify your answer.

5. **A flawed shuffle**

   Consider the following bad method for "shuffling" (i.e. randomly permuting the elements of) a 52-element array A.

   (a) Initialize an array *newA* to contain 52 "empty" indicators.

   (b) For $k = 0$ to 51, do the following:

      i. Repeatedly generate a random integer j between 0 and 51 until $A[j]$ isn't empty.

      ii. Copy $A[j]$ to $newA[k]$, and set $A[j]$ to "empty".

   (c) Copy *newA*, which now contains the shuffled elements, back to A.

   Determine the expected number of random integers j that will be generated to produce $newA[k]$.

6. **Find the joker**

   (a) Suppose you take an ordinary deck of 52 playing cards, and add a single joker. If you shuffle it and turn up the cards one at a time until the joker appears, on average how many cards are required until you see the joker?

   (b) Now take a deck of 52 playing cards, add two jokers, shuffle, and turn up cards one at a time until the first time that a joker appears. On average, how many cards are required until you see the first joker?

7. **Games**

   Consider the following game: Alice and Bob will each roll a fair, six-sided die. If Alice's die comes up with a number higher than Bob's, Alice wins \$3 from Bob. If Bob's number comes up higher, or if they tie, Bob wins \$2 from Alice. Is this game a good deal for Alice? Explain.

8. **Variance**

   You have a die which has one side with a 0, one side with a 2, and four sides with 1s. (So the six sides are 0,1,1,1,1,2.) You roll the die twice.
   Let X be the **product** of the two rolls.

   a. Compute $E[X]$.

   b. Compute $Var[X]$.

9. **St. Petersburg Paradox**
   Toss a fair coin repeatedly until it comes up heads; then stop. If it first comes up heads on the i-th toss, you win $\$2^i$. Let X denote how many dollars you win after playing this game once. Calculate $\mathbf{E}[X]$.

10. **Chopping up DNA**

    (a) In a certain biological experiment, a piece of DNA consisting of a linear sequence (or string) of 4000 nucleotides is subjected to bombardment by various enzymes. The effect of the bombardment is to randomly cut the string between pairs of adjacent nucleotides: each of the 3999 possible cuts occurs independently and with probability $\frac{1}{500}$. What is the expected number of pieces into which the string is cut? What is the variance? Justify your calculation.

    (b) Suppose that the cuts are no longer independent, but highly correlated: when a cut occurs in a particular location, nearby locations are much more likely to be cut as well. The probability of each individual cut remains $\frac{1}{500}$. Does the expected number of pieces increase, decrease, or stay the same? Justify your answer with a precise explanation. Can you say the same about variance?

11. **Random variables modulo p**

    Let the random variables X and Y be distributed independently and uniformly at random in the set $\{0,1,\ldots,p-1\}$, where $p > 2$ is a prime.

    (a) What is the expectation $\mathbf{E}[X]$?

    (b) Let $S = (X+Y) \bmod p$ and $T = XY \bmod p$. What are the distributions of S and T?

    (c) What are the expectations $\mathbf{E}[S]$ and $\mathbf{E}[T]$?

    (d) By linearity of expectation, we might expect that $\mathbf{E}[S] \equiv (\mathbf{E}[X] + \mathbf{E}[Y]) \pmod{p}$. Explain why this does not hold in the present context; i.e., why does the value for $\mathbf{E}[S]$ obtained in part (b) not contradict linearity of expectation?

    (e) Since X and Y are independent, we might expect that $\mathbf{E}[T] \equiv \mathbf{E}[X]\mathbf{E}[Y] \pmod{p}$. Does this hold in this case? Explain why/why not?

12. **Random Graph**

    You create a graph on n nodes by adding edge $[i,j]$, for any two of nodes i,j, with probability p.

    (a) What is the probability that the graph has no edge? That it is the complete graph? (In this and the following two questions, give your answer as an expression, however complicated, involving n and/or p.)

(b) What is the probability that the nodes $5,7$, and 9 will form a triangle (i.e., that all three edges are present)?

(c) What is the expected number of triangles in the graph?

(d) Now you are told that $n = 10$, that $p \leq .5$, and that the variance of the number of edges in the graph is 10.8. What is p? (In this and the next question your answer can be an expression involving real numbers.)

(e) Use Chebyshev's inequality to bound the probability that the graph has at least 40 edges.

13. **The martingale**

    Consider a *fair game* in a casino: on each play, you may stake any amount \$$S$; you win or lose with probability $\frac{1}{2}$ each (all plays being independent); if you win you get your stake back plus \$$S$; if you lose you lose your stake.

    (a) What is the expected number of plays before your first win (including the play on which you win)?

    (b) The following gambling strategy, known as the "martingale," was popular in European casinos in the 18th century: on the first play, stake \$1; on the second play \$2; on the third play \$4; on the kth play \$$2^{k-1}$. Stop (and leave the casino!) when you first win.

    Show that, if you follow the martingale strategy, and assuming you have unlimited funds available, you will leave the casino \$1 richer with probability 1.

    (c) To discover the catch in this seemingly infallible strategy, let X be the random variable that measures your maximum loss before winning (i.e., the amount of money you have lost *before* the play on which you win). Show that $\mathbf{E}[X] = \infty$. What does this imply about your ability to play the martingale strategy in practice?

14. **The evolution of a social network**

    (We give a simplified analysis of the connectivity of a social network.)

    Say one person in a class of n people knows a secret, perhaps where the midterm is. Occasionally a randomly chosen person A **who doesn't know the secret** calls a randomly chosen person B ($B \neq A$) and learns the secret if B knows it.

    Let X_2 be a random variable that represents the number of calls (no two calls are simultaneous) until two people know the secret.

    (a) What is the distribution of X_2?

    (b) What is $E[X_2]$?

    (c) Let X_i be the number of calls needed to go from $i-1$ people knowing the secret to i people. What is $E[X_i]$?

    (d) What is the expected time for everyone to know the secret?

    (e) Bound your expression to within a constant factor for large n. Your expression should not have a summation. (You may use $\Theta(\cdot)$ notation, recall that $2n^2 - 5n + 2 = \Theta(n^2)$.)

15. **Independent Poisson Variables**

    Suppose you see two kinds of blips, red and blue. The blue blips are Poisson with intensity α and the red blips are independent and Poisson with intensity β. Suppose you are color blind and see all blips as being of a single kind. Show that the blips you see are Poisson distributed with intensity $\alpha + \beta$.

16. **Memorylessness**

We begin by proving two very useful properties of the exponential distribution. We then use them to solve a problem about the expected life of a package of batteries.

(a) Let r.v. X have geometric distribution with parameter p. Show that, for any positive integers m, n, we have

$$\Pr[X > m+n \mid X > m] = \Pr[X > n].$$

NOTE: This is called the "memoryless" property of the geometric distribution, because it says that conditioning on the past does not change the future distribution.

(b) Let r.v. X have exponential distribution with parameter λ. Show that, for any positive s, t, we have

$$\Pr[X > s+t \mid X > t] = \Pr[X > s].$$

NOTE: This is the memoryless property of the exponential distribution.

(c) Let r.v.'s X_1, X_2 be independent and exponentially distributed with parameters λ_1, λ_2. Show that the r.v. $Y = \min\{X_1, X_2\}$ is exponentially distributed with parameter $\lambda_1 + \lambda_2$. [Hint: work with cdf 's.]

(d) You have a digital camera that requires two batteries to operate. You purchase n batteries, labeled $1, 2, \ldots, n$, each of which has a lifetime that is exponentially distributed with parameter λ and is independent of all the other batteries. Initially you install batteries 1 and 2. Each time a battery fails, you replace it with the lowest-numbered unused battery. At the end of this process you will be left with just one working battery. What is the expected total time until the end of the process? Justify your answer.

(e) In the scenario of part (d), what is the probability that battery i is the last remaining working battery, as a function of i?