Please simplify your answers to the extent reasonable without a calculator, show your work, and explain your answers, concisely. If you set up an integral or a sum that you cannot evaluate, leave it as it is; and if the result is needed for the next part, say how you would use the result if you had it.

1. Suppose Bob is trying to guess a specific natural number \( x^* \in \{1, \ldots, N\} \). On his first guess he chooses a number \( X_0 \) uniformly at random. For \( t \in \mathbb{N} \), if \( X_t = x^* \) the game ends; if \( X_t \neq x^* \), he guesses \( X_{t+1} \) uniformly at random from among the numbers different from \( X_t \).
   a. [5 points] What is the expected number of guesses it takes Bob to find \( x^* \)?
   b. [5 points] Suppose Bob has a bad memory and can’t remember the number he guessed previously, so that he guesses \( X_{t+1} \) uniformly at random from among \( \{1, \ldots, N\} \). In this case what is the expected number of guesses it takes him to find \( x^* \)?
   c. [5 points] Suppose Bob has a good memory, and at each step guesses uniformly at random among the numbers he has not guessed at any previous step. Now what is the expected number of guesses it takes him to find \( x^* \)?

2. [15 points] \( 2 \leq n \in \mathbb{N} \) is a prime number if its only divisors are 1 and itself. The Prime Number Theorem says that the primes are distributed approximately as if they came from an inhomogeneous Poisson process \( P(x) \) with intensity \( \lambda(x) = 1/\ln x \). [We have to say approximately since (1) the primes are integers, not general real numbers, and (2) the primes take determined, not random, values.] Use this theorem to estimate the number of primes in the interval \( [2, N] \).

3. Suppose we flip a fair coin repeatedly.
   a. [5 points] What is the expected number of flips until we see Head followed by Tail?
   b. [5 points] What is the expected number of flips until we see Head followed by Head?
   c. [5 points] Give an intuitive explanation for why your answers in (a) and (b) are the same or different.

4. [15 points] Here is a list of the 23 prime numbers between 3 and 100, together with their remainders when divided by 3:

\[
\begin{array}{cccccccccccccccc}
\end{array}
\]

\[
\begin{array}{cccccccccccccccc}
2 & 1 & 2 & 1 & 2 & 1 & 2 & 2 & 1 & 1 & 2 & 1 & 2 & 2 & 1 & 1 & 2 & 1 & 1 & 2 & 1 & 1 & 1 & 2 & 2 & 1 \\
\end{array}
\]

In this list 1 follows 1 three times; 2 follows 1 seven times; 1 follows 2 eight times; and 2 follows 2 four times, so we might imagine that the sequence of remainders of prime numbers divided by 3 are the outcome of a Markov process with transition matrix

\[
P = \frac{1}{2} \begin{pmatrix} 1 & 2 \\ 3/10 & 7/10 \\ 2/3 & 1/3 \end{pmatrix}.
\]

If this were true, what fraction of all prime numbers would we expect to have remainder 1 when divided by 3?
5. Let $X(t)$ be a Poisson process on $\mathbb{R}_{\geq 0}$, with intensity $\lambda$. Suppose $Y_i \sim \text{Poisson}(\mu)$ are independent of one another and of $X(t)$. Let

$$Y(t) = \sum_{i=1}^{X(t)} Y_i.$$ 

a. [6 points] What is $E[Y(t)]$? What is $\text{Var}[Y(t)]$?

b. [6 points] What is $E[Y(t)|X(t) = n]$? What is $\text{Var}[Y(t)|X(t) = n]$?

c. [4 points] Is $Y(t)$ a Poisson process? Why or why not?

d. [4 points] Give a different probability function for $Y_i$ that makes $Y(t)$ a Poisson process.

6. There is a 500 m. $\times$ 1000 m. plot of land on Barro Colorado Island in Panama where a careful census has been made of the trees. The locations of trees of one common species, *Alseis blackiana*, are indicated by dots in the graphic below, where the size of each dot represents the size of the tree.

a. [10 points] Do you think the locations of these trees should be modeled as arising from a Poisson process? Why or why not?

b. [10 points] There are 7599 dots in this graphic. If we partition the plot into four congruent rectangles, by dividing it in half vertically and horizontally, the number of dots in each rectangle is 2115, 1950, 1708 and 1826, moving counterclockwise from the northeast rectangle. What is the probability of observing this distribution if the tree locations arise from a homogeneous Poisson process?
7. Let $Y_1, Y_2 \in \{0, 1\}$ be Bernoulli random variables. Suppose

$$\Pr(Y_2 = 0 \mid Y_1 = 0) = r \quad \Pr(Y_2 = 1 \mid Y_1 = 1) = s$$

and $\Pr(Y_1 = 1) = p$.

a. [4 points] What is the joint probability function for $Y_1$ and $Y_2$?

b. [4 points] For what values of $r$, $s$ and $p$ are $Y_1$ and $Y_2$ independent?

c. [4 points] What is $\Pr(Y_2 = 1)$?

d. [4 points] For what values of $r$, $s$ and $p$ is $\Pr(Y_2 = 1) = \Pr(Y_1 = 1)$?

8. Consider a sequence of customers entering a store. Let $Y_i \in \{0, 1\}$ denote the number of items the $i^{th}$ customer buys, for $0 < i \in \mathbb{N}$. Suppose every customer, after the first, sees what the previous customer does; if the previous customer bought nothing, the current customer also buys nothing, with probability $r$, and if the previous customer bought an item, the current customer does too, with probability $s$.

a. [4 points] Suppose $0 < r = s < 1$. Without doing any calculation, explain what is the fraction of customers who buy an item in the infinite number of customers limit.

b. [4 points] For general $0 < r, s < 1$, what is $\lim_{n \to \infty} \Pr(Y_n = 1)$?

c. [8 points] Still for general $0 < r, s < 1$, what is $\lim_{n \to \infty} \Pr(Y_n = 1 \text{ and } Y_{n+1} = 1)$?

9. [16 points] Now suppose people enter a store according to a Poisson process with rate $\lambda$. The store owner wants to know if someone buying something makes the next person more likely to buy something (as it did in problem 2 for $s > 1/2$). To help answer this question, suppose each person who enters the store buys something with probability $p$, independently of what anyone else does. Let $C(t)$ be the number of people who buy something and are followed by the next customer also buying something, both before time $t$. What is $\mathbb{E}[C(t)]$?

Extra Credit. [5 points] Suppose instead of the people’s purchasing decisions being independent, they are dependent as in problem 2. Now what is $\mathbb{E}[C(t)]$?
10. Let \( \{X(t) \mid t \in \mathbb{R}_{\geq 0}\} \) be a Poisson point process with rate \( \lambda \) on \( \mathbb{R}_{\geq 0} \). For each point \( 0 < i \in \mathbb{N} \) of the process, let \( D_i \) be the distance to its nearest neighbor.

a. [3 points] Write \( D_i \) in terms of the sojourn times \( \{S_j \mid j \in \mathbb{N}\} \).

b. [3 points] Are the \( \{D_i \mid 0 < i \in \mathbb{N}\} \) independent?

c. [10 points] What is the probability density function for each \( D_i \)?

11. [20 points] An ant starts at one vertex of a cube and at each time step walks along an edge to an adjacent vertex. What is the probability the ant returns to its original vertex before reaching the opposite vertex?

12. Let \( X, Y \) be random variables with a bivariate normal distribution such that \( \mathbb{E}[Y] = 0 \) and \( \text{Var}[X] = 1 = \text{Var}[Y] \). Suppose \( \mathbb{E}[X \mid Y = y] = 2 - y/3 \).

a. [6 points] What is \( \mathbb{E}[X] \)?

b. [10 points] What is \( \text{Cov}[X, Y] \)?