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 \(x^*\)?

We can describe this as a Markov process on two states \(x^*\) and \(\bar{x}^* = \{1, \ldots, N\} \setminus \{x^*\}\), with transition probability matrix:

\[
P = \begin{pmatrix}
1 & 0 \\
\frac{1}{(N-1)} & \frac{(N-2)/(N-1)}
\end{pmatrix}.
\]

Let \(\nu\) be the expected number of steps it takes to reach \(x^*\) from \(\bar{x}^*\). Then from \(P\) we have

\[
\nu = 1 + \frac{N-2}{N-1} \nu,
\]

which we can solve to find \(\nu = N - 1\). Then the expected number of guesses it takes Bob to find \(x^*\) is

\[
\frac{1}{N} \cdot (0 + 1) + \frac{N-1}{N} \cdot (N-1 + 1) = N - 1 + \frac{1}{N}.
\]

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 \(x^*\)?

Bob’s bad memory changes the transition probability matrix to:

\[
P = \begin{pmatrix}
1 & 0 \\
\frac{1}{N} & \frac{(N-1)/N}{N}
\end{pmatrix}.
\]

Now

\[
\nu = 1 + \frac{N-1}{N} \nu,
\]

so \(\nu = N\) and the expected number of guesses is

\[
\frac{1}{N} \cdot (0 + 1) + \frac{N-1}{N} \cdot (N + 1) = N.
\]
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 $x^*$?

Bob has equal probability $1/N$ of guessing correctly on any of guesses 1 through $N$, so the expected number of guesses in this case is

$$\sum_{k=1}^{N} k \cdot \frac{1}{N} = \frac{N(N-1)}{2} \cdot \frac{1}{N} = \frac{N-1}{2}.$$ 

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 $[2, N]$.

Since $P(x)$ is an inhomogeneous Poisson process, the expected number of events in the interval $[2, N]$ is

$$E[P(N)] - E[P(2)] = \int_{2}^{N} \frac{1}{\ln x} \, dx.$$ 

This estimates the number of primes in this interval, and is approximately $N/\ln N$.

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?

We can think of this as a Markov process on three states: $H$, meaning the most recent flip was Head; $HT$ meaning the most recent two flips were Head then Tail; and $0$ meaning anything else. Then the transition probability matrix is

$$P = \begin{pmatrix}
0 & H & HT \\
0 & 1/2 & 1/2 & 0 \\
HT & 0 & 1/2 & 1/2 \\
0 & 0 & 1
\end{pmatrix}.$$ 

Using $\nu_{HT} = E[\# \text{flips}|HT] = 0$, from $P$ we have

$$\nu_0 = E[\# \text{flips}|0] = 1 + \frac{1}{2} \nu_0 + \frac{1}{2} \nu_H,$$

$$\nu_H = E[\# \text{flips}|H] = 1 + \frac{1}{2} \nu_H.$$ 

Solving this system of linear equations gives $\nu_0 = 4$. 


b. [5 points] What is the expected number of flips until we see Head followed by Head? Again this a Markov process on three states; now HH, meaning the most recent two flip were both Heads; H meaning only the most recent flips was Head; and 0 meaning anything else. Then the transition probability matrix is

\[
P = \begin{pmatrix}
0 & H & HH \\
0 & 1/2 & 1/2 & 0 \\
H & 1/2 & 0 & 1/2 \\
HH & 0 & 0 & 1
\end{pmatrix}.
\]

Using \( \nu_{HH} = E[\#\text{flips}|HH] = 0 \), from \( P \) we have

\[
\nu_0 = E[\#\text{flips}|0] = 1 + \frac{1}{2} \nu_0 + \frac{1}{2} \nu_H,
\]

\[
\nu_H = E[\#\text{flips}|H] = 1 + \frac{1}{2} \nu_0.
\]

Solving this system of linear equations gives \( \nu_0 = 6 \).

c. [5 points] Give an intuitive explanation for why your answers in (a) and (b) are the same or different.

In the Markov chain in case (a), once it reaches state H it cannot return to 0, but in case (b), it can. This makes the expected number of steps to reach the absorbing state longer in case (b).

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

5  7  11  13  17  19  23  29  31  37  41  43  47  53  59  61  67  71  73  79  83  89  97
2  1  2  1  2  1  2  2  1  1  2  1  2  2  1  1  2  1  1  2  2  1

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?

\( P \) is regular, so the limiting distribution for this Markov process is its stationary state, i.e., its left eigenvector with eigenvalue 1: \( (p \quad 1-p) \ P = (p \quad 1-p) \), so

\[
\frac{3}{10}p + \frac{2}{3}(1-p) = p,
\]

which implies \( p = 20/41 \); this is the fraction of all primes we would expect to have remainder 1 when divided by 3, if this were a good model for the prime numbers.
In fact, a generalization of the theorem mentioned in problem 2, the Prime Number Theorem for arithmetic sequences, says that in the limit as \( N \to \infty \), \( \frac{1}{2} \) of prime numbers less than \( N \) have remainder 1, and half have remainder 2, when divided by 3. But a recent paper by Robert J. Lemke Oliver and Kannan Soundararajan, “Unexpected biases in the distribution of consecutive primes”, arXiv:1603.03720 [math.NT], shows that even for very large \( N \), successive primes more often have different than the same remainders when divided by 3. They show that a conjecture of Hardy and Littlewood implies that this bias goes away, slowly, as \( N \to \infty \), but that conjecture remains unproved.

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 \( \mathbb{E}[Y(t)] \)? What is \( \text{Var}[Y(t)] \)?
\[
\mathbb{E}[Y(t)] = \lambda t \cdot \mu. \quad \text{Var}[Y(t)] = \lambda t \cdot \mu + \mu^2 \cdot \lambda t.
\]

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

c. [4 points] Is \( Y(t) \) a Poisson process? Why or why not?
No. One way to see this is that \( \mathbb{E}[Y(t)] \neq \text{Var}[Y(t)] \), so \( Y(t) \) is not a Poisson random variable. Another way to see it is that \( Y(t) \) does not make jumps of size 1 the way a Poisson process does.

d. [4 points] Give a different probability function for \( Y_i \) that makes \( Y(t) \) a Poisson process.
The only possibility is \( Y_i \sim \text{Bernoulli}(p) \).
6. There is a 500\text{m.} \times 1000\text{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.

![Tree Locations Graphic](image)

a. [10 points] Do you think the locations of these trees should be modeled as arising from a Poisson process? Why or why not?
I think not. At best it appears the Poisson process would be quite inhomogeneous. Also, since a tree cannot grow within another tree, the number of trees within the radius of a tree of its center does not follow a pre-specified, even inhomogeneous, Poisson distribution. Finally, although it is maybe hard to see this without doing some calculations, the numbers of trees in disjoint regions seem not to be independent.

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?
Conditioning on there being 7599 dots, a homogeneous Poisson process restricted to this rectangle becomes 7599 samples of a uniform process on the rectangle. Since there is probability \( \frac{1}{4} \) for each sample to fall into each subrectangle, the probability of seeing this distribution is

\[
\frac{7599!}{2115!1950!1708!1826!} \left( \frac{1}{4} \right)^{7599} \approx 6.5 \times 10^{-17}.
\]

Even without a calculator, we can see that this will be a very small number because the numbers of trees in each subrectangle are so different; if this were really 7599 samples from a uniform distribution it is much more probable that they all would be close to 1900.