1. Coin Toss Bernoulli Process

(a)

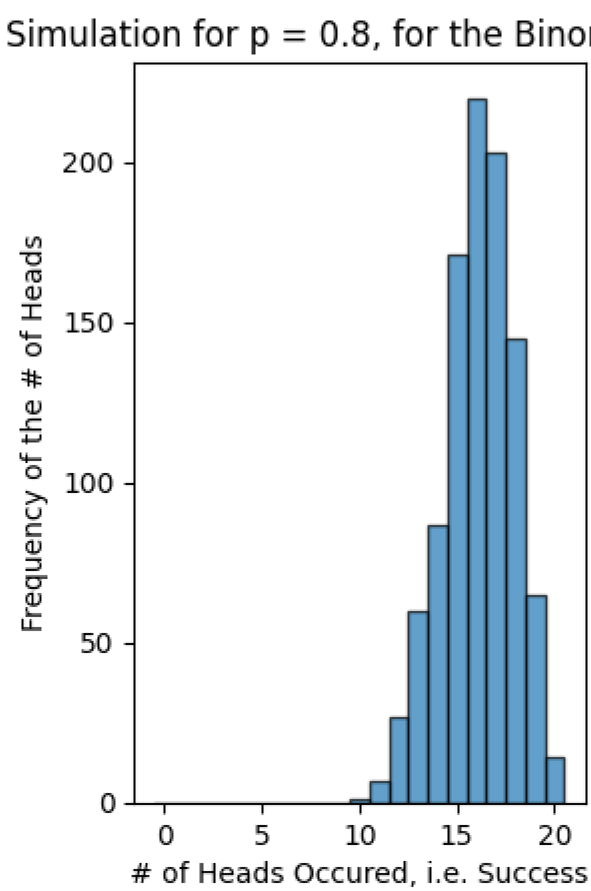
The Bernoulli process with (p = 0.8) means that we have a higher probability of success in each trial. This implies that we expect more "heads" when a coin is tossed. In 1000 simulations of tossing the coin 20 times each, we would expect to see more successes.

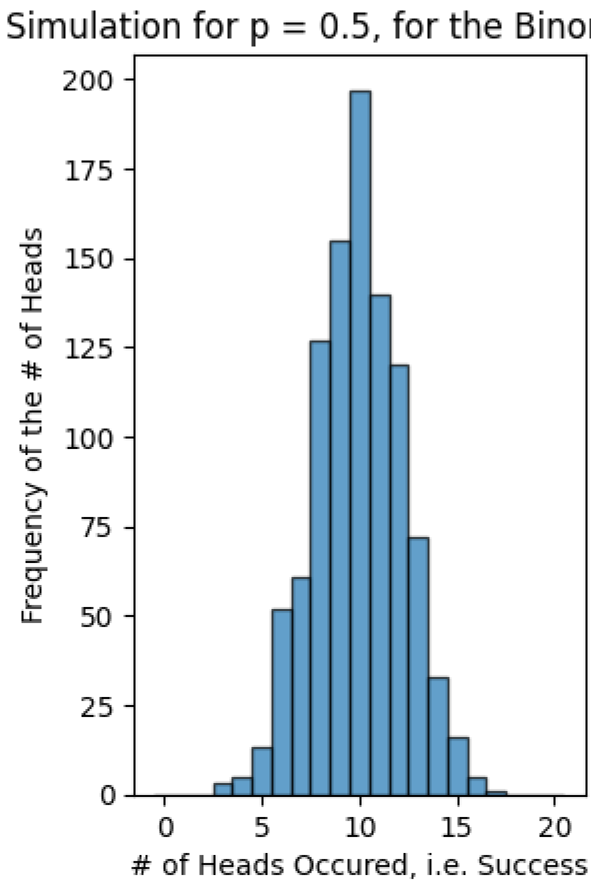
(b)

For (p = 0.5), the probability of success (getting "heads") is exactly 50%, which is equivalent to a fair coin toss. In 1000 simulations, we would expect an equal number of "heads" and "tails" on average.

Therefore, for part (a), where p = 0.8, you will see a histogram with more successes on average compared to part (b), where p = 0.5.

This is because a higher value of p leads to a higher probability of success in each trial.





Inference from the plots -

The average number of heads, based on performing the experiment 1000 times, corresponds to n * p, where,

n <- # of coin tosses (viz. 20 in our case, as per the question)
p <- Probability of success (viz. getting a head in our case, as per question)

This follows from the fact that # of bernoulli success, being a RV, follows a binomial distribution, & has an expected value = n * p.

Therefore, for part (a),
the average # of heads = 0.8 * 20 = 16 (can be inferred as an approximation, from the plot)

& for part (b),
the average # of heads = 0.5 * 20 = 10 (can be inferred approximately, from the plot)

Theoretical Reasoning -

Recall that the Moment Generating Function (MGF) of a random variable X is defined as:

$$M_X(t) = E[e^{tX}]$$

If (X) and (Y) are independent random variables, then:

$$M_{X+Y}(t) = M_X(t) * M_Y(t)$$

Now, let's consider (n) independent Bernoulli random variables $(X_1, X_2, \ldots, X_n)$ with the same probability of success (p). The MGF of a Bernoulli random variable (X_i) is:

$$M_{X_i}(t) = E[e^{tX_i}] = pe^t + (1-p)$$

The MGF of the sum $(Y = X_1 + X_2 + \ldots + X_n)$ is:

$$M_Y(t) = M_{X_1}(t) * M_{X_2}(t) * \ldots * M_{X_n}(t) = (pe^t + (1-p))^n$$

This MGF corresponds to the MGF of a Binomial random variable with parameters (n) and (p).

This confirms that the MGF of the sum of (n) independent Bernoulli random variables is the same as the MGF of a Binomial random variable with parameters (n) and (p).

2. Hospital Emergency Room Poisson Process

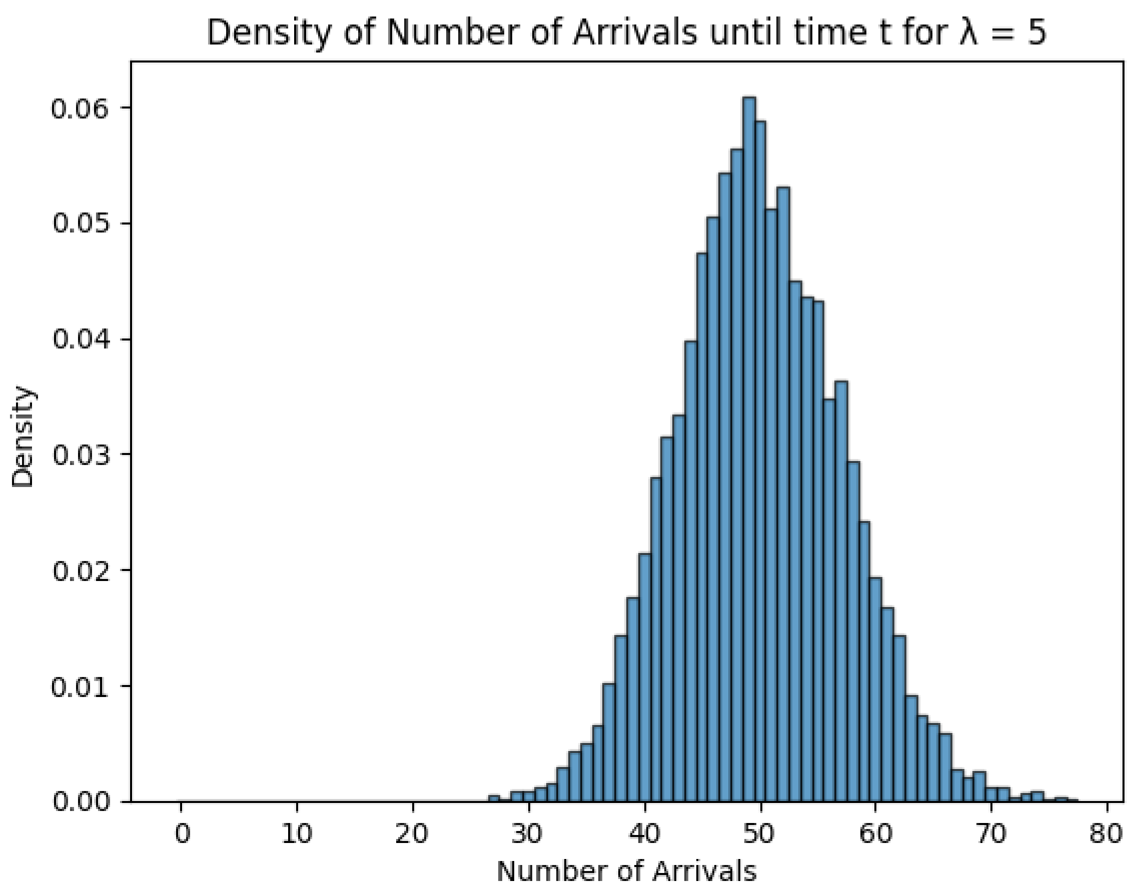
(a)

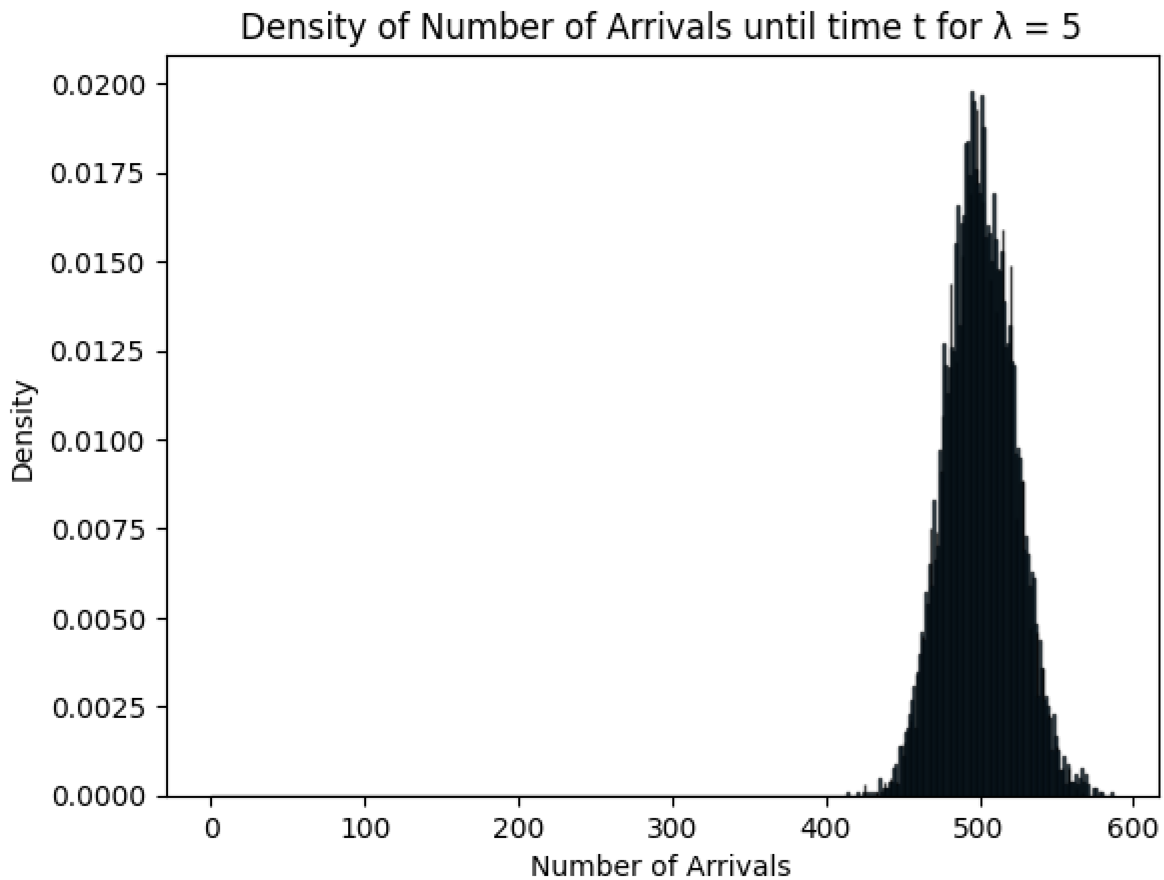
For a Poisson process with a rate of 5 patients per hour, we expect an average of 5 patients to arrive in one hour. In this simulation, we're interested in the number of arrivals in a specific time interval (0, t]. We simulate this process and plot a histogram.

The mean comes out to be, $\lambda * t = 50$ (approximately).

Here, t <- time interval (viz. 10 in my implementation - can be adjusted in the code, as per requirement)

This implies that we're expecting an average of 50 patients, in 10 hours.





Similarly, here, the mean comes out to be, λ * t = 500 (approximately).

Here, t <- time interval (viz. 100 in my implementation - can be adjusted in the code, as per requirement)

This implies that we're expecting an average of 500 patients, in 100 hours.

(b)

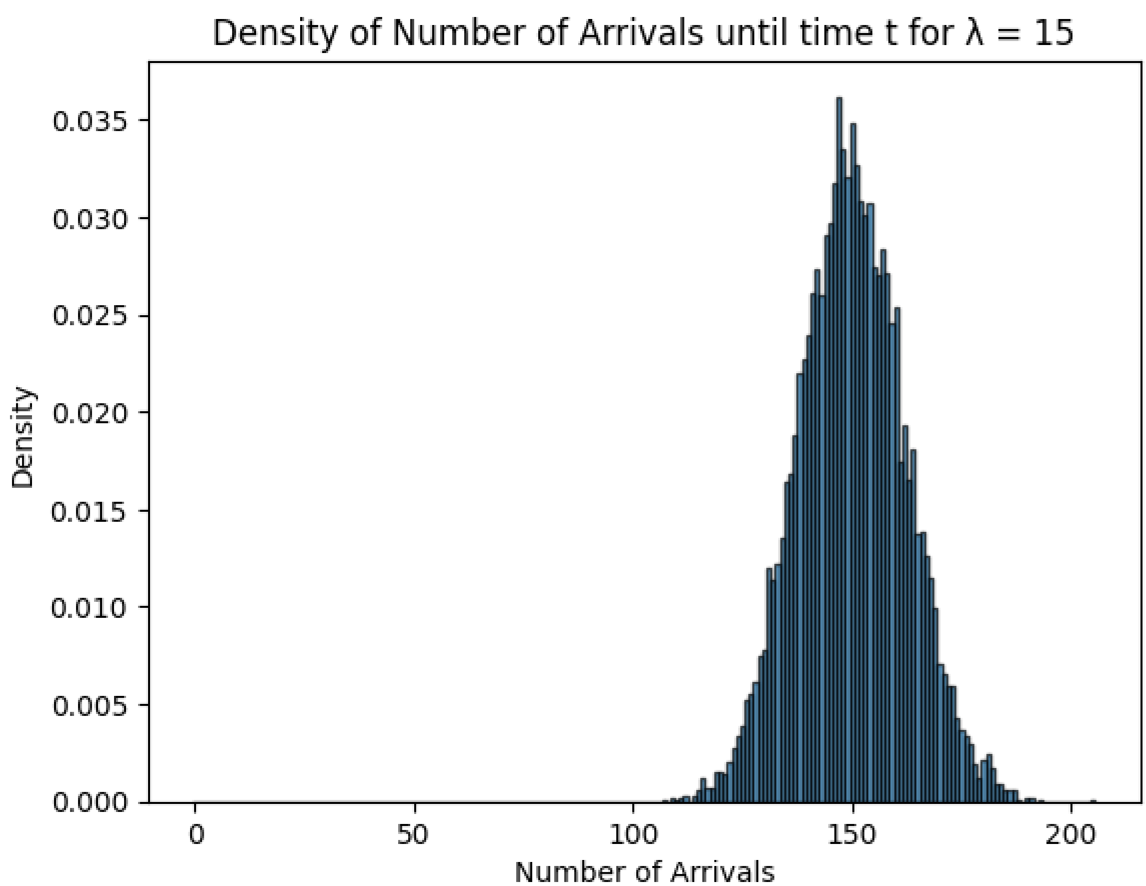
Increasing the rate parameter to (lambda = 15) means we're now expecting an average of 15 patients in an hour. This will lead to a higher expected number of arrivals in the same time interval compared to when (lambda = 5).

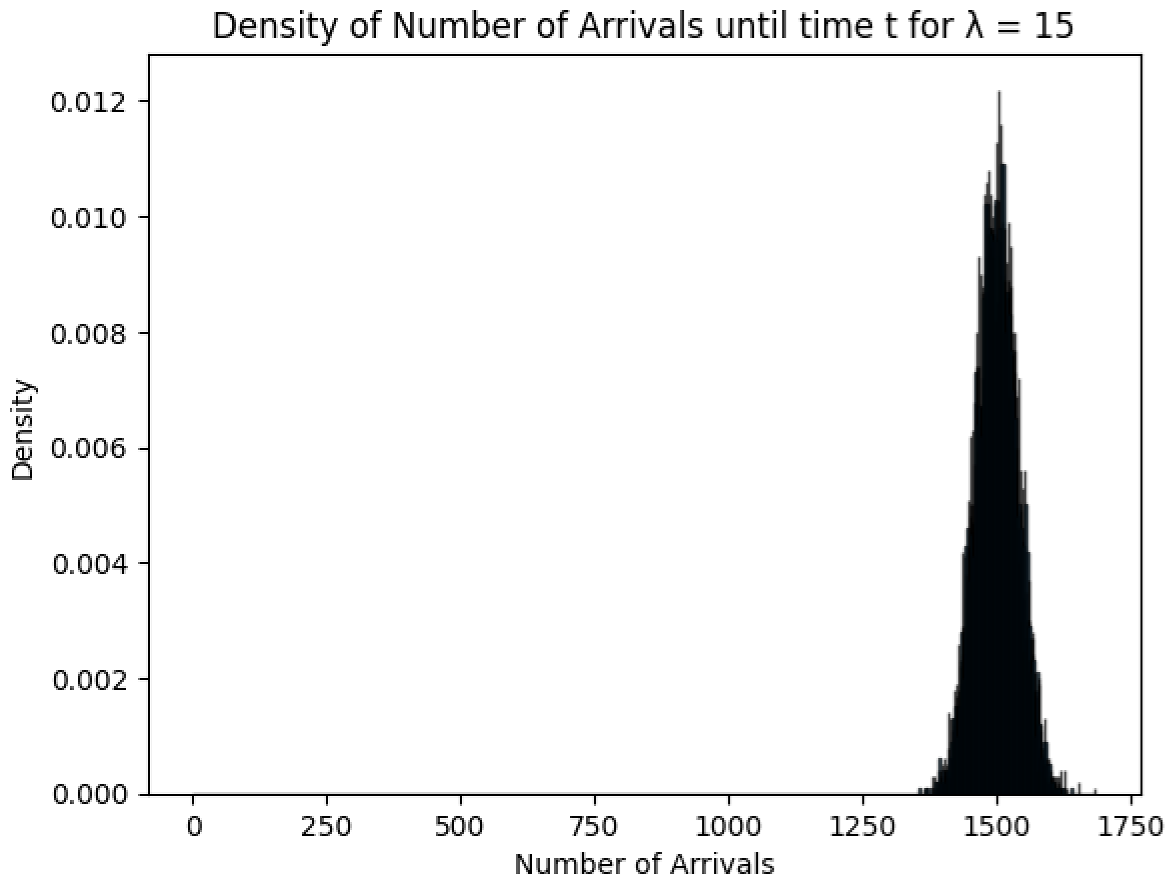
Comparing the two simulations, we find that now the average number of patients becomes 15 per hour.

Therefore, now that the simulation is being run for t = 10,
this implies that,

the average number of arrivals in time t = 10, becomes = 150

This result is, however, approximated - since there’s always a marginal error while computing probabilities on a smaller scale.





Similarly, here, the mean comes out to be, $\lambda * t = 1500$ (approximately).

Here, t <- time interval (viz. 100 in my implementation - can be adjusted in the code, as per requirement)

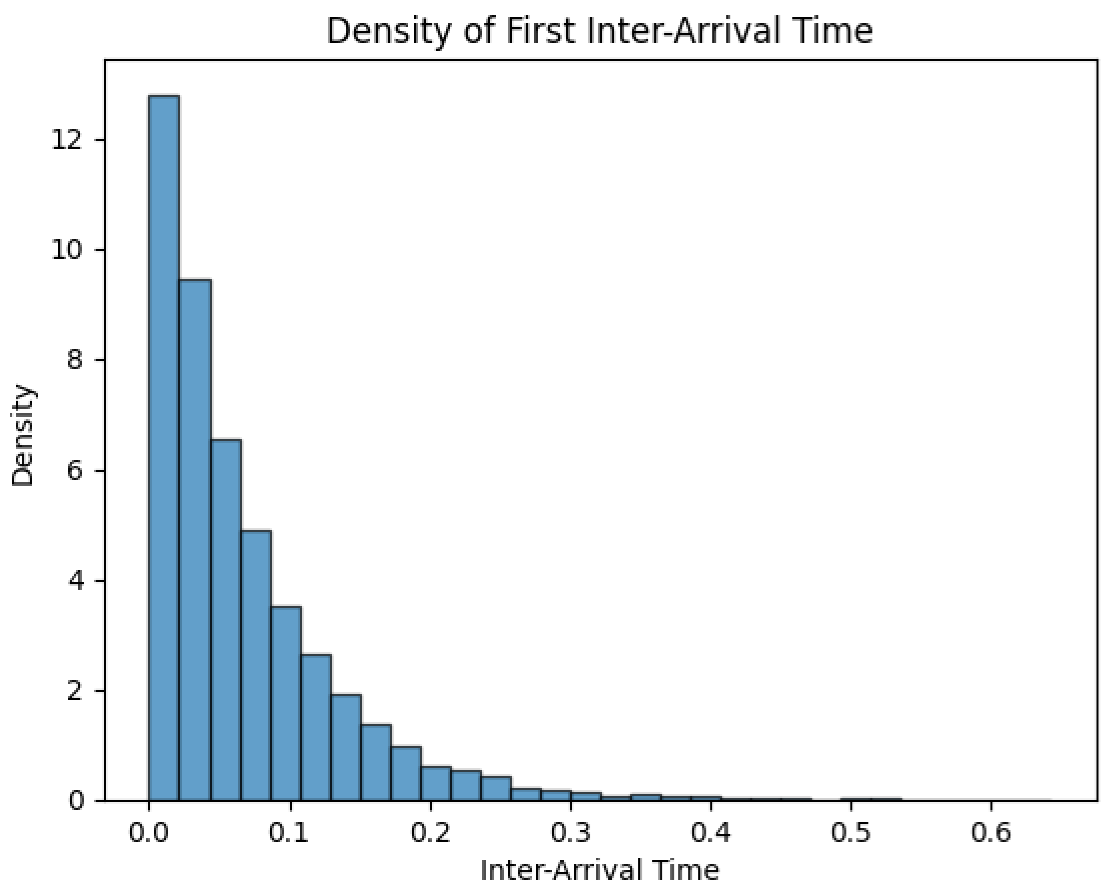
This implies that we're expecting an average of 1500 patients, in 100 hours.

(c)

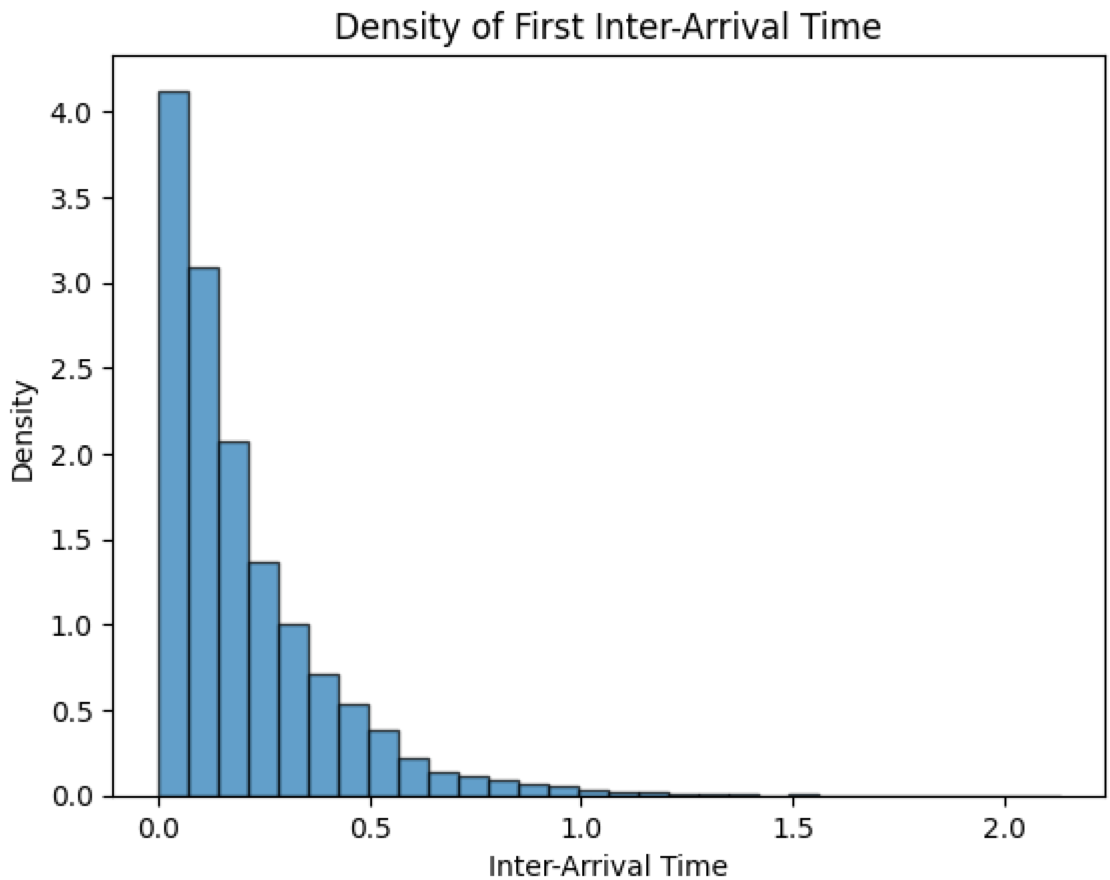
The first inter-arrival time in a Poisson process is the time until the first event (in this case, the first patient arrives). This is typically modeled by an exponential distribution. This simulation will provide insights into how long it takes, on average, for the first patient to arrive at the emergency room.

All the different graphs, per simulation, are added on the .ipynb file, viz. uploaded with the assignment.

*Note - The below plot occurs for Lambda = 15.



*Note - The below plot occurs for Lambda = 5.



*Note:

In a Poisson process, the inter-arrival times follow an exponential distribution with a mean of 1/λ, where λ is the rate of arrivals per unit time.

In the exponential distribution, the probability density function (PDF) is given by:

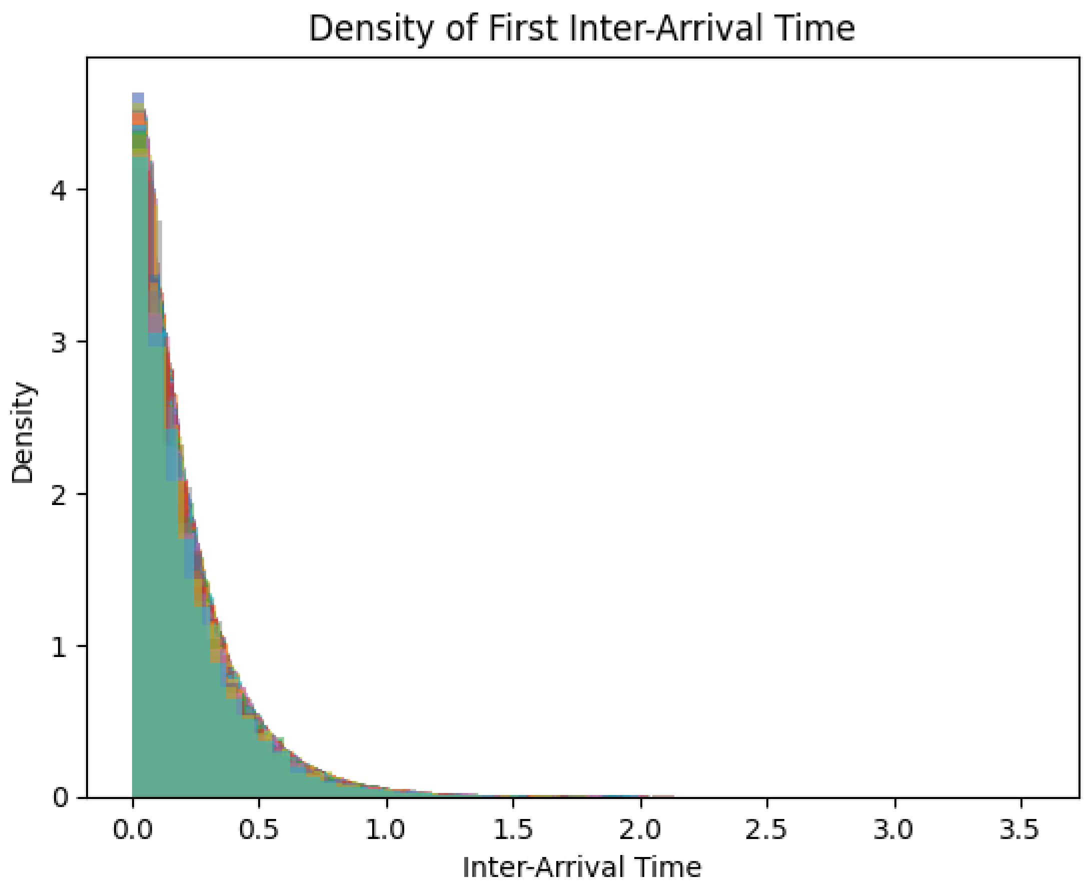
f(x) = λ * e^(-λx)

In this case, λ is the rate parameter which you've set to 4.

So, when you plot the histogram of inter-arrival times, the density value represents the probability density of observing a specific inter-arrival time. It might be around 4 because of the specific parameters you've chosen for your simulation.

Now, the below code does the exact same thing as Q2 part (c), except that it uses 'num_simulations' as a parameter too.

This plots 1000 simulations of the exponential RV on the graph, & the result is that, the graph appears smoother, & the differences for each simulation is clearly visible.



Theoretical Reasoning -

In a Poisson process, the number of arrivals in a fixed time interval follows a Poisson distribution. The mean value of the number of arrivals can be derived as follows:

Let $(N(t))$ be the number of arrivals in the time interval $((0, t])$. This is a random variable.

The probability of observing (k) arrivals in this interval is given by the Poisson probability mass function:

$P(N(t) = k) = (((\lambda t)^k) * e^{(-\lambda t)}) / (k)!$

where:
λ is the average rate of arrivals per unit time.
t is the length of the time interval.

The expected value (mean) of a Poisson random variable is given by:

$E(N(t)) = \Sigma\, k * P(N(t) = k)$ $\quad\quad$ $0 <= k <$ infinity

Substituting the Poisson probability mass function:

$E(N(t)) = \Sigma(k=0 \text{ to } \infty)\ k * ((\lambda t)^k * e^{(-\lambda t)}) / k!$

Now, let's simplify the sum:

$E(N(t)) = \Sigma(k=1 \text{ to } \infty)\ ((\lambda t)^k * e^{(-\lambda t)}) / (k-1)!$

Notice that we can perform a change of index to simplify further:

$E(N(t)) = \lambda t * \Sigma(k=0 \text{ to } \infty)\ ((\lambda t)^k * e^{(-\lambda t)}) / k!$

Recall that $\Sigma(k=0 \text{ to } \infty)\ (x^k) / k! = e^x$. Applying this property:

$E(N(t)) = \lambda t * e^{(-\lambda t)} * \Sigma(k=0 \text{ to } \infty)\ ((\lambda t)^k) / k!$

$E(N(t)) = \lambda t * e^{(-\lambda t)} * e^{(\lambda t)}$

$E(N(t)) = \lambda t$

Therefore, the mean value of the number of arrivals in a time interval $((0, t])$ in a Poisson process with rate (λ) is (λt). This is an important property of Poisson processes.

To prove that the inter-arrival times of a Poisson random variable follow an Exponential distribution, we'll proceed as follows -

Let $N(t)$ be a Poisson random variable representing the number of arrivals in the time interval $(0, t]$ with rate λ. The probability mass function (PMF) of $N(t)$ is given by:

$P(N(t) = k) = ((\lambda t)^k * e^{(-\lambda t)}) / k!$

Now, let T1 be the first inter-arrival time, which is the time until the first arrival in the process. This can be represented as $T1 = t$ where $N(t) = 1$. Therefore, the cumulative distribution function (CDF) of T1 is:

$F_{T1}(t) = P(T1 \leq t) = P(N(t) \geq 1) = 1 - P(N(t) = 0) = 1 - e^{(-\lambda t)}$

Next, we differentiate $F_{T1}(t)$ to get the probability density function (PDF) of T1, denoted as $f_{T1}(t)$:

$f_{T1}(t) = (d/dt) F_{T1}(t) = \lambda * e^{(-\lambda t)}$

This is the PDF of an exponential random variable with rate parameter λ, which is exactly the definition of an exponential distribution.

Therefore, we have shown that the first inter-arrival time T1 in a Poisson process with rate λ follows an exponential distribution with rate parameter λ. By extension, the inter-arrival times of the entire process follow the same exponential distribution.