ASSIGNMENT-1
STOCHASTIC PROCESSES AND APPLICATIONS (MTH 371)

Ques1. OBSERVATION ON MAGNITUDE OF EARTHQUAKE

There is research on earthquakes related to their magnitude and we have data from 2000-2022.So this process (research) can be modeled as Bernoulli Process from the following points given below :-

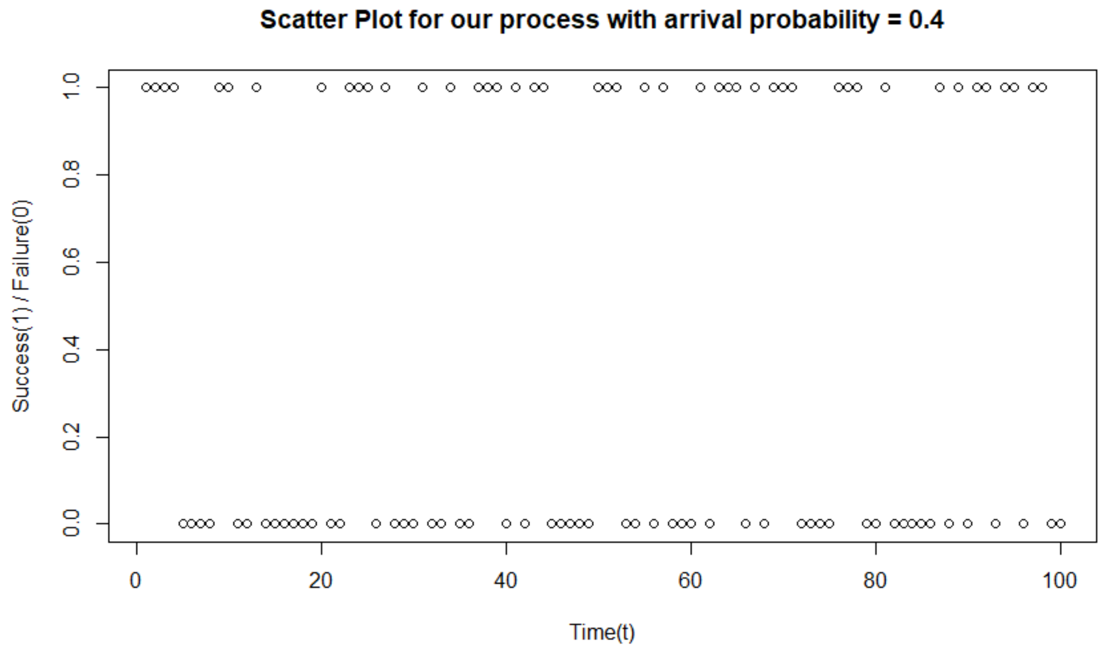
- Researchers only interested in earthquakes of magnitude >= 6. So we can classify this in failure and success which means we have only 2 possible outcomes i.e., we will take all those earthquakes of magnitude >= 6 as success ($Z_t = 1$) and others earthquakes as failure($Z_t = 0$).
- The Occurrences of earthquakes are independent of each other.
- Each earthquake of magnitude ≥ 6 occurs with the same probability p.
- Occurrences of earthquakes is discrete time process i.e., at one instance of time there will be at max 1 earthquake observed.
- Our process will start from t = 0 and the arrival of earthquakes will start from t > 0.

(a) Given :- We have Random Variable Z_k such that :

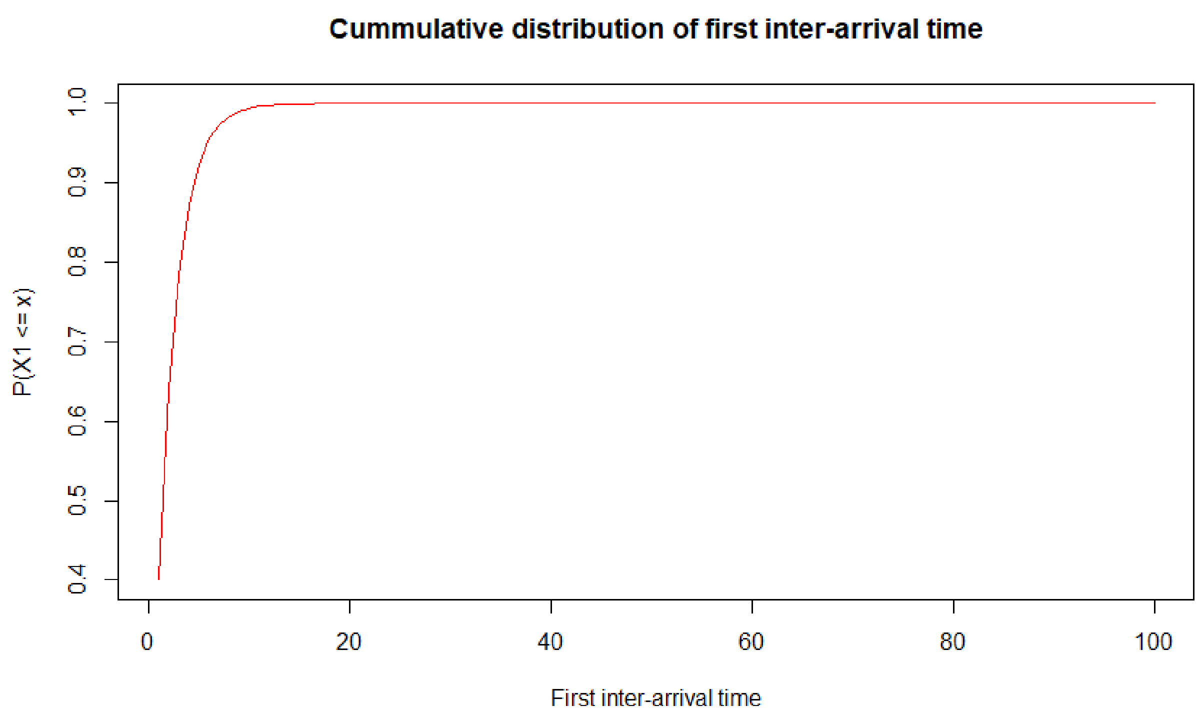
$Z_k = 1$, magnitude of earthquake >= 6
$Z_K = 0$, otherwise

Probability of occurrence of an earthquake with magnitude >= 6 is 0.4 i.e., p = 0.4,

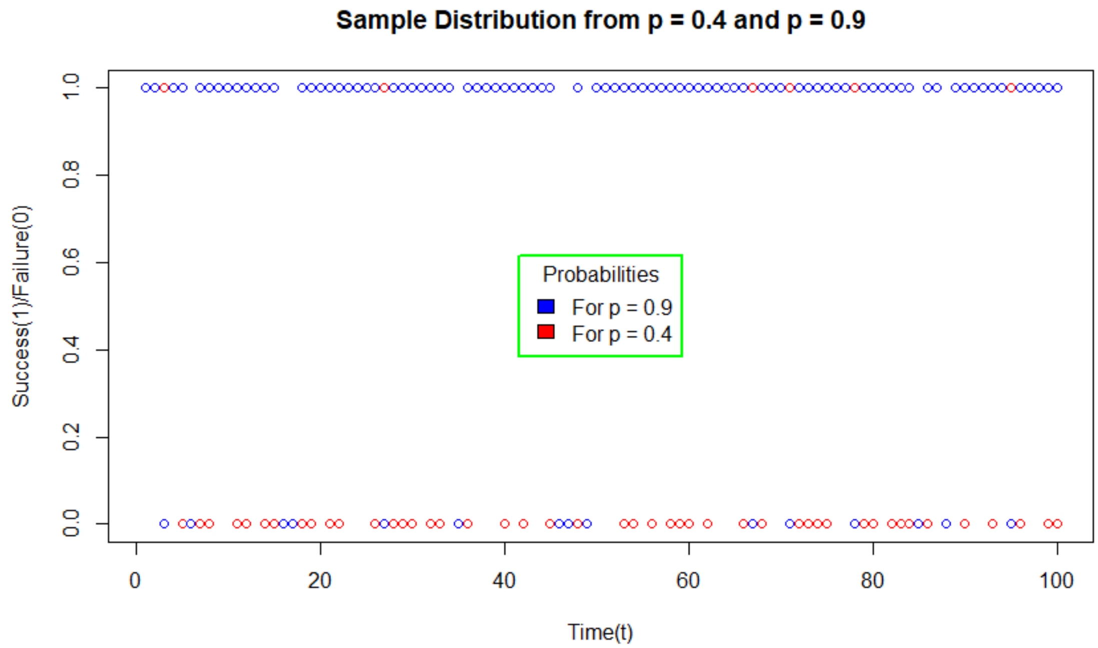
After simulating the process for t = 100 we plot a scatter plot taking x-axis as time or number of observations and y-axis taking for success or failure representation and below our plot is given,

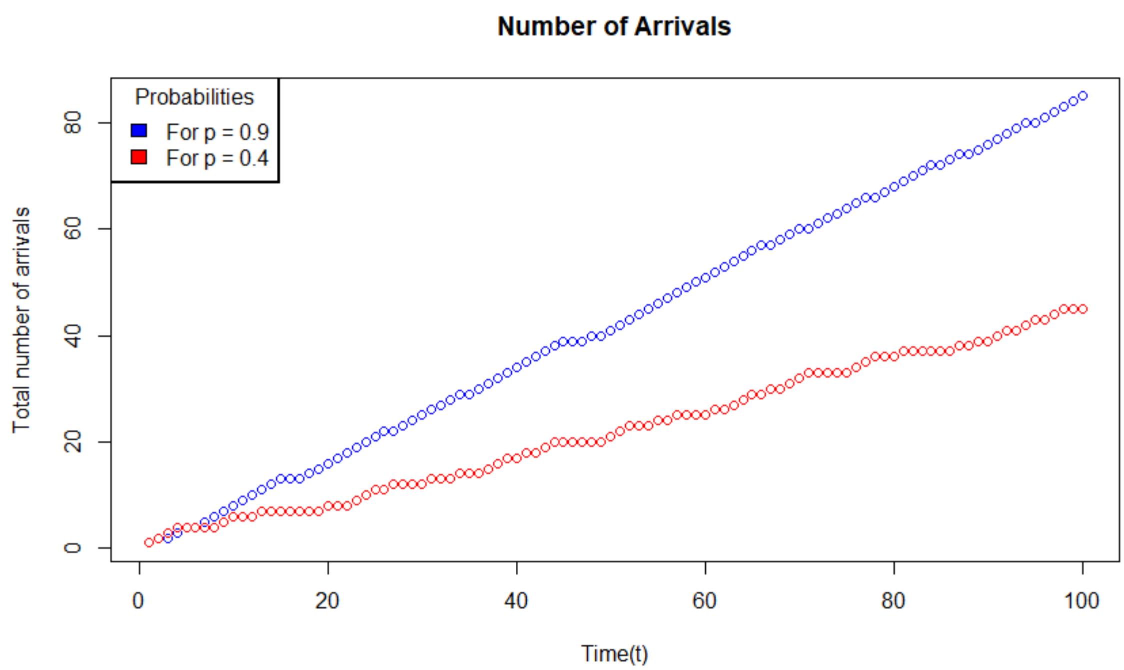


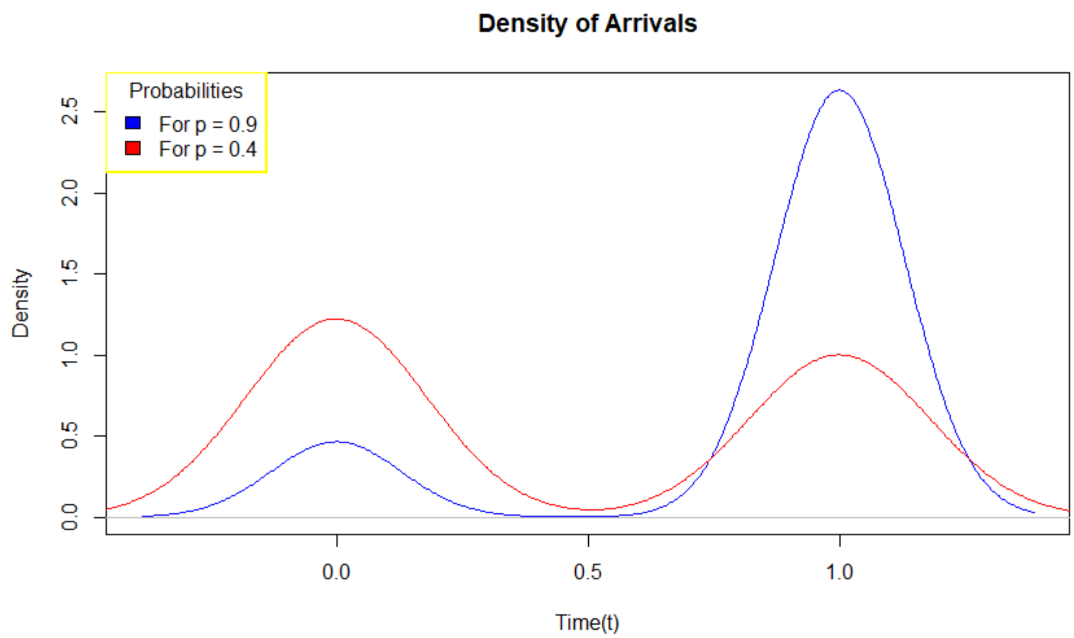
(b) We know that inter-arrival times in bernoulli process follows geometric distribution.So,we will calculate the CDF at each instance using PMF of geometric distribution and plot the CDF of first inter-arrival time and we will take y-axis to represent the CDF at any instance and x-axis to represent the number of instances or time where first inter-arrival happens which we will take till 100 instances.

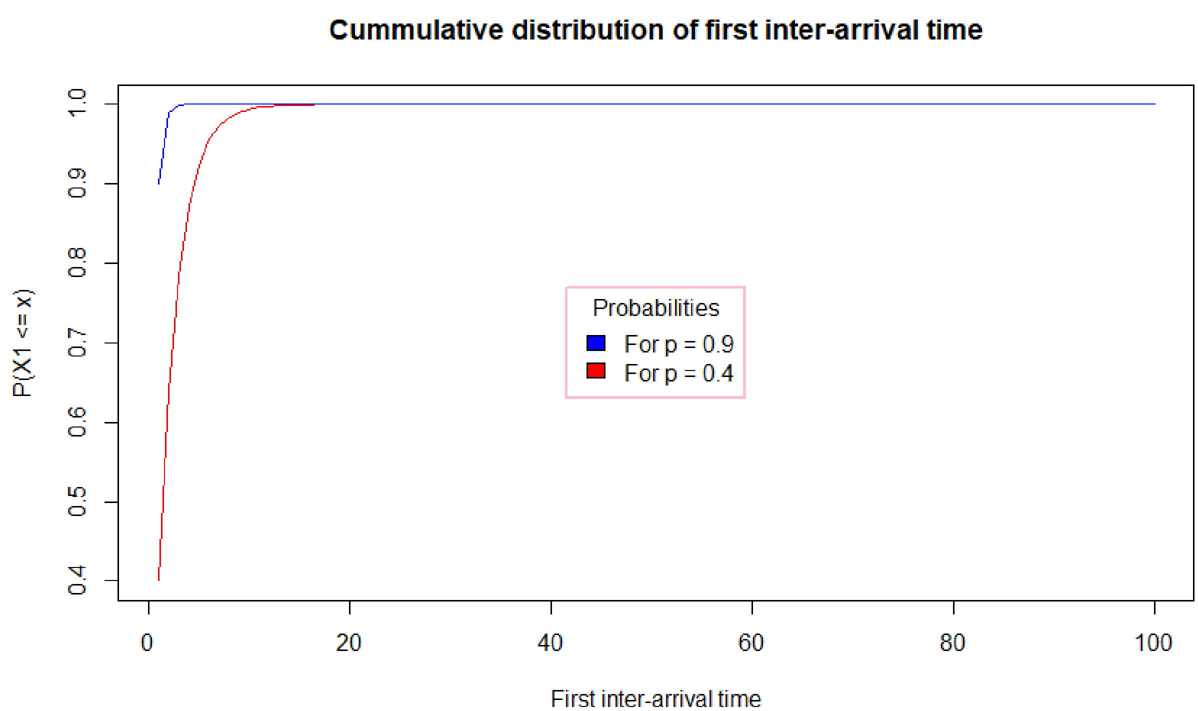


(c) Studying and Comparing change in the number of arrivals for p = 0.4 and p = 0.9 :-









Observations on number of arrivals from plot we made :-

- In the sample Distribution plot we can see that more arrivals happen in p = 0.9 than p = 0.4 because p = 0.9 signifies that we have a 90 percent chance of getting a success and in p = 0.4 we have only a 40 percent chance of getting a success.

- Total number of arrivals for p = 0.9 are more than 80 but in the case of p = 0.4, the total number of arrivals are between 40 to 50 which also shows that arrivals are faster in p = 0.9 than p = 0.4 and this is because of we have higher probability of getting an arrival in p = 0.9 than p = 0.4.

- From the density plot we can observe that success density is greater in p = 0.9 than p = 0.4 but failure density is greater in p = 0.4 than p = 0.9 and this is our desired result because in point 2 we are getting higher number of arrivals in p = 0.9 than p = 0.4.So density must be greater in p = 0.9 than p = 0.4 and same reason for higher density of failure in p = 0.4 than p = 0.9.

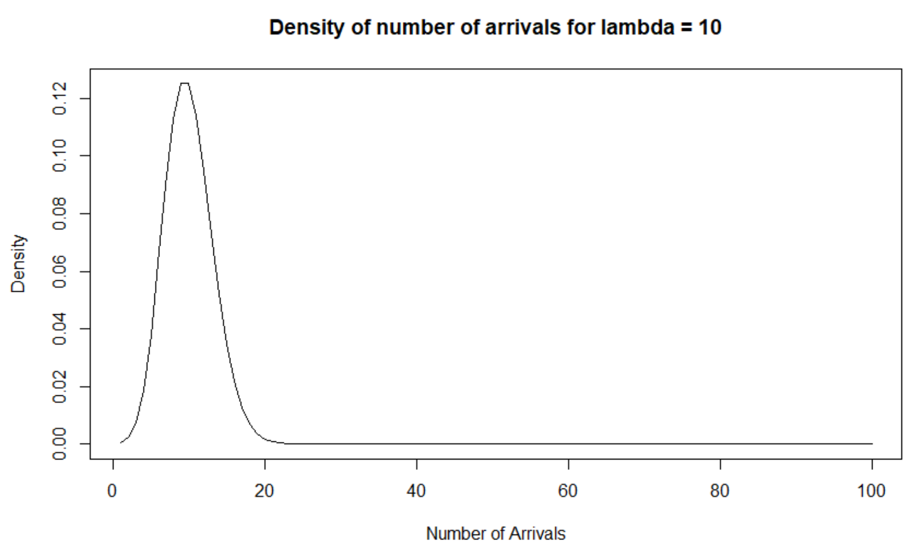
- From the cumulative distribution plot we can observe that CDF of first inter-arrival time in case of p = 0.9 is faster approaching to 1 than p = 0.4 and from this observation we can claim that CDF of 2^{nd},3^{rd} ….. K^{th} inter-arrival time we have always p = 0.9 faster approaching 1 than p = 0.4 because of the fresh start property.

Ques2. **Observation on number of visitors in a website**

We have given that a website receives an average of 10 visitors per hour.So we have to study the number of visitors in a time interval (0,t] where t is continuous and also given that this process is modeled as the Poisson process.

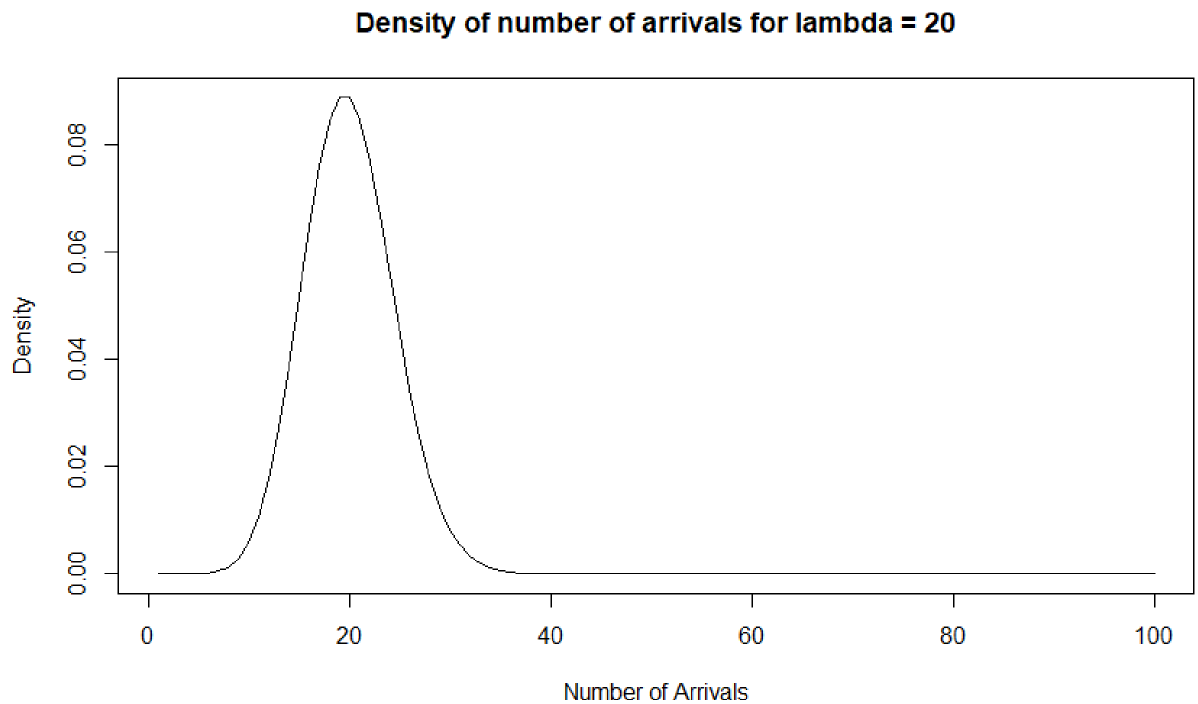
We will denote the rate of arrivals as lambda.
Mean of Poisson Distribution = lambda = Variance of Poisson Distribution,

(a) We have given a rate of arrival as 10 visitors per hour, which means lambda = 10,
So after simulating density of number of arrivals, we get a plot :-



We observe that we have a maximum density at lambda number of arrivals which is 10 this is because our mean is also equal to lambda in case of Poisson Distribution and also our variance also equals to lambda and we can see that from the graph that we have a min value of 0 and max value of 20 i.e., 10 + 10 as max and 10-10 as min value.

(b) Now we have a lambda = 20,
So after simulating the density of number of arrivals,we get a plot :-



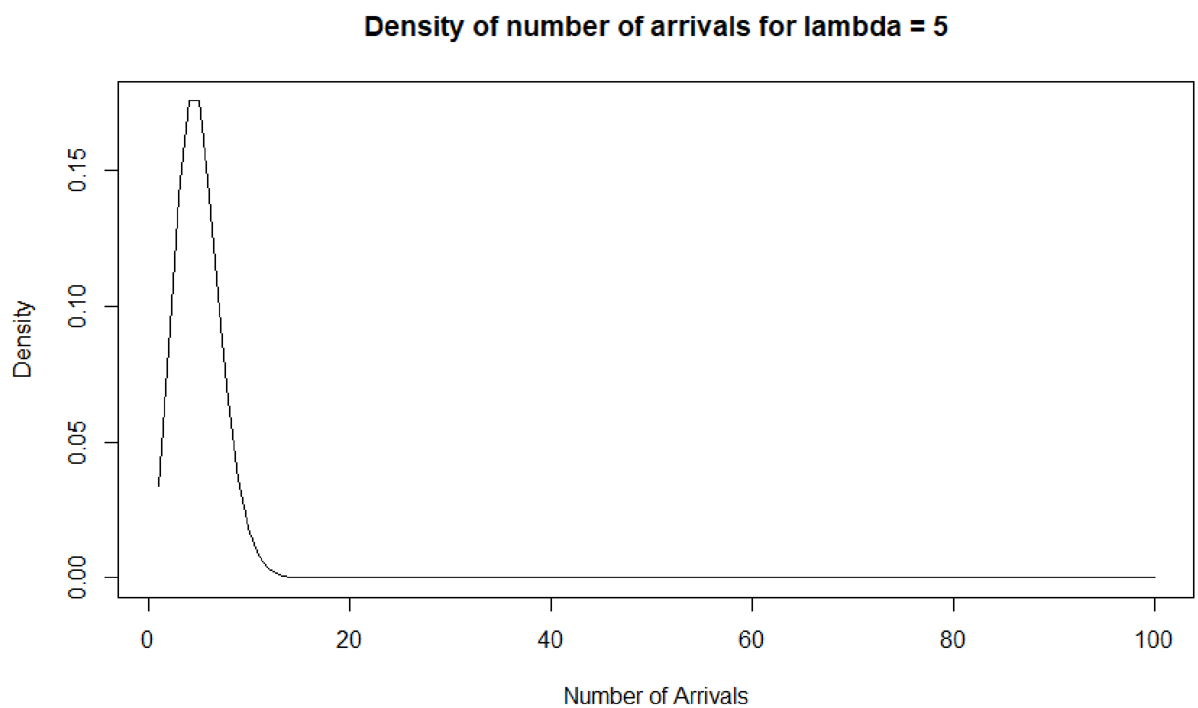
So, we have a maximum density at number of arrivals = 20 because here lambda we have 20 which is equal to mean and variance also we have 20 which shows that we have max val of 40 arrivals and min value of 0 arrivals which we can confirm from the plot also.

Comparison between results for lambda = 10 and lambda = 20 :-

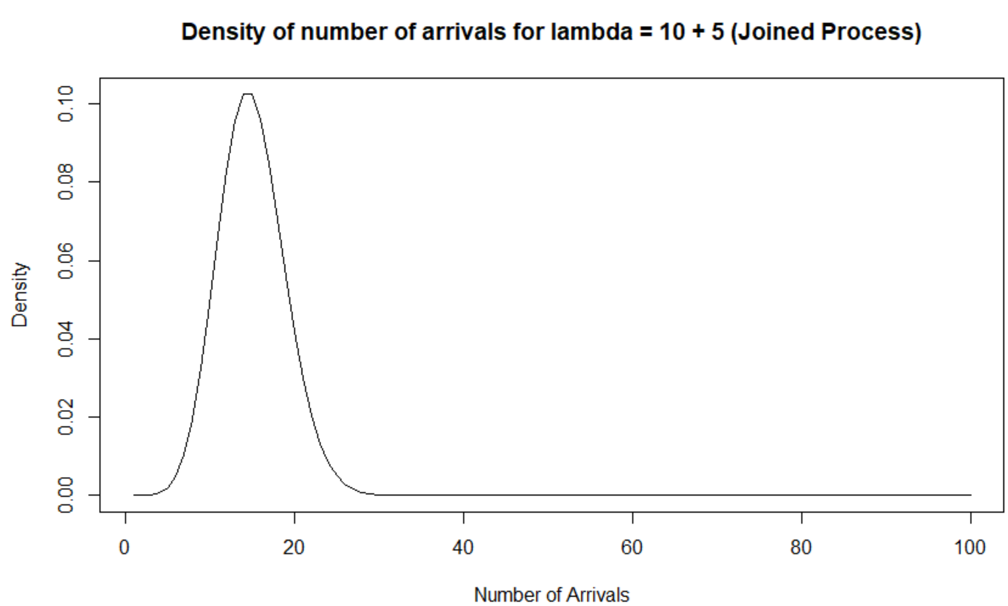
- Maximum density in lambda = 10 we get is around 0.13 and in case of lambda = 20 max density is around 0.09, So there is decrease in density value at lambda = 20 and this is because of higher variance in case of lambda = 20 than lambda = 10.
- At lambda = 10 the maximum density we get at number of arrivals = 10 and at lambda = 20 the maximum density we get at number of arrivals = 20, this is because of change in mean value which is equal to our lambda in poisson distribution.

By comparing the results of both the plots we have observed that if we increases the lambda the highest density point will move rightwards i.e., number of arrivals at which density is maximum and decreases highest density value or we can say that bandwidth increases peak value decreases as we increase the lambda.

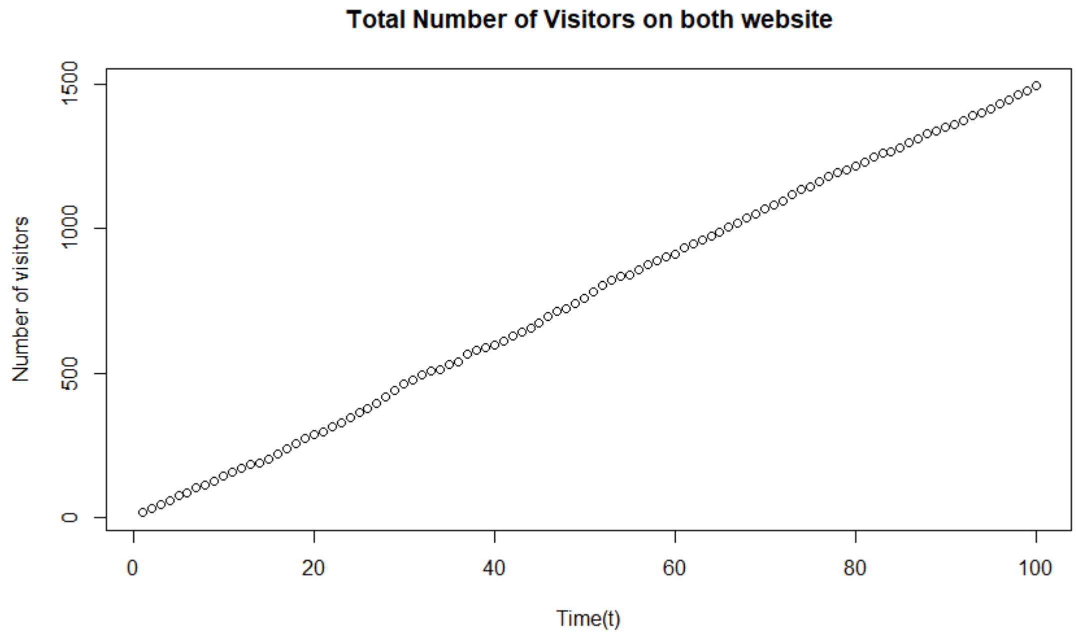
(c) Now, we have a new independent website with lambda = 5, we know that two independent poisson process can be combine as one and the rate of that process will be sum of both the rates,
So firstly we will plot the density graph for lambda = 5,



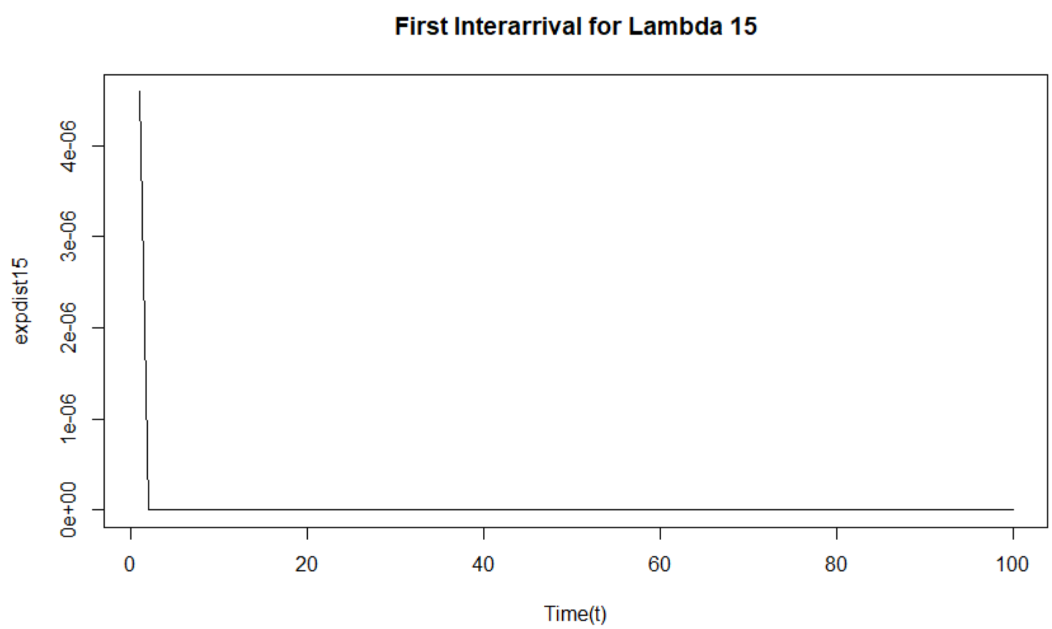
Now we will plot density graph for combines process which means lambda = 15,



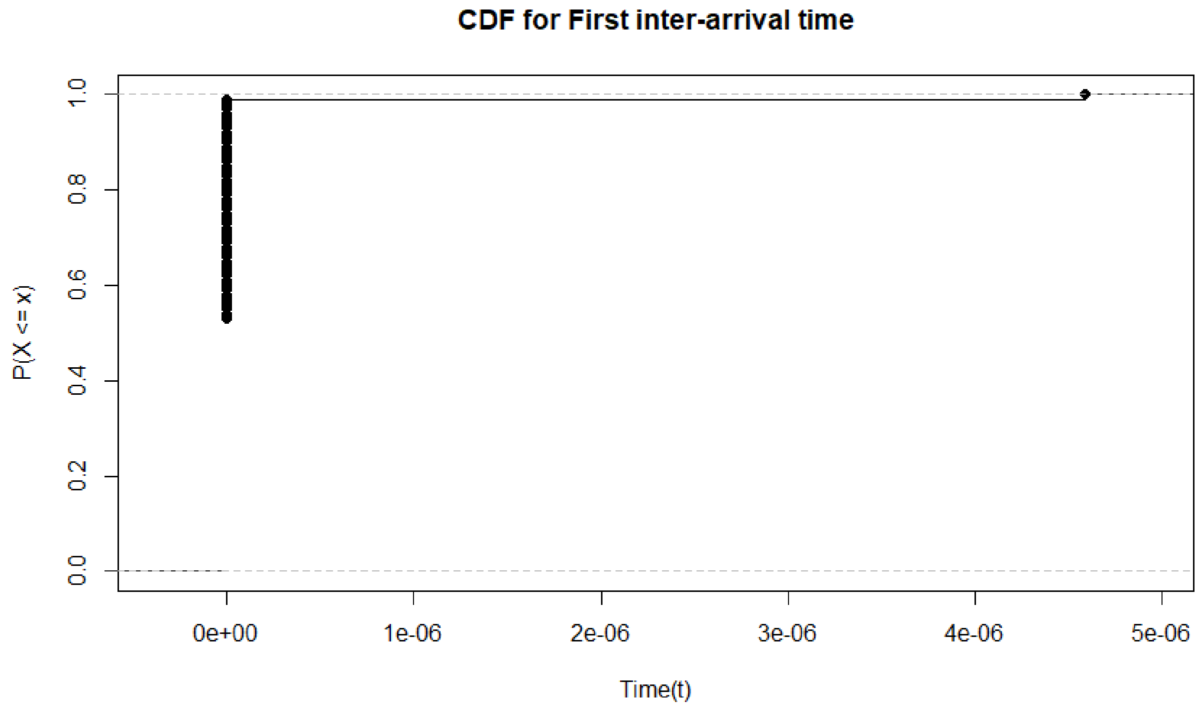
The plotting of the above 2 graphs can be described same as we describe for lambda = 10 and 20 in previous part,
Now we will plot total number of visitors from both the website and we get,



We get around 1500 visitors and this is because we have arrival rate = 15 and total hrs we observer is 100 so 15*100 = 1500 will be our total visitors.
Simulation for first inter-arrival time, We know that inter-arrival time for Poisson Process follows exponential distribution. So we can plot the graph by using formula P(N = 1) = (lambda * t) * e^-(lambda*t)



Now, Corresponding CDF we have,



We can see that CDF value gets 1 in very less time and this is because of the lambda, if the lambda is higher than the first arrival happens in very less time and vice versa.