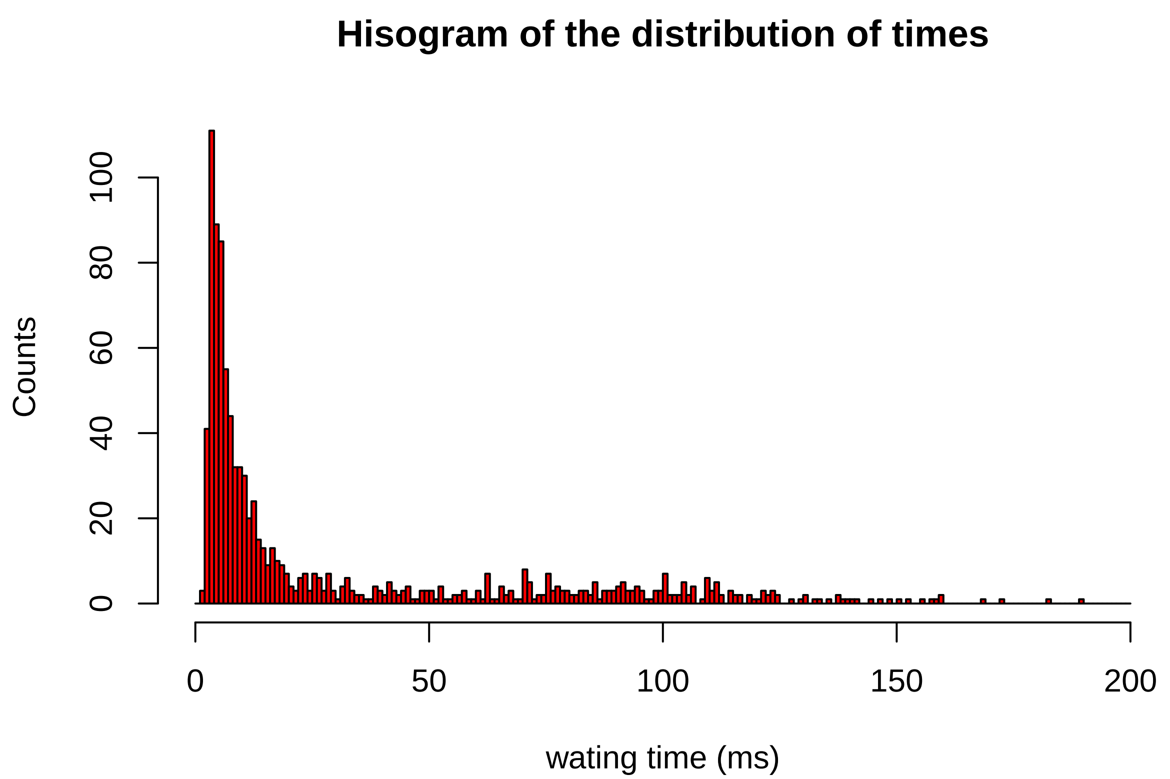
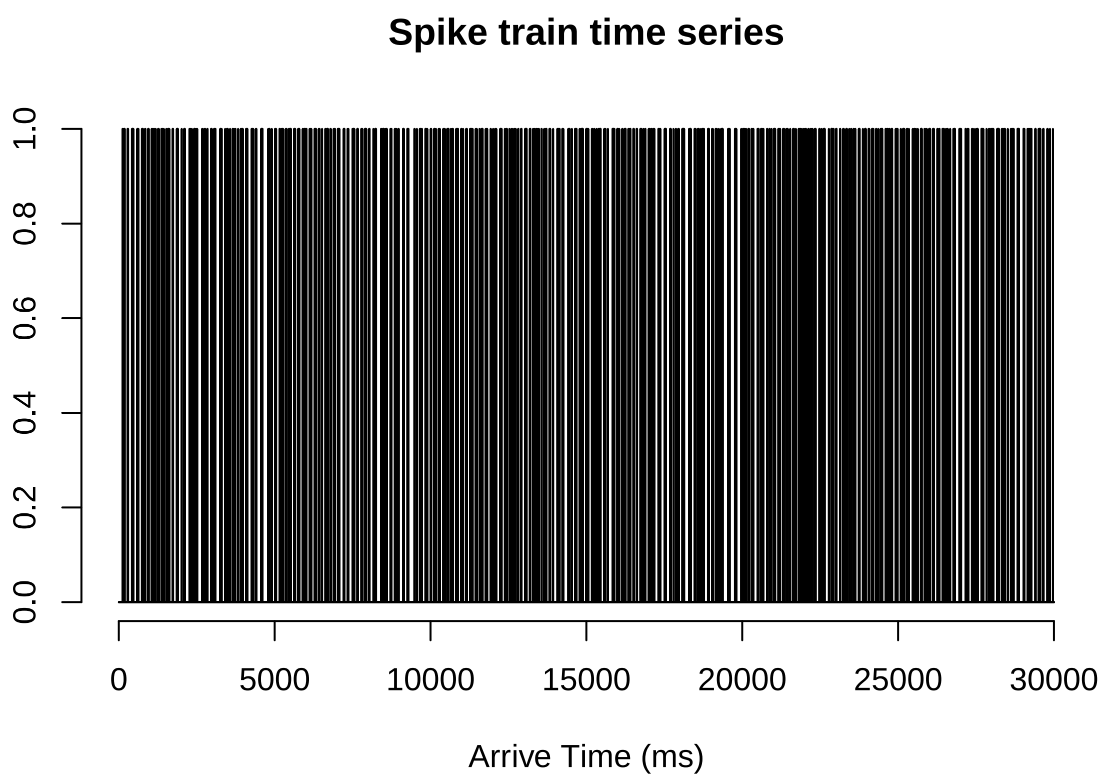
Prob1_MA568

Hengchang Hu

9/29/18

1.Download data file Retinal_ISIs.txt which contains wating times in milliseconds & Plot spiking activity as a histogram of the distribution of the times and as a spike train time series & Describe the spiking properties .

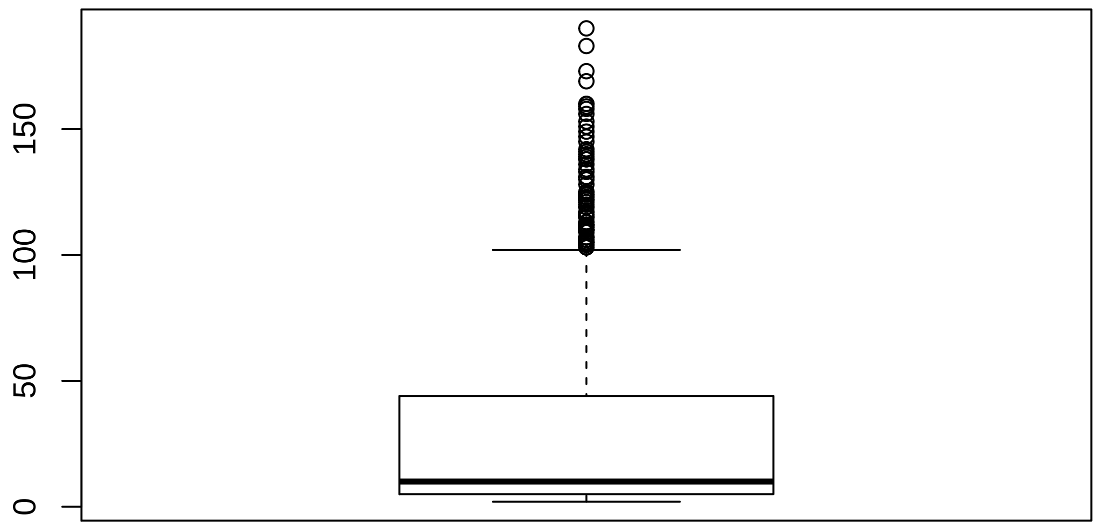




Conclusion: The histogram shows that most of the waiting times are 3-30 ms, and the waiting times between 3-11 ms consists of more than 50 % of observed data. So a homogeneous Poisson Process may not well fitted in this data.

2.Compute a 5-number summary & box plot for ISI distribution

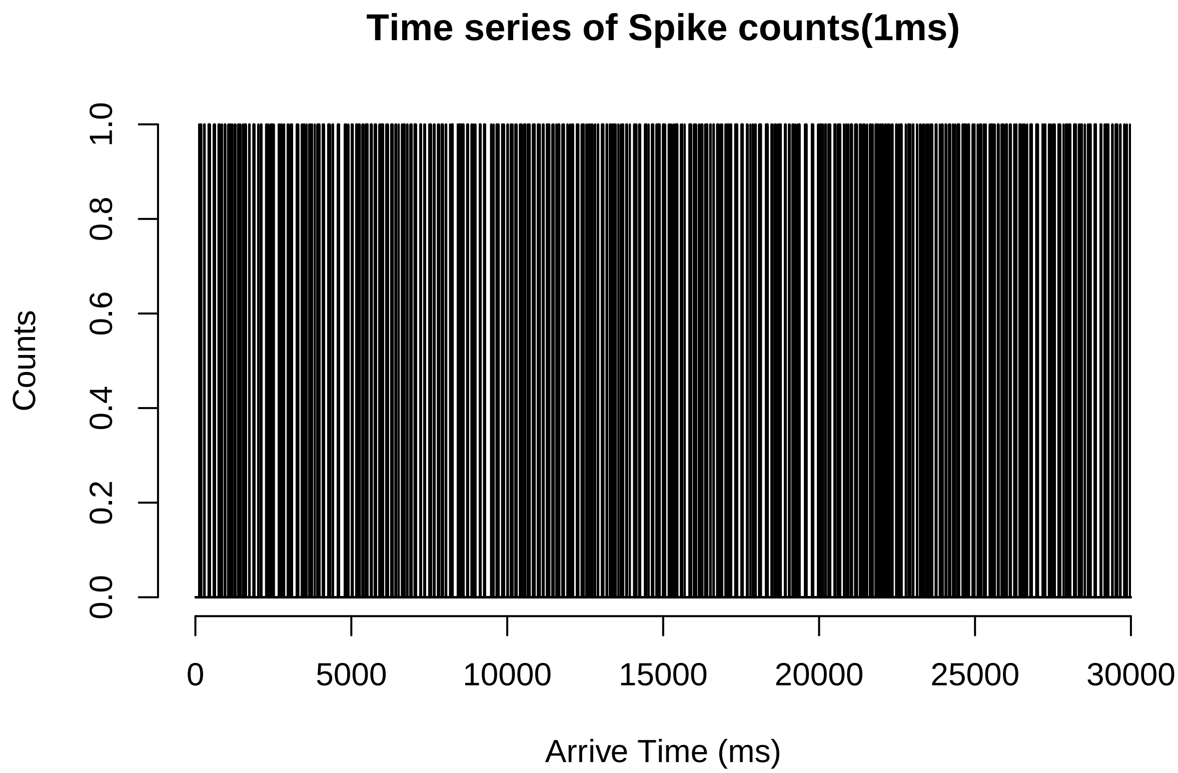
```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    2.00    5.00   10.00   30.83   44.00  190.00
```

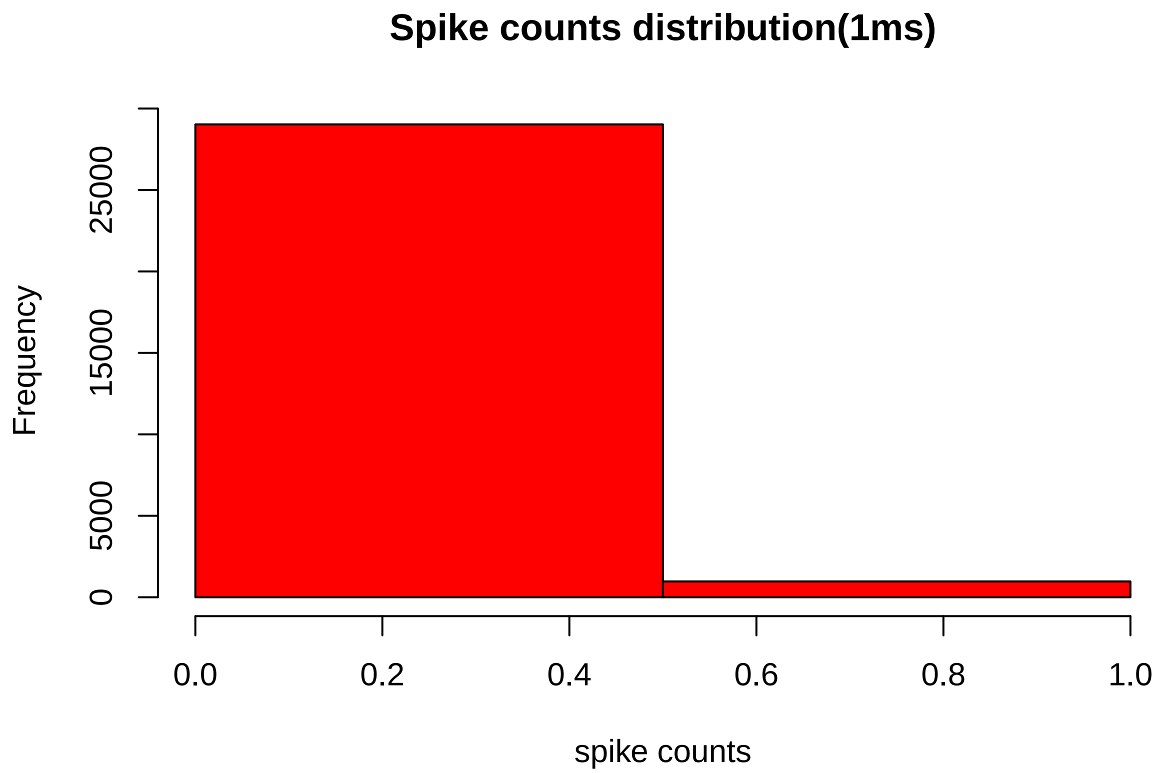


Conclusion: 5-number summary of this data is (2, 5, 10, 44, 190) corresponding to (min, .25 quantile, median, .75 quantile, max).

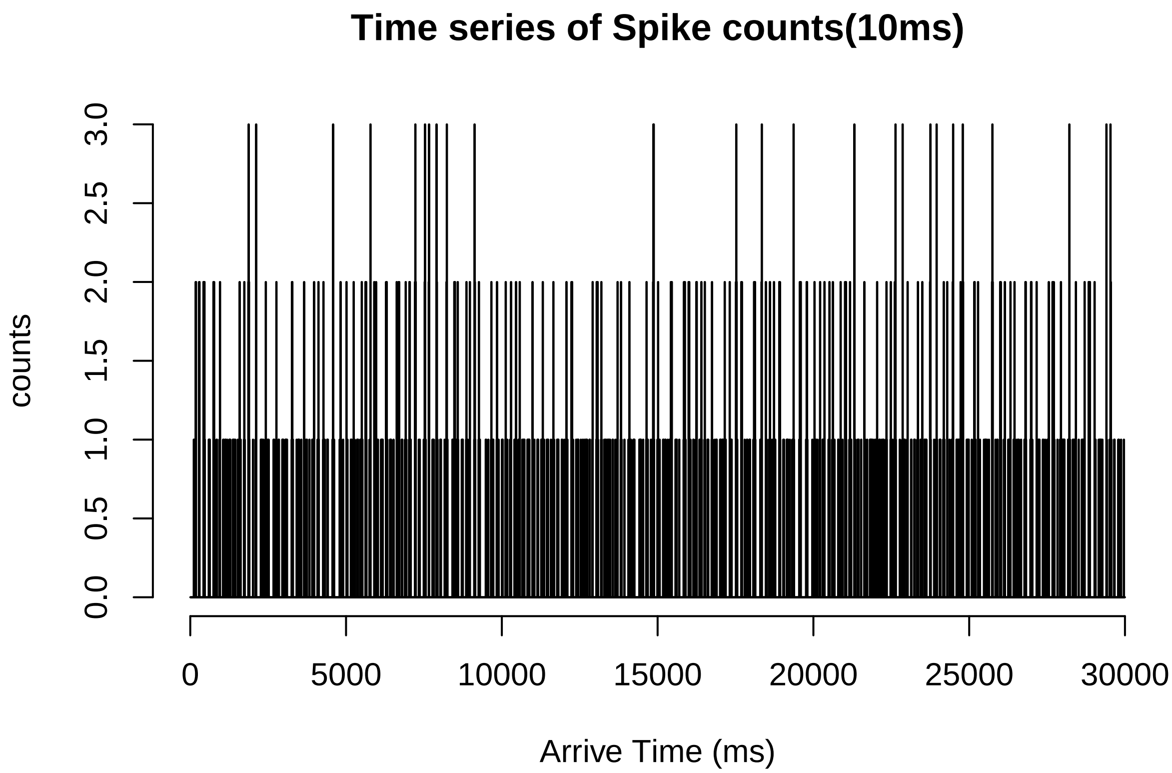
- From the 5-number summary, we can see that most of waiting time data are concentrated in 5-44 ms.
- From the boxplot, we can tell that this is a long-tailed distribution because there are lots of big data over 100 ms although most of data are concentrated in 5-44 ms.

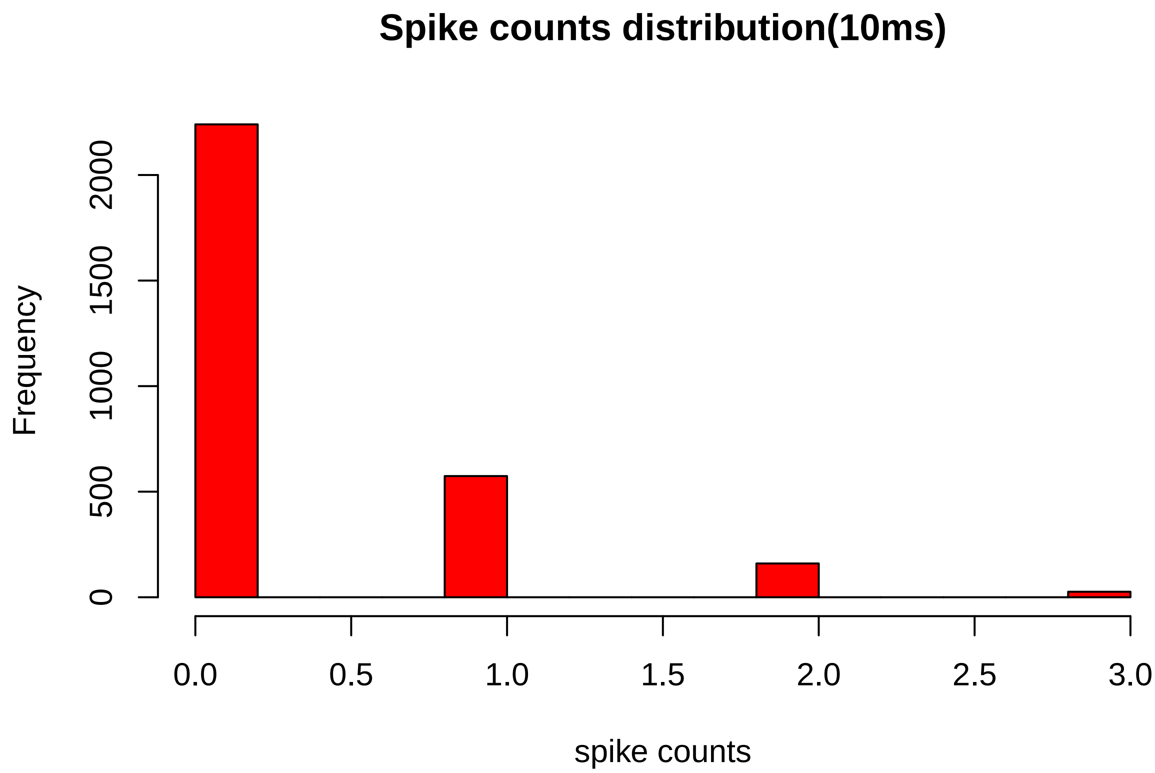
3.Bin the spike train data from Retinal_ISIs.txt into time bins of width 1ms., 10ms., 100ms & Plot time series of spike counts and distribution of spike counts as histogram for each bin width

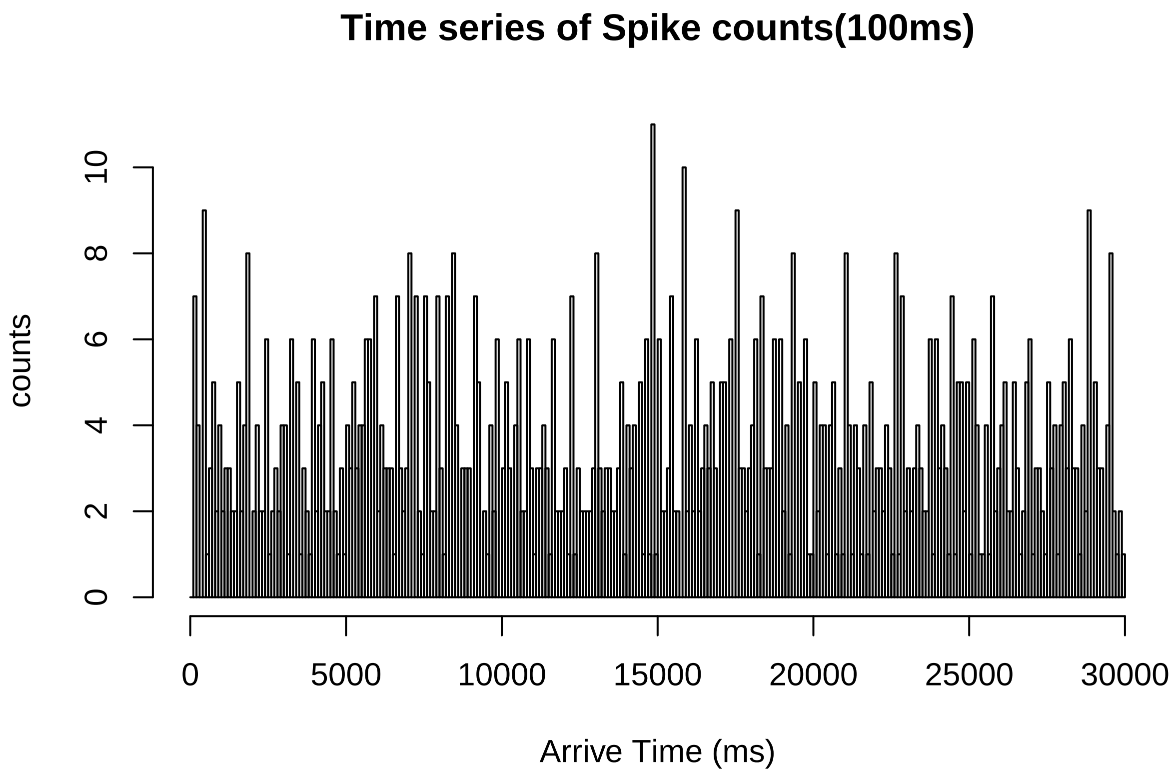

Time series of Spike counts(1ms)
Counts
0.0
0.2
0.4
0.6
0.8
1.0
0
5000
10000
15000
20000
25000
30000
Arrive Time (ms)

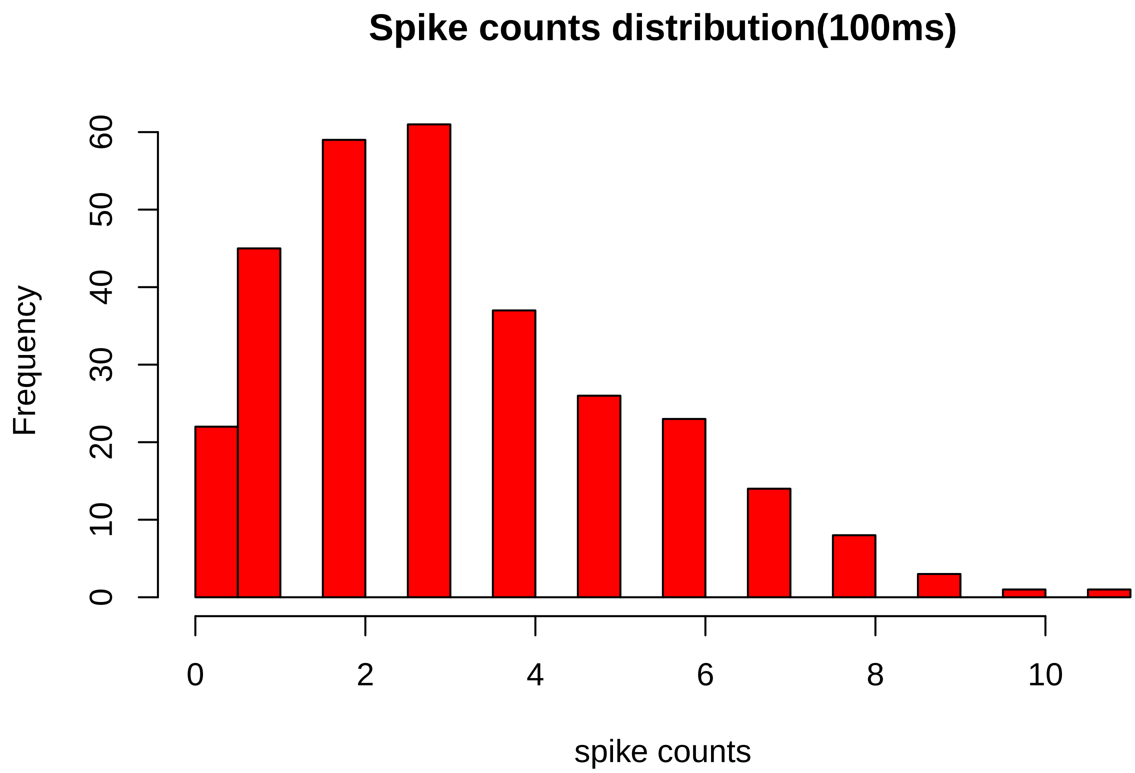


Spike counts distribution(1ms)
Frequency
0
5000
15000
25000
0.0
0.2
0.4
0.6
0.8
1.0
spike counts


Time series of Spike counts(10ms)
counts
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
5000
10000
15000
20000
25000
30000
Arrive Time (ms)

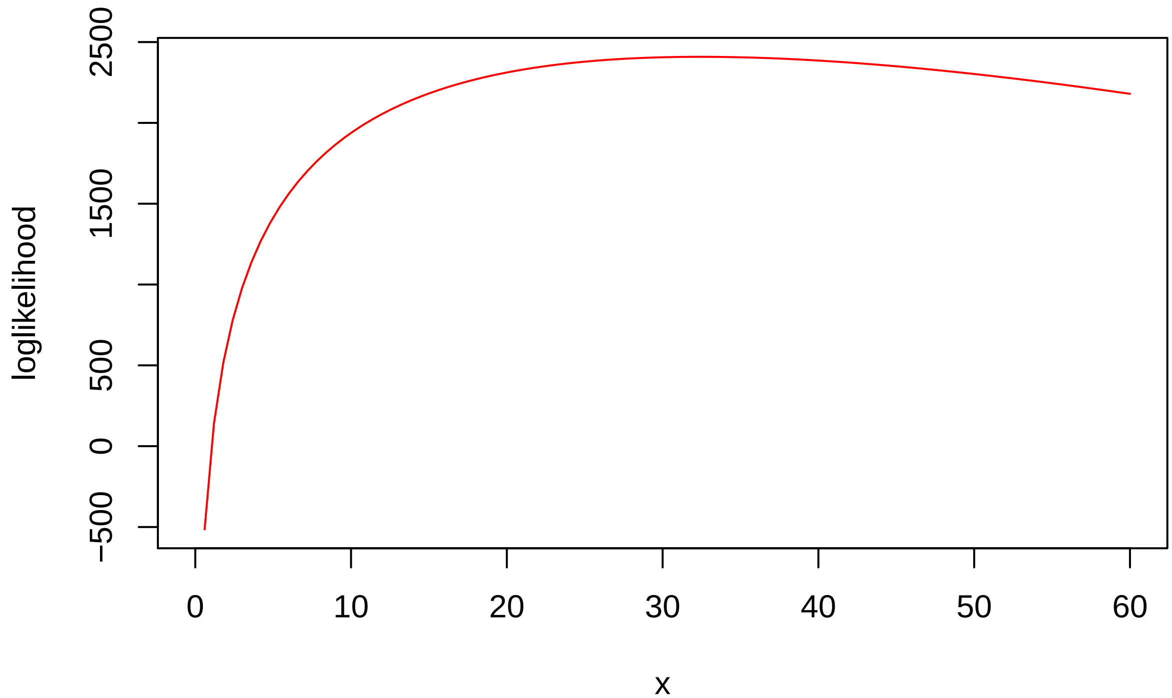

Spike counts distribution(10ms)
Frequency
0
500
1000
1500
2000
0.0
0.5
1.0
1.5
2.0
2.5
3.0
spike counts


Time series of Spike counts(100ms)
counts
0
2
4
6
8
10
0
5000
10000
15000
20000
25000
30000
Arrive Time (ms)



Conclusion: Comparing to model for increments of a Poisson process, the distribution of spike counts are long-tailed especially in 100 ms width time bins.

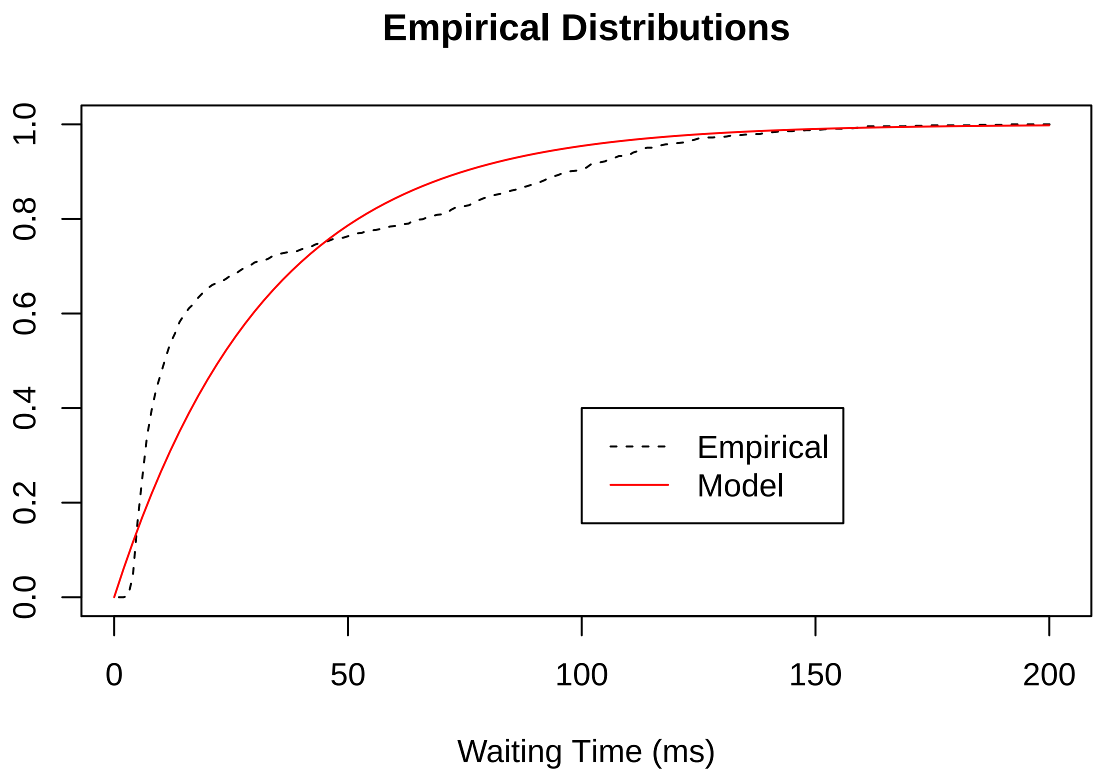
4.Plot the likelihood as a function of λ for values of λ between 0 Hz to 60 Hz & Find value $\hat{\lambda}_{ML}$ that maximize the likelihood & Provide an approximate 95% confidence interval for $\hat{\lambda}_{ML}$.



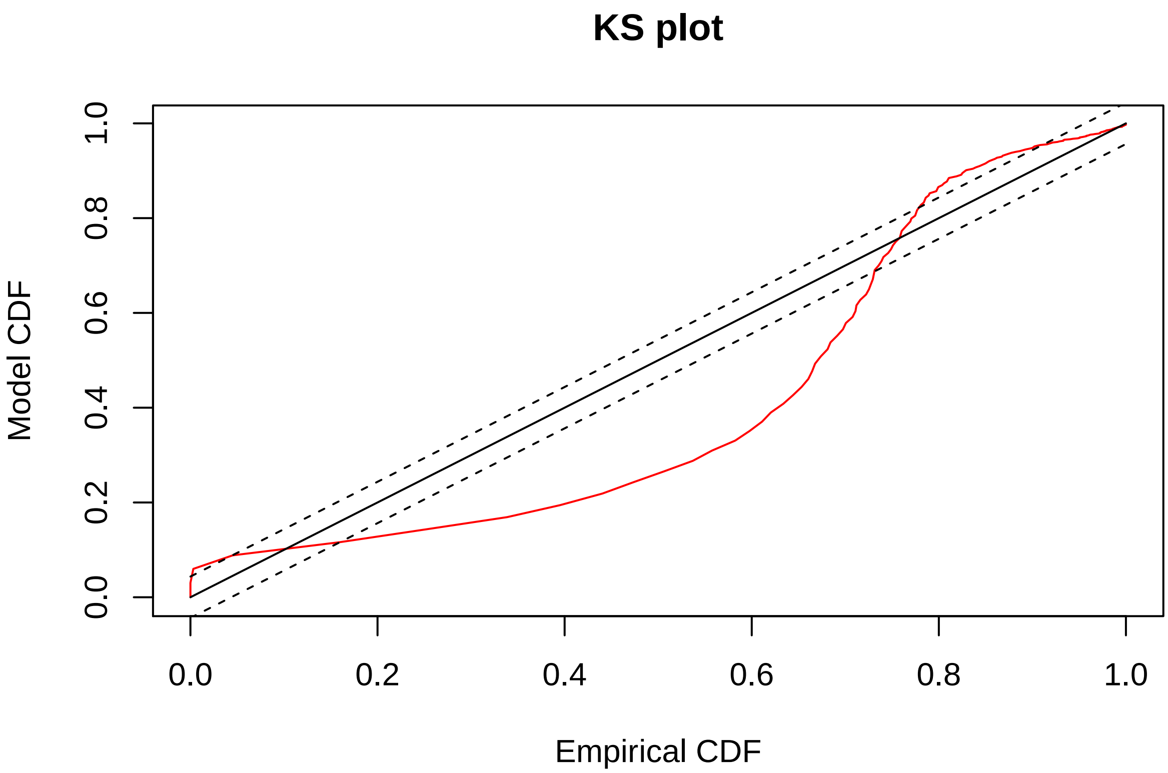

```
## $maximum
## [1] 32.4
##
## $objective
## [1] 2408.77
```

Conclusion: The value of $\hat{\lambda}_{ML}$ that maximaize the likelihood is 32.4, and the approximate 95% confidence interval for $\hat{\lambda}_{ML}$ is $[32.4 - 1.96 \times \frac{\sqrt{972}}{30}, 32.4 + 1.96 \times \frac{\sqrt{972}}{30}]$, i.e. $[30.36311, 34.43689]$.

5.Plot an empirical CDF of the interspike intervals for the data & Plot the exponential CDF on the same plot as your empirical CDF & Construct a KS plot of the empirical CDF on the x-axis against the model CDF on the y-axis

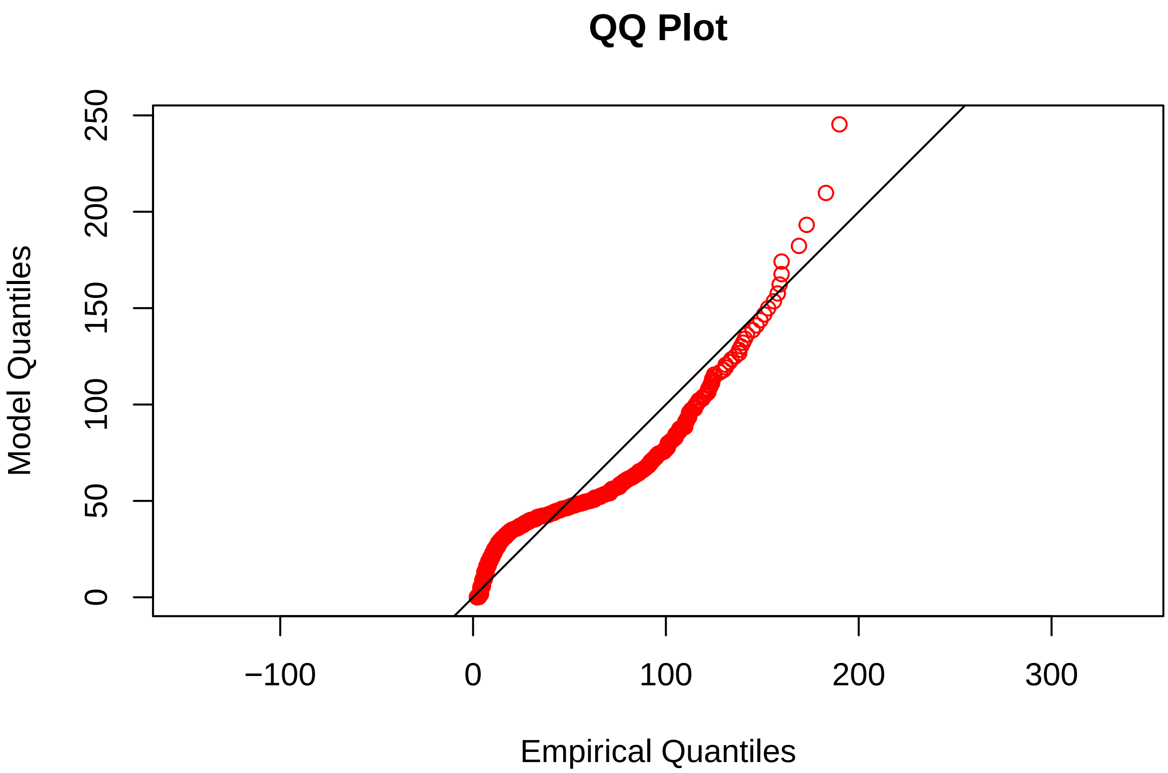


Empirical Distributions
1.0
0.8
0.6
0.4
0.2
0.0
0
50
100
150
200
Waiting Time (ms)
Empirical
Model




```
## [1] "KS statistics is 0.251797527798863"
```

6.Construct a QQ plot of the empirical vs model quantiles



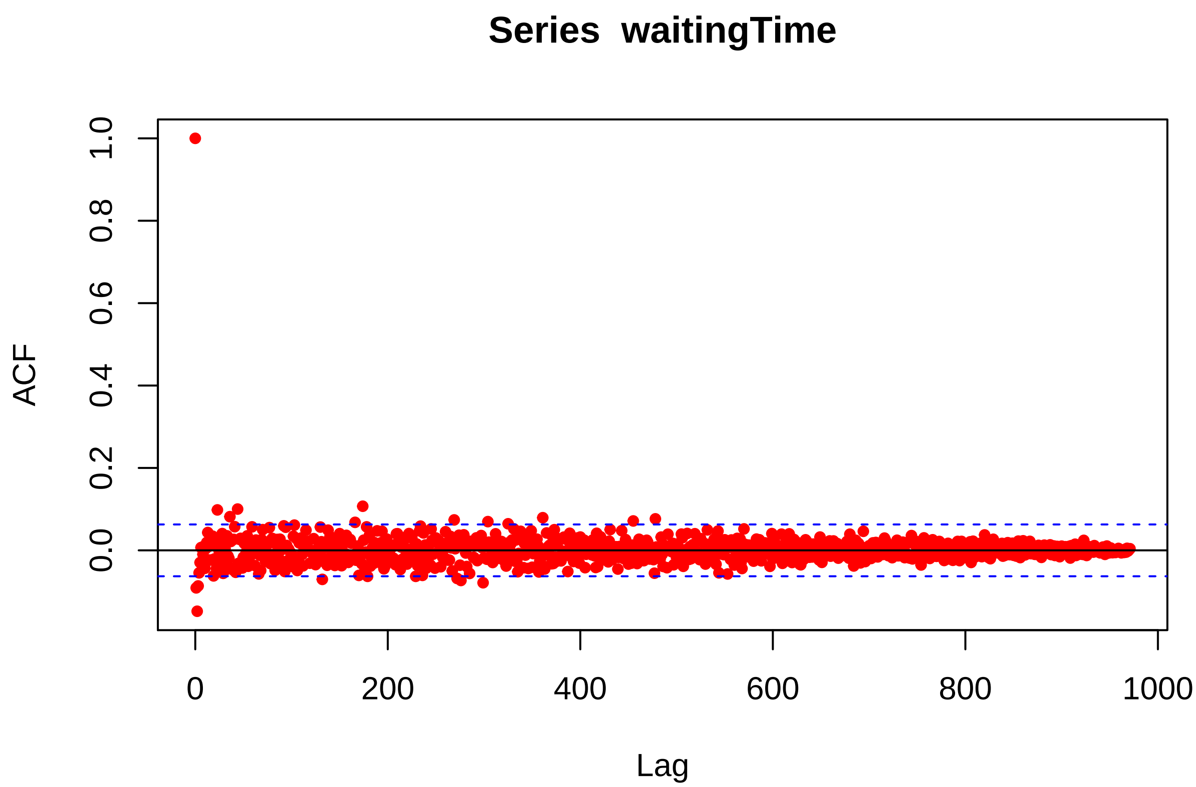
Conclusion: From the Q-Q Plot, we can see that QQ plot deviates most from expected for ISI values between 50-150 ms.

7.Compute the Fano Factor for increments process binned at 1ms, 10ms, 100ms*

```
## [1] "The sample Fano Factor for the increments process binned at 1ms is  0.96763225440848"
## [1] "The 95% confidence interval of sample Fano Factor for a Poisson Process at 1ms bin length is [ (
## [1] "The sample Fano Factor for the increments process binned at 10ms is  1.16610063436783"
## [1] "The 95% confidence interval of sample Fano Factor for a Poisson Process at 10ms bin length is [
## [1] "The sample Fano Factor for the increments process binned at 100ms is  1.45827655972584"
## [1] "The 95% confidence interval of sample Fano Factor for a Poisson Process at 100ms bin length is
```

Conclusion: These statistics do not fall into the 95% confidence interval for a Poisson process at each of these bin lengths.

8.Plot the autocorrelation function of the observed interspike intervals with 95% confidence bounds



Conclusion: From the plot of autocorrelation function, we can see that most of data can be assumed as independent data because there are only a few data falls out the confidence bounds.

9.What conclusions can you draw about the spontaneous firing properties of these retinal ganglion neurons & what other variables could be added to improve goodness-of-fit

Conclusion: A homogeneous Poisson process fails on many goodness-of-fit tests, so a homogeneous Poisson Process may not be a good choice. The Fano Factor value of 1 ms increments shows that it is more similar to Poisson process, and Fano Factor values of 10 ms and 100 ms increments are significantly greater than 1, which means that comparing to Poisson Process they are less stable. And autocorrelation function shows that data are almost independent, however some of data do not fall into the confidence bounds.