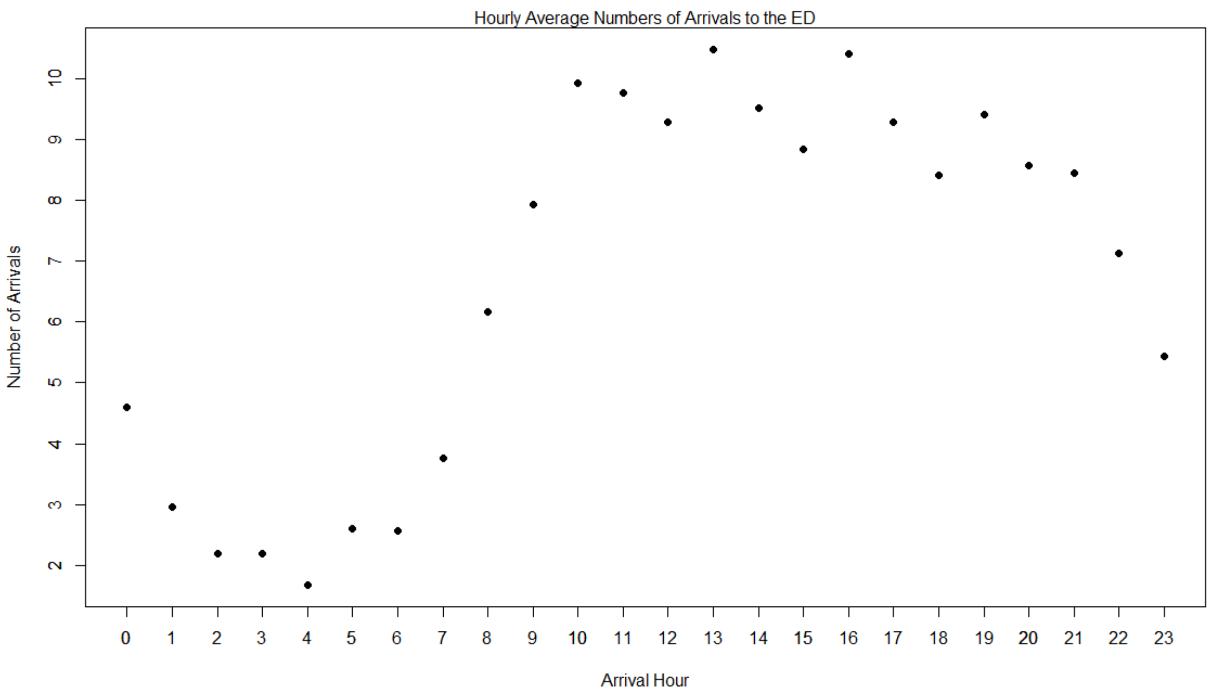
# Case Study

Henry Cui

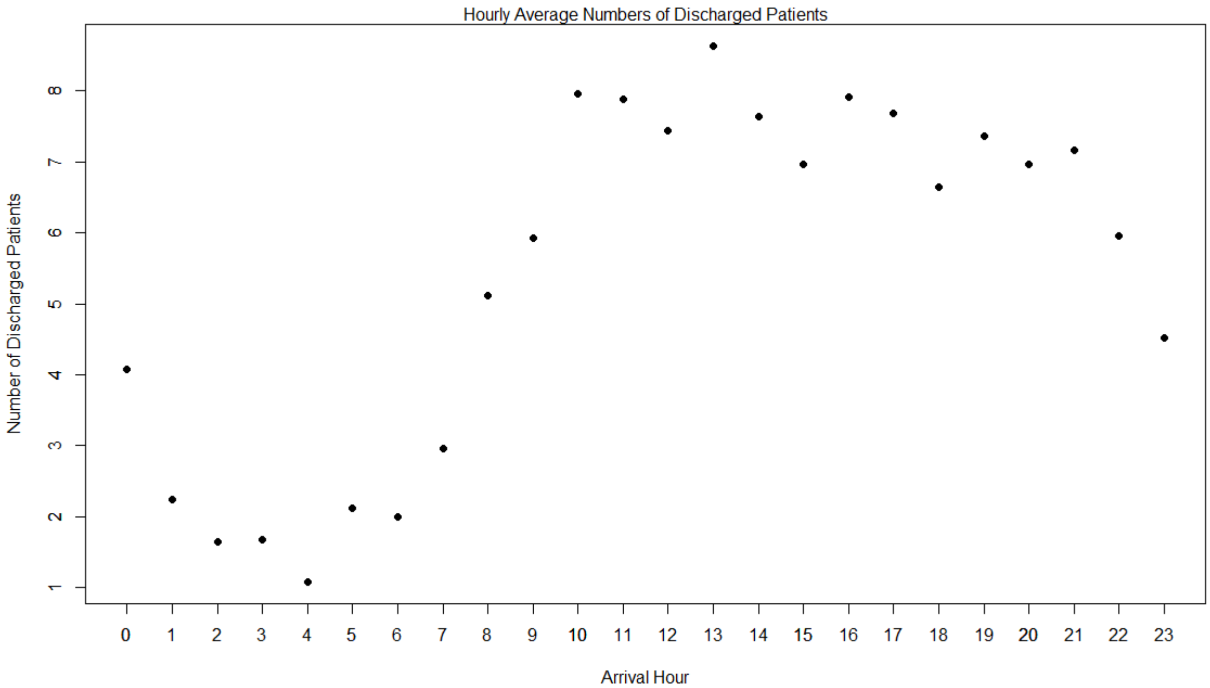
2022/4/17

## Part I. Arrival Process

### (a)



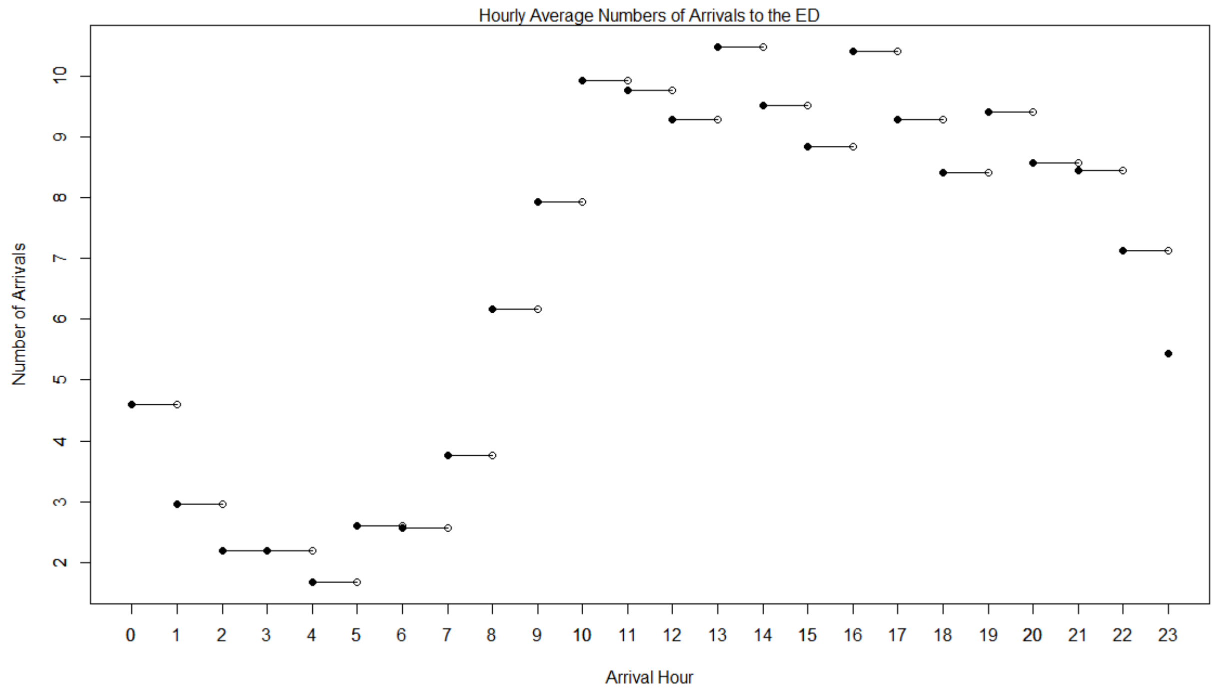
### (b)




### (c)

The approximate arrival rate function for each hour is one constant. We can refer to the following plot for each hourly arrival rate as constants.

The probability is about 0.0042.



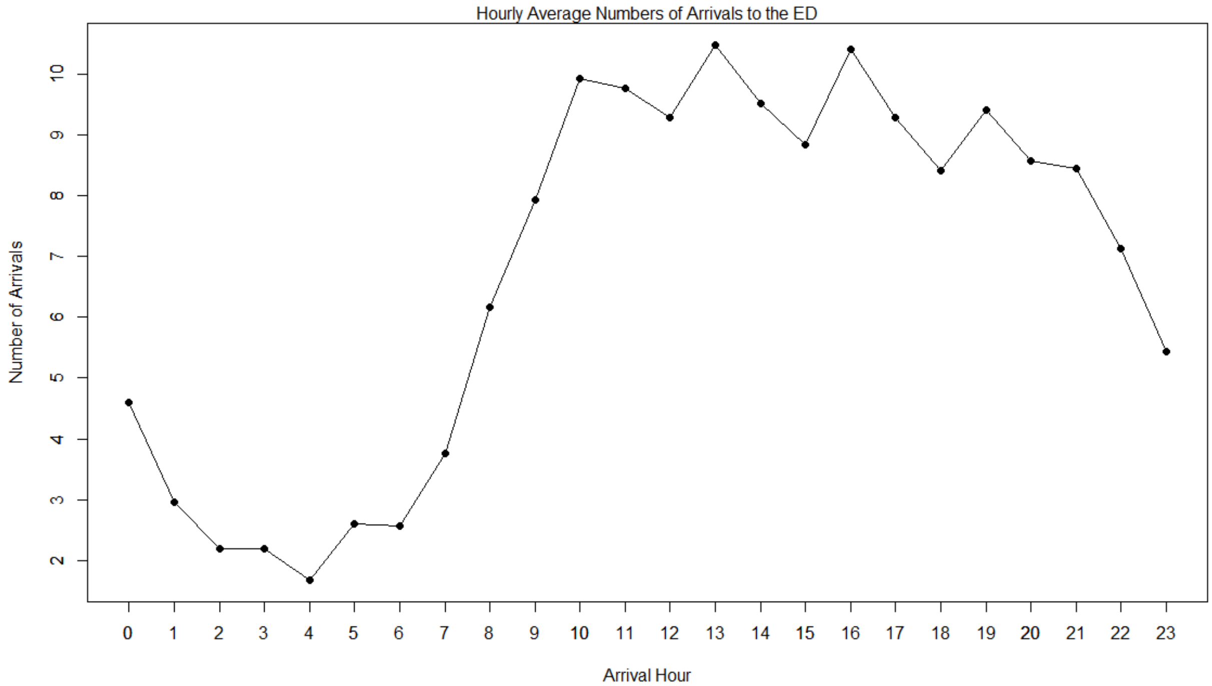

```r
1 - cul
```

```
## [1] 0.004163416
```

## (d)

The approximate arrival rate function for each hour is one linear function in the interval [i,i+1], where i can be 0 to 22. We can refer to the following plot for the piece wise linear arrival rate function.

The probability is about 0.0037.




```
1-cul2
```

```
## [1] 0.003658564
```

## (e)

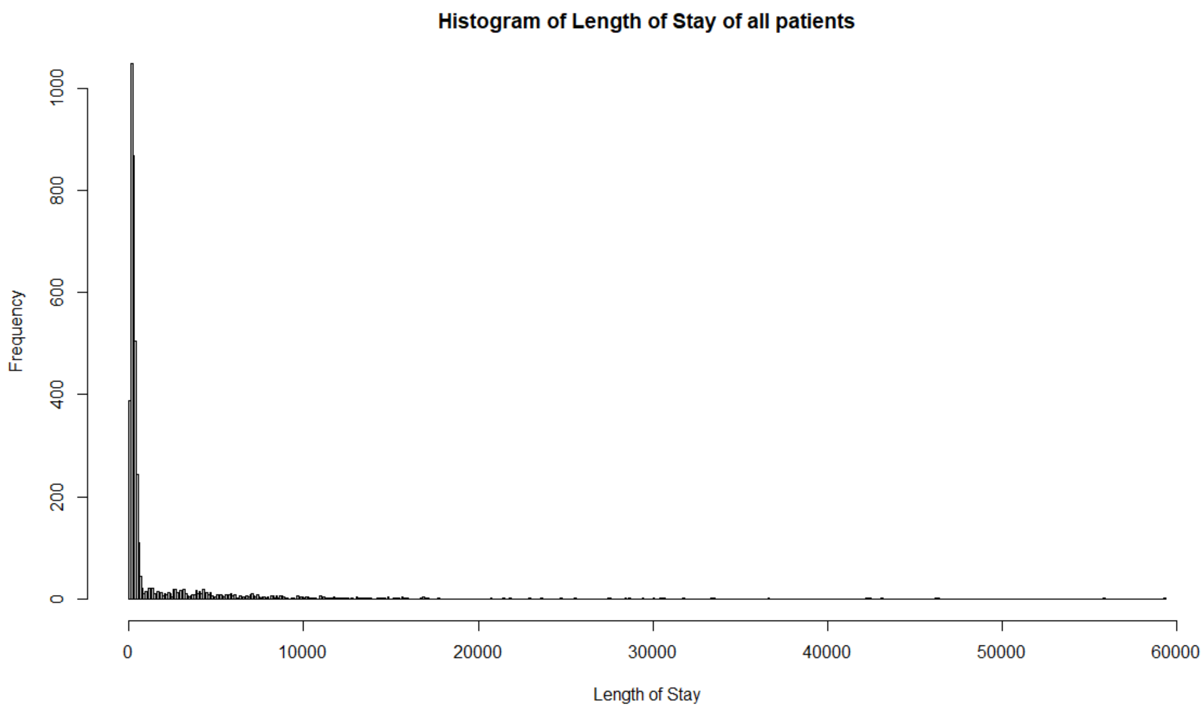
IDC for each hour is printed below.

```
##  [1] 0.8876812 0.6610360 2.6195652 0.5752688 0.9597701 0.3750000 0.4915459
##  [8] 0.8181004 0.8758170 0.6316926 1.0484190 0.8412534 1.3586957 1.1734234
## [15] 0.8360768 0.8428030 1.4431525 1.0883405 0.6686047 1.8240741 0.9544025
## [22] 0.9126984 0.8072344 1.9196429
```

```
## [1] "mean of IDC is 1.02559576614775 and standard deviation of IDC is 0.505348831988967"
```

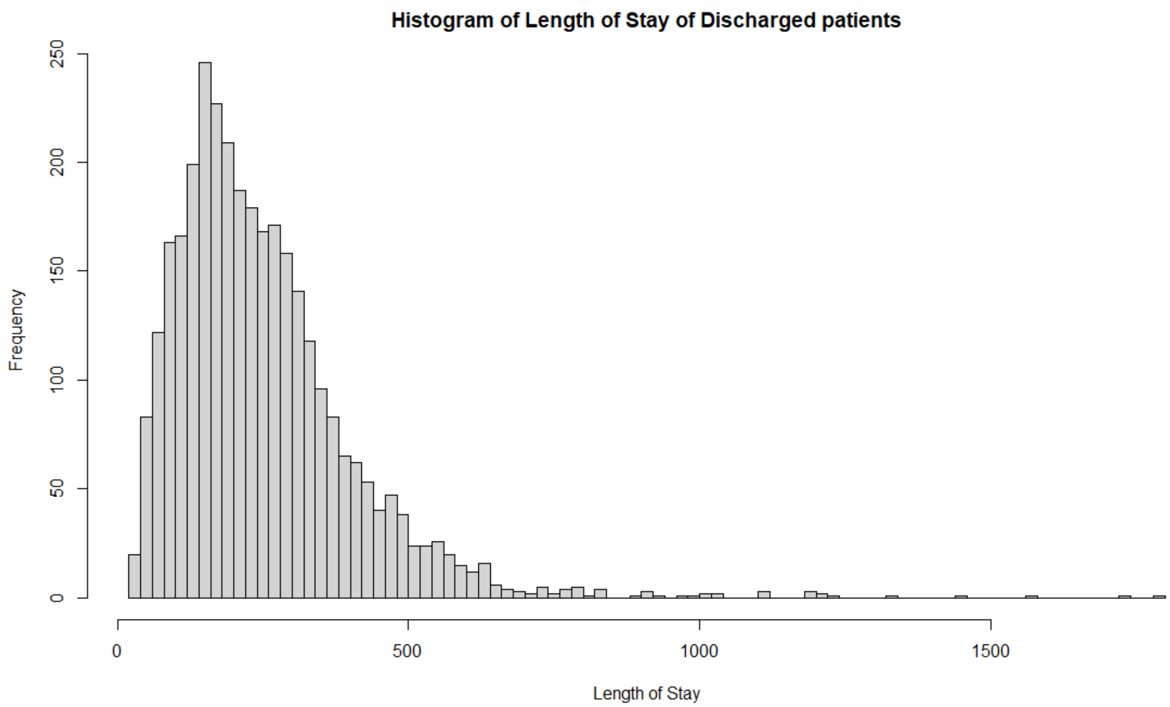
# Part II. Length of Stay

## (a)

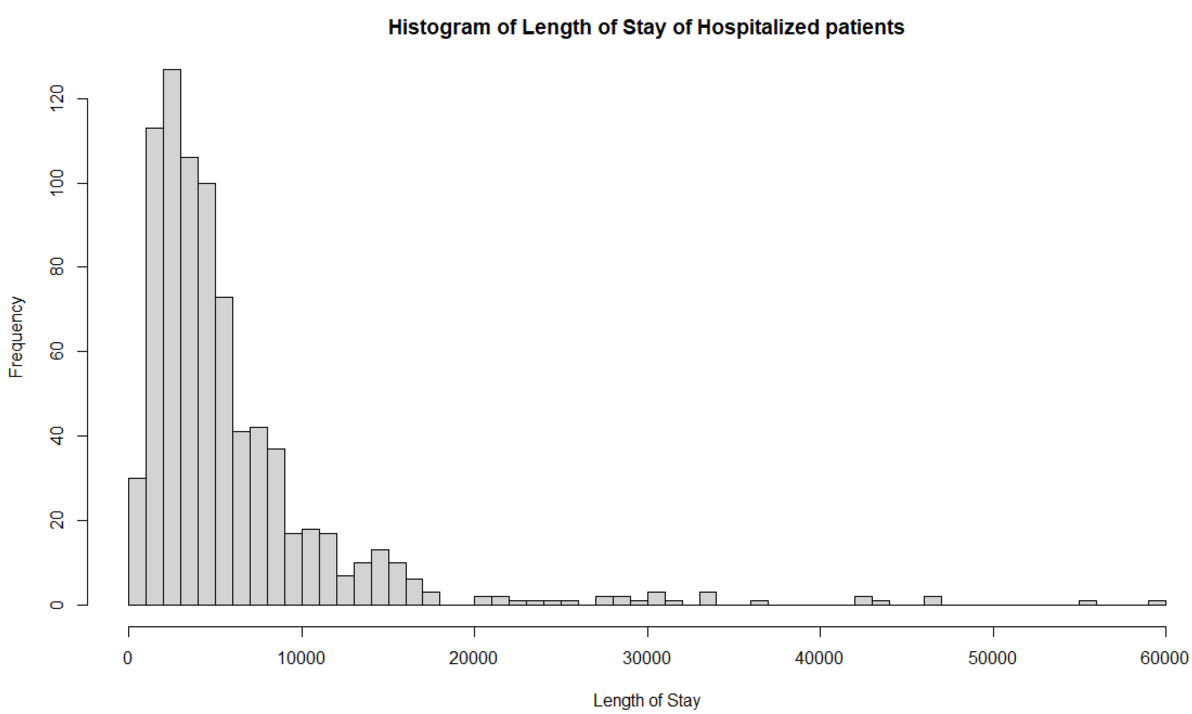



```
## [1] "mean is 1390.86945751796 and standard deviation is 3700.22089074266"
```

(b)




```
## [1] "for discharged patients, mean is 251.560049397962 and standard deviation is 159.331614353288"
```




```
## [1] "for hospitalized patients, mean is 6015.20927318296 and standard deviation is 6522.62578755142"
```

### (c)

The first row printed below is the mean of the length of stay of the patients in different age groups. The second row printed below is the standard deviation of the length of stay of the patients in different age groups.

```
##  [1]  293.5402  586.5309  697.1473  897.8148 1262.8670 1918.7511 2664.1547
##  [8] 3009.1807 3081.8018 3203.4699
```

```
##  [1]  605.497 2019.376 2444.174 2533.225 3472.601 4379.028 5784.636 5374.051
##  [9] 4446.876 3845.592
```

### (d)

Parameter value is 0.00072.

```
## [1] "parameter value is:  0.000718974735259 86"
```

### (e)

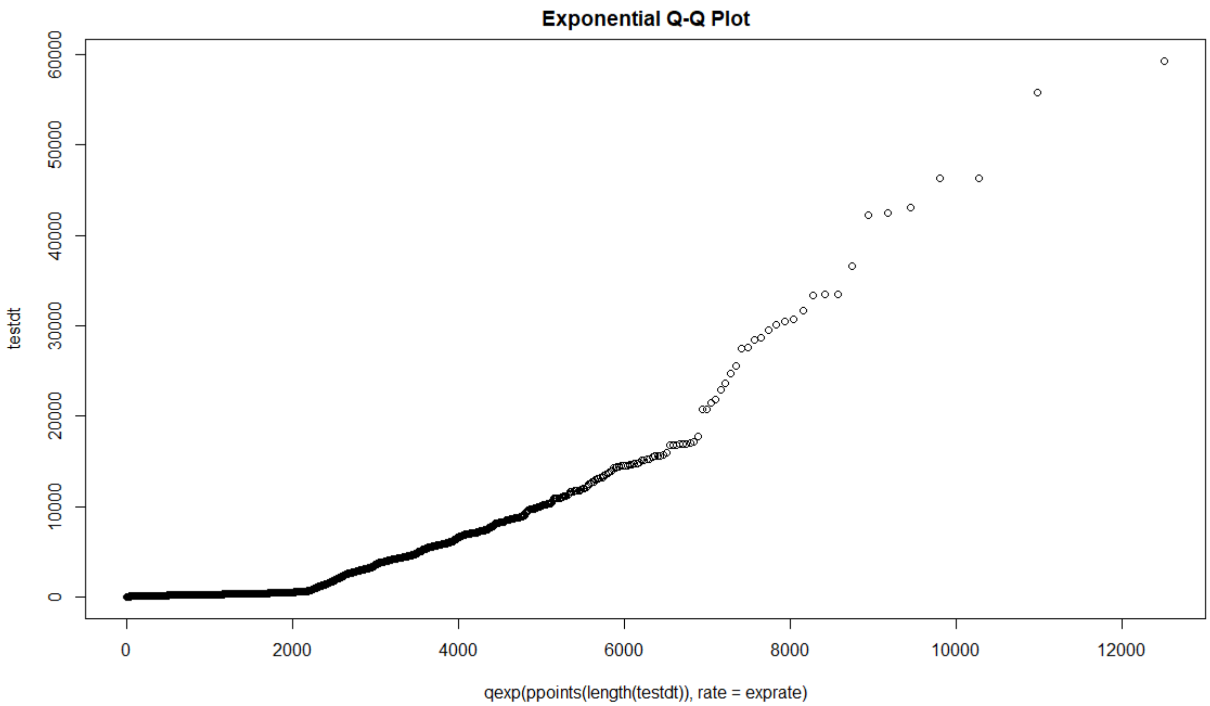
By Q-Q plot, the distribution should be Gamma. By Cullen and Frey Graph, the distribution should be beta or Gamma distribution.

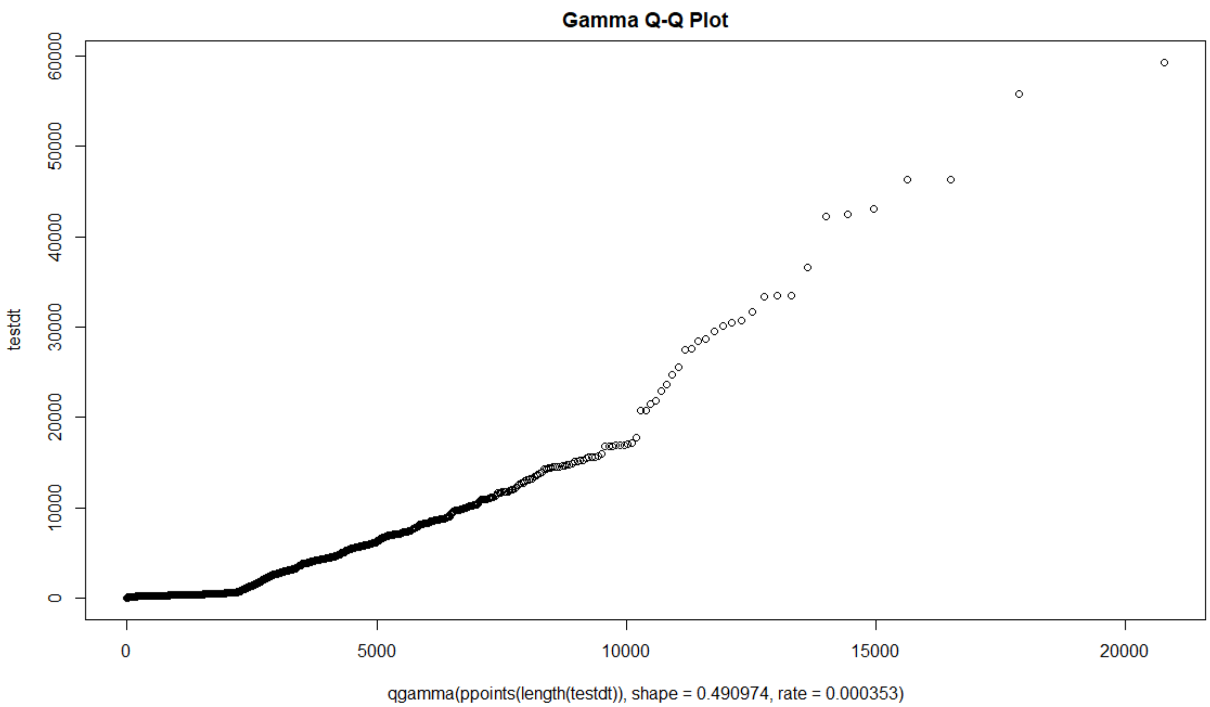
```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  testdt
## D^+ = 0.45482, p-value < 2.2e-16
## alternative hypothesis: the CDF of x lies above the null hypothesis
```

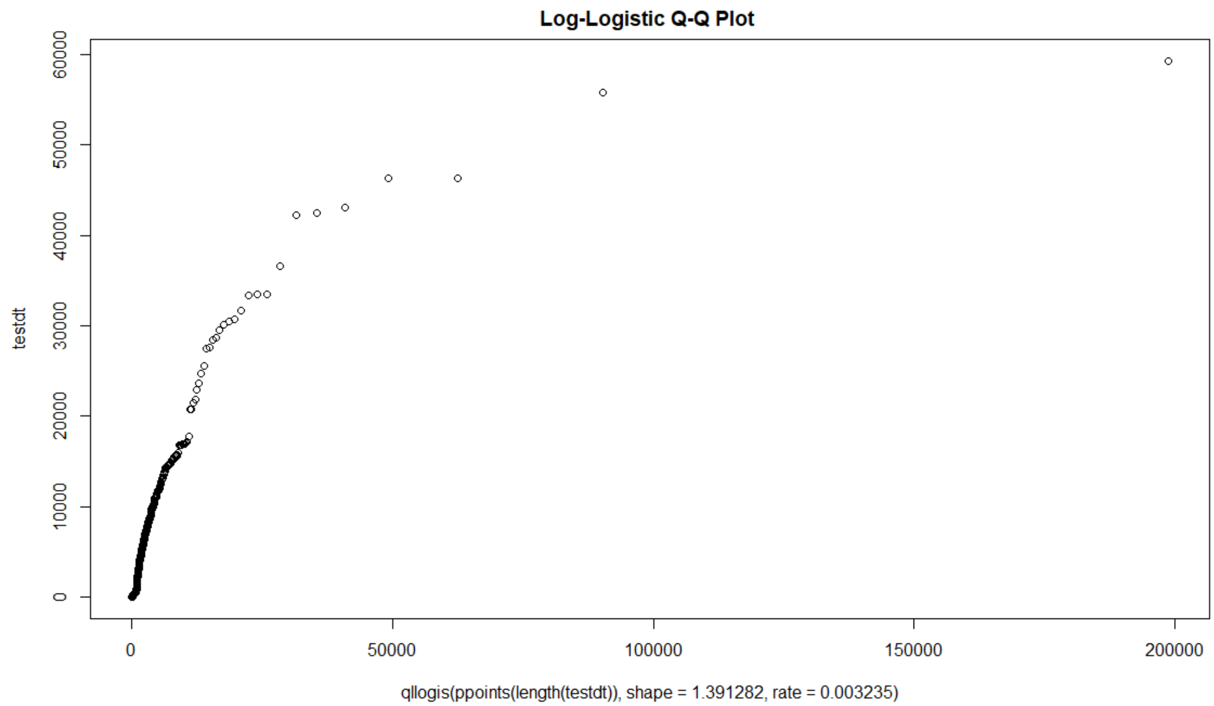
```
## Maximum likelihood estimates for the Gamma model
##    shape      rate
## 0.490974  0.000353
```

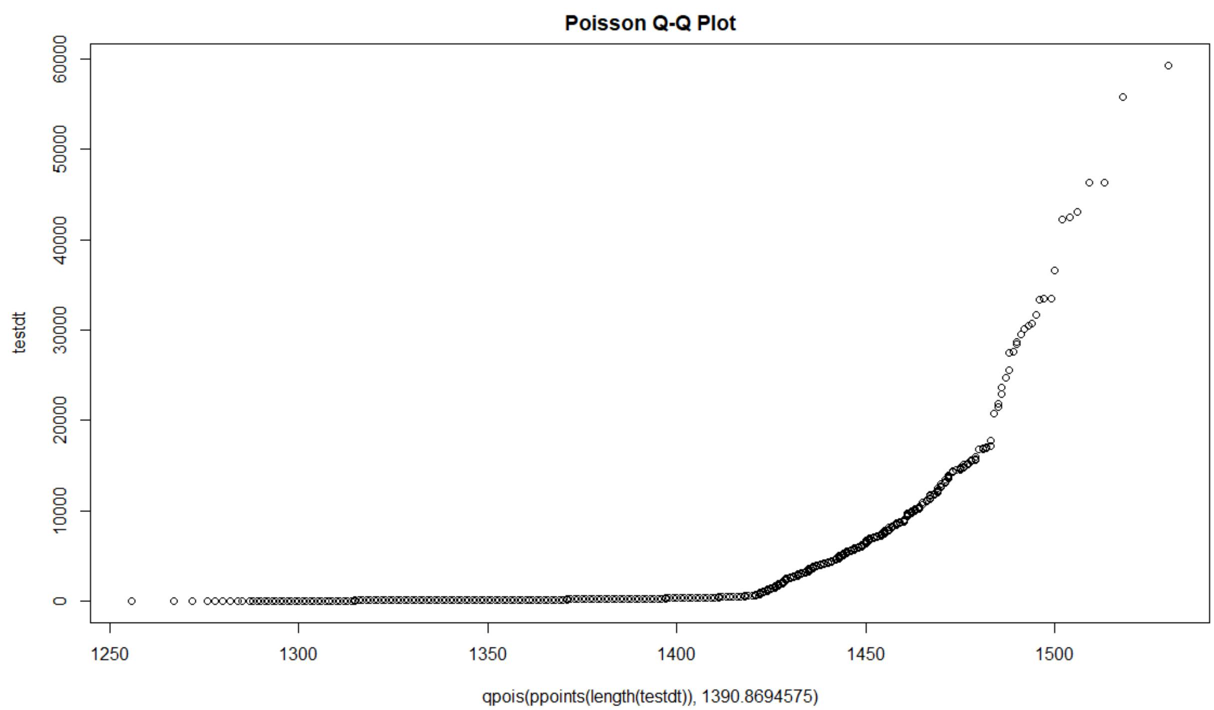
```
## Maximum likelihood estimates for the Loglogistic model
##    shape      rate
## 1.391282  0.003235
```

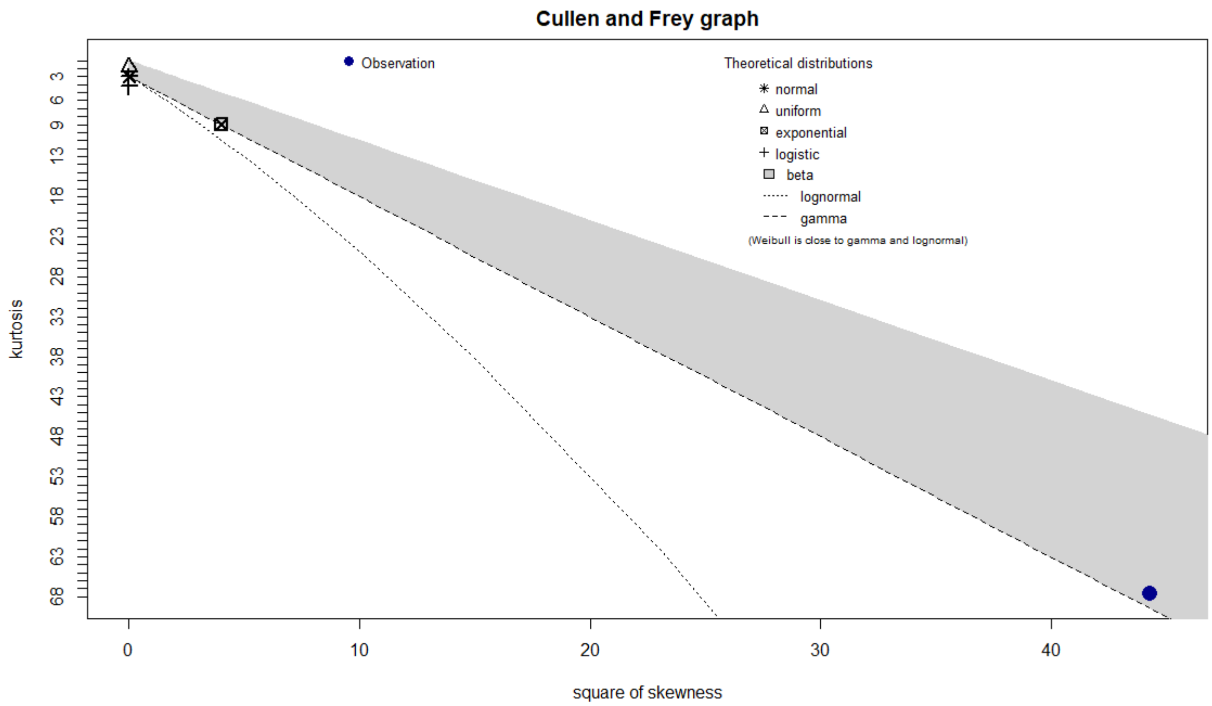
```
##       lambda
##   1390.8694575
##  (   0.5869672)
```


Exponential Q-Q Plot
testdt
0
10000
20000
30000
40000
50000
60000
0
2000
4000
6000
8000
10000
12000
qexp(ppoints(length(testdt)), rate = exprate)


Gamma Q-Q Plot
testdt
0
10000
20000
30000
40000
50000
60000
0
5000
10000
15000
20000
qgamma(ppoints(length(testdt)), shape = 0.490974, rate = 0.000353)


Log-Logistic Q-Q Plot
testdt
0
10000
20000
30000
40000
50000
60000
0
50000
100000
150000
200000
qllogis(ppoints(length(testdt)), shape = 1.391282, rate = 0.003235)

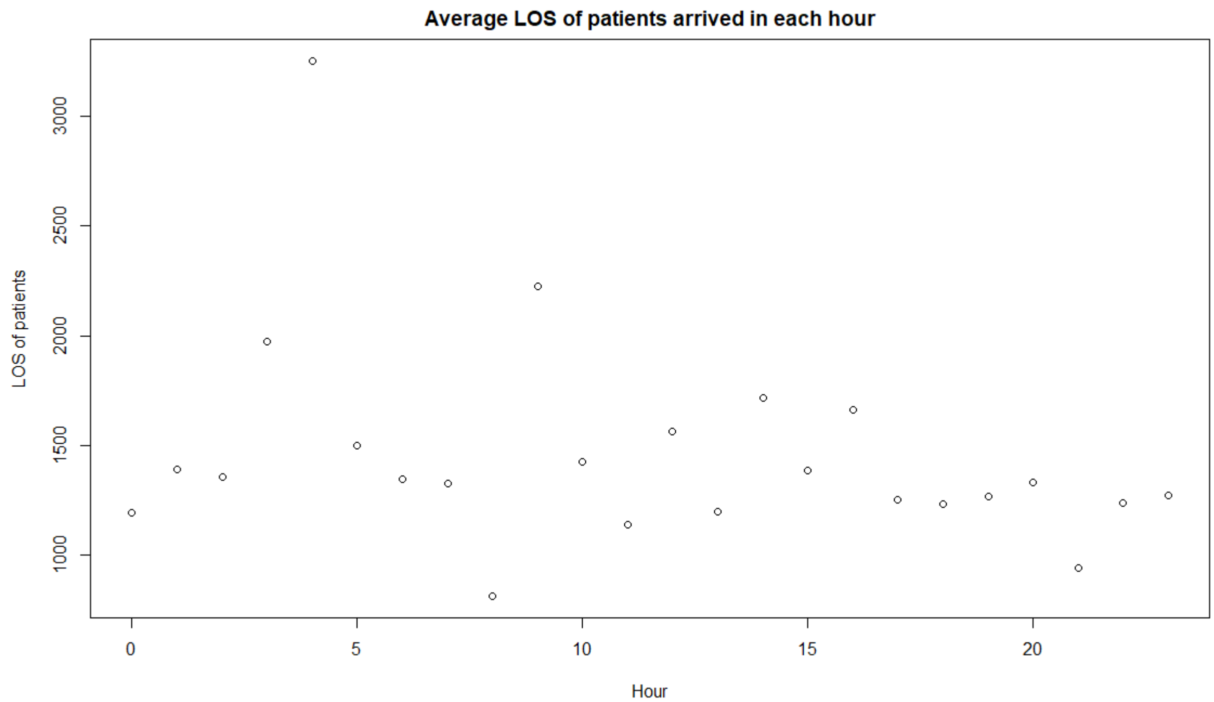

Poisson Q-Q Plot
testdt
0
10000
20000
30000
40000
50000
60000
1250
1300
1350
1400
1450
1500
qpois(ppoints(length(testdt)), 1390.8694575)




```
## summary statistics
## ------
## min:  20   max:  59342
## median:  266
## mean:  1390.869
## estimated sd:  3700.221
## estimated skewness:  6.654149
## estimated kurtosis:  67.67478
```

## (f)

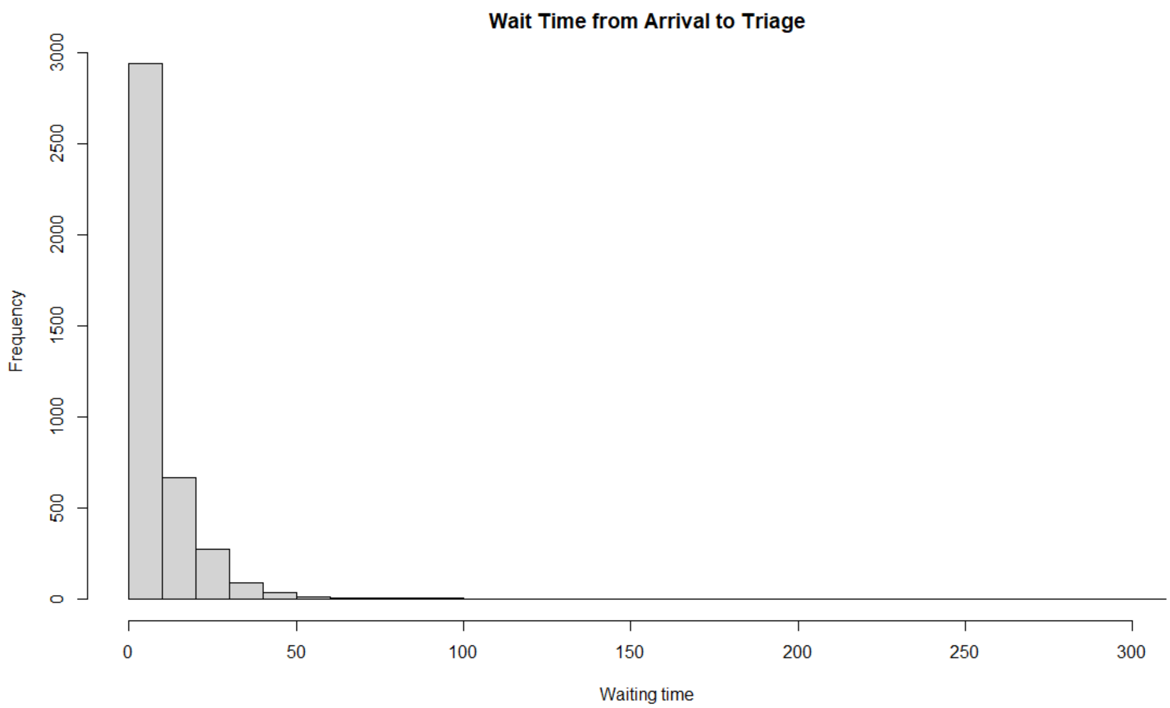
By Pearson's product-moment correlation test, the null hypothesis is that the correlation is zero and we fail to reject the null. Therefore, the correlation between the two should be zero. LOS is not dependent upon the arrival time.




```
##
##  Pearson's product-moment correlation
##
## data:  x and averagelos
## t = -1.5744, df = 22, p-value = 0.1297
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.63951706  0.09772929
## sample estimates:
##        cor
## -0.3182125
```

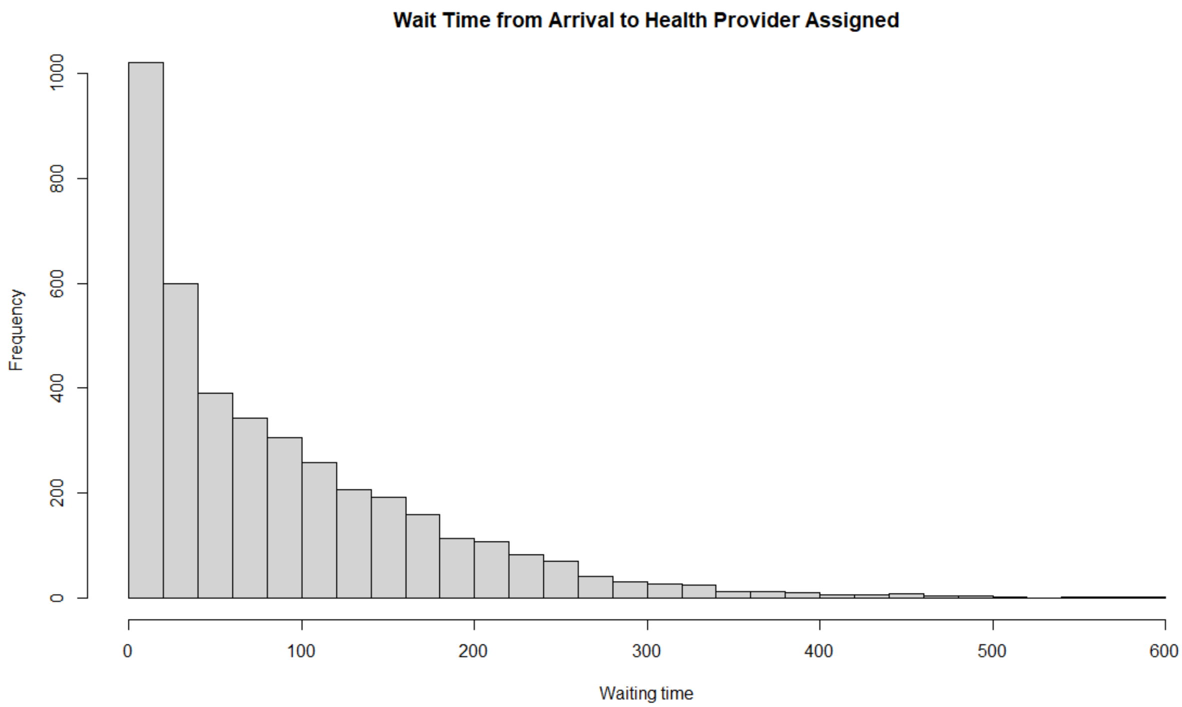
## Part III. Wait time

### (a)




```
## [1] "mean is 8.21724052514243"
## [1] "variance is 142.994172462698"
```

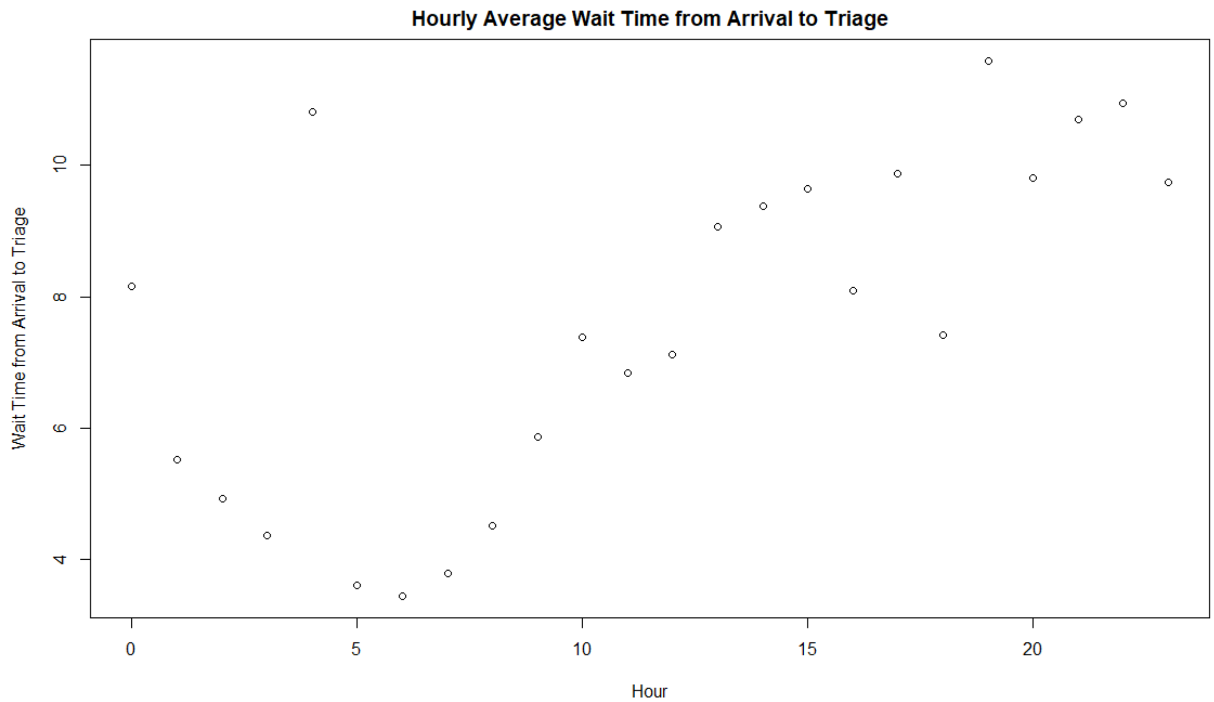
## (b)



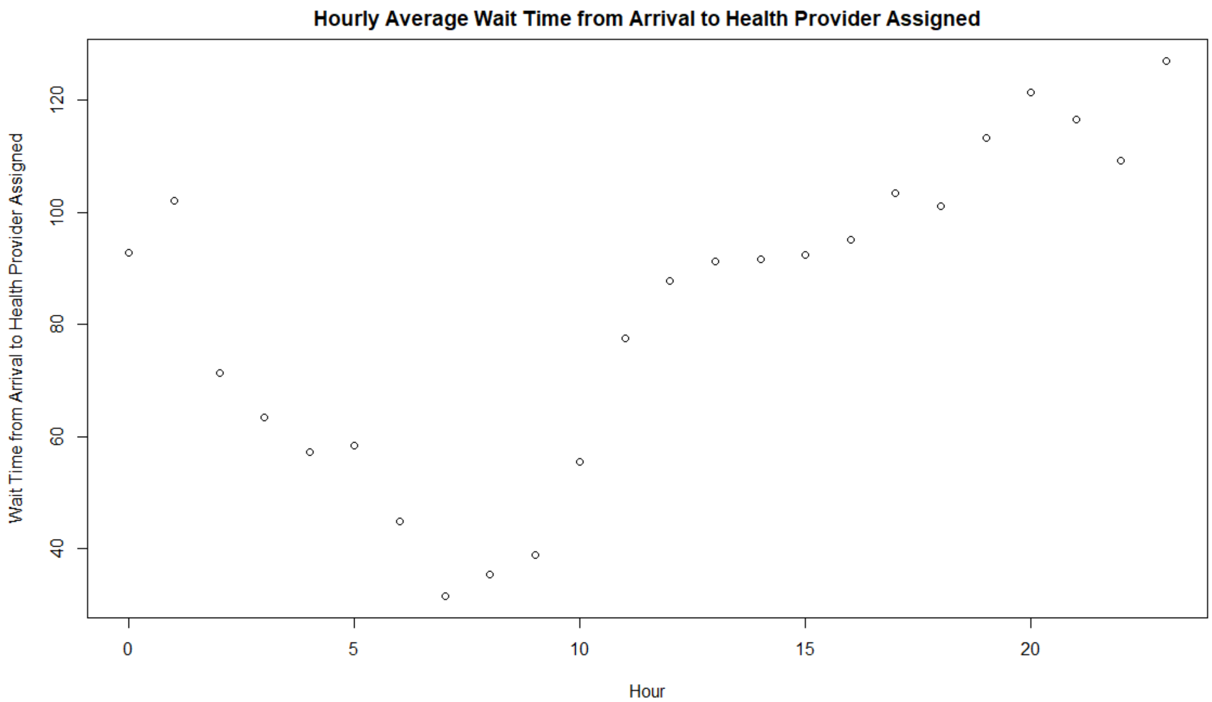

```
## [1] "mean is 87.667079514491"
## [1] "variance is 7452.42828415943"
```

## (c)

By Pearson's product-moment correlation Test, the p-value is small enough. Therefore, the two variables both depend on the arrival time.



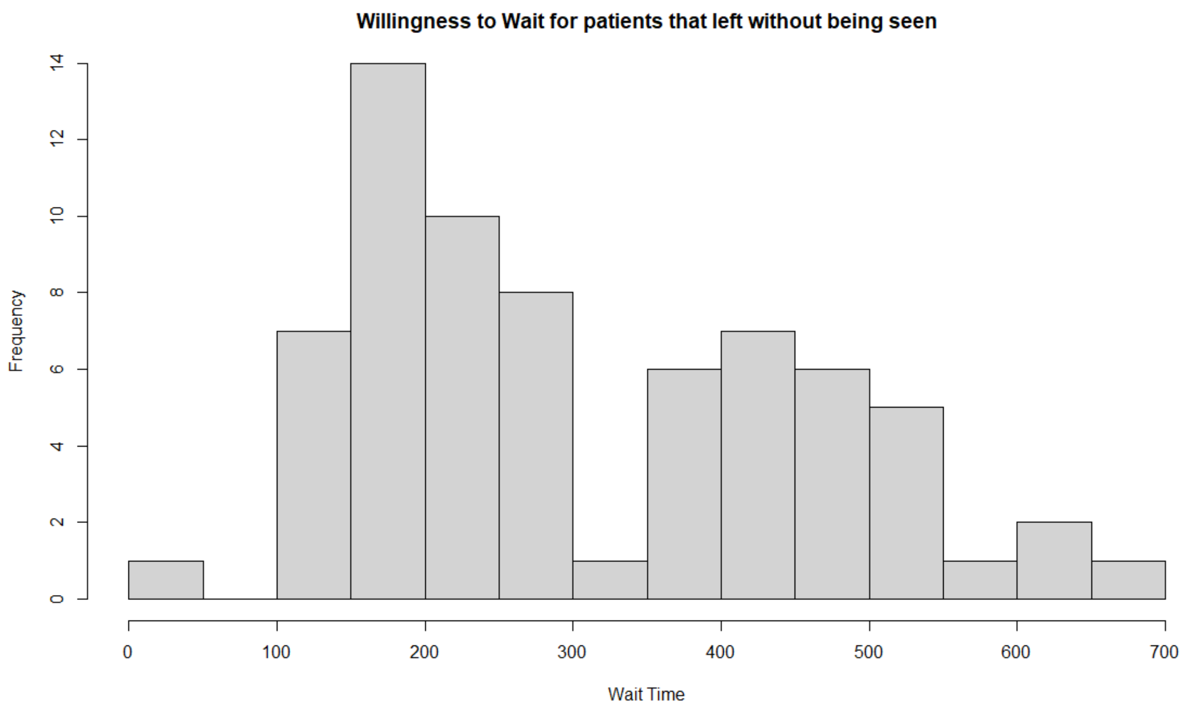

```
##
##  Pearson's product-moment correlation
##
## data:  x and averagetri
## t = 4.4659, df = 22, p-value = 0.0001933
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.3964426 0.8550985
## sample estimates:
##       cor
## 0.6895621
```




```
##
##  Pearson's product-moment correlation
##
## data:  x and averagepro
## t = 4.1935, df = 22, p-value = 0.0003761
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.3598749 0.8431991
## sample estimates:
##       cor
## 0.6665124
```

# Part IV. Patients willingness to wait

## (a)

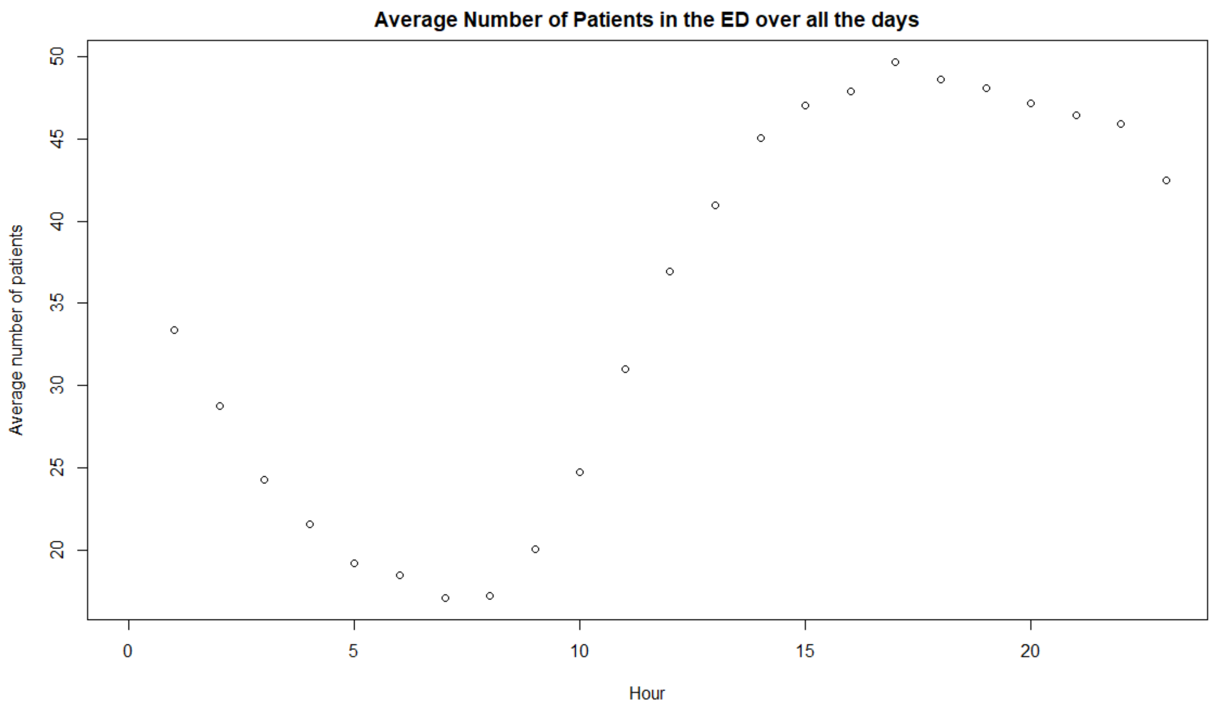



## (b)

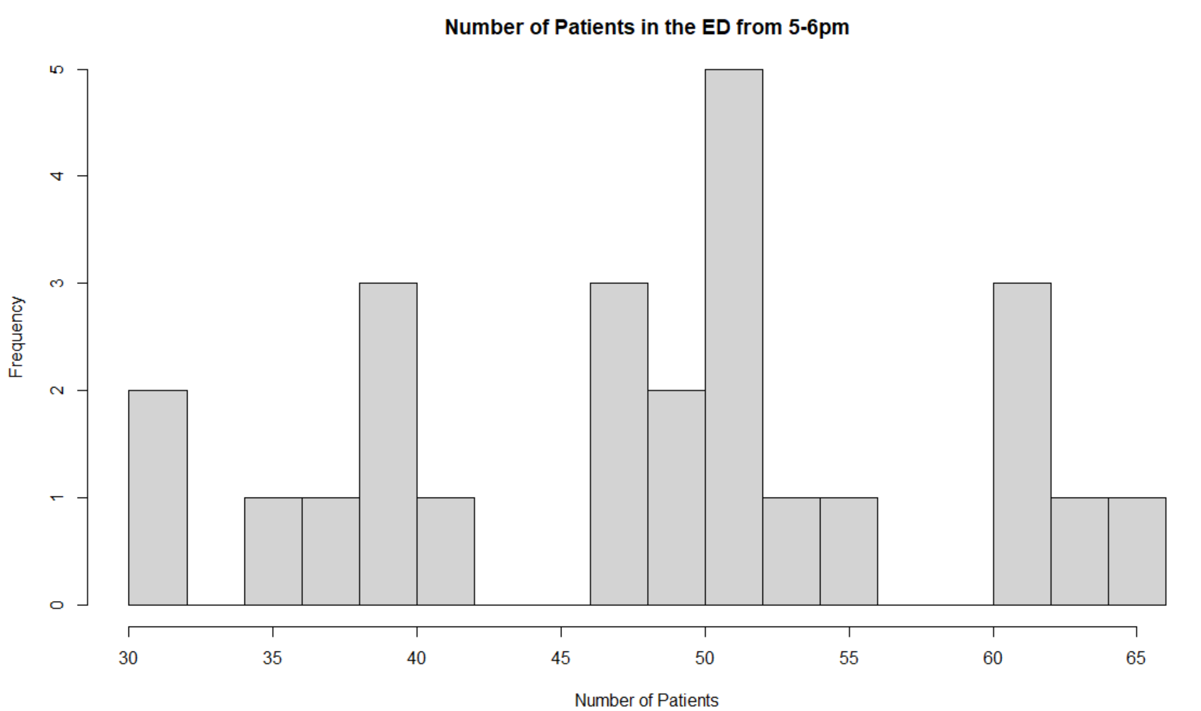
We can observe that it follows bimodal distribution.

## Part V. Number of patients in the ED

### (a)

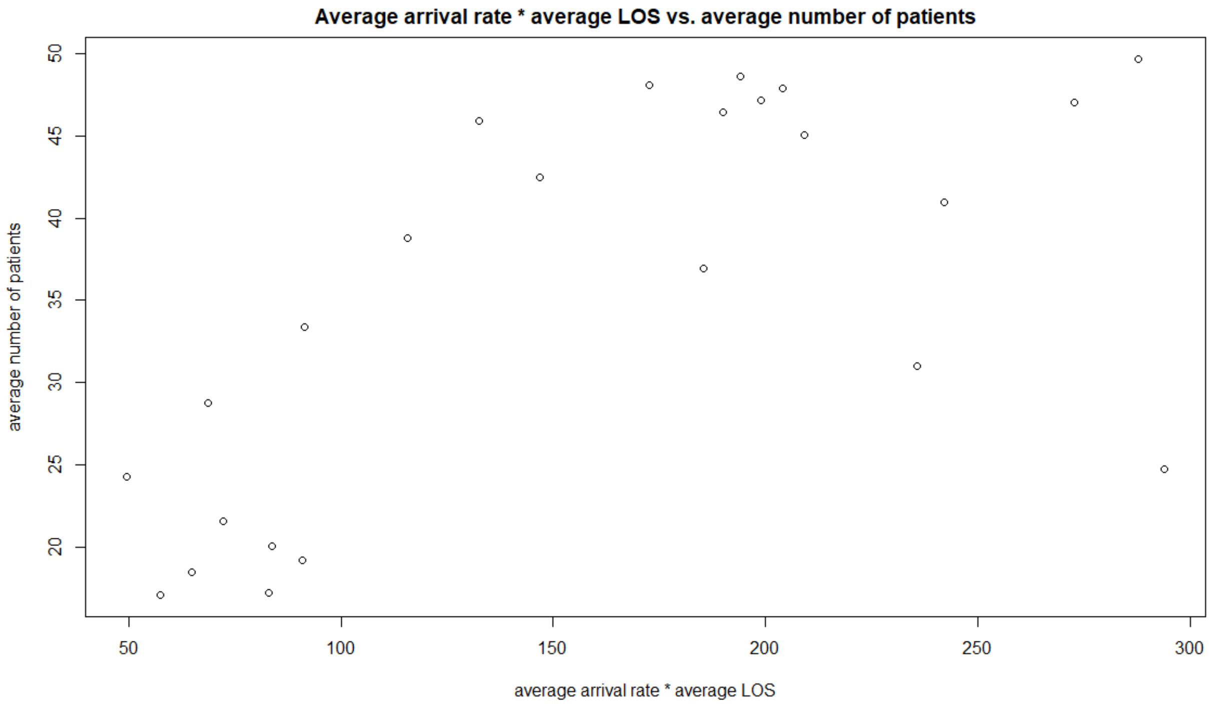


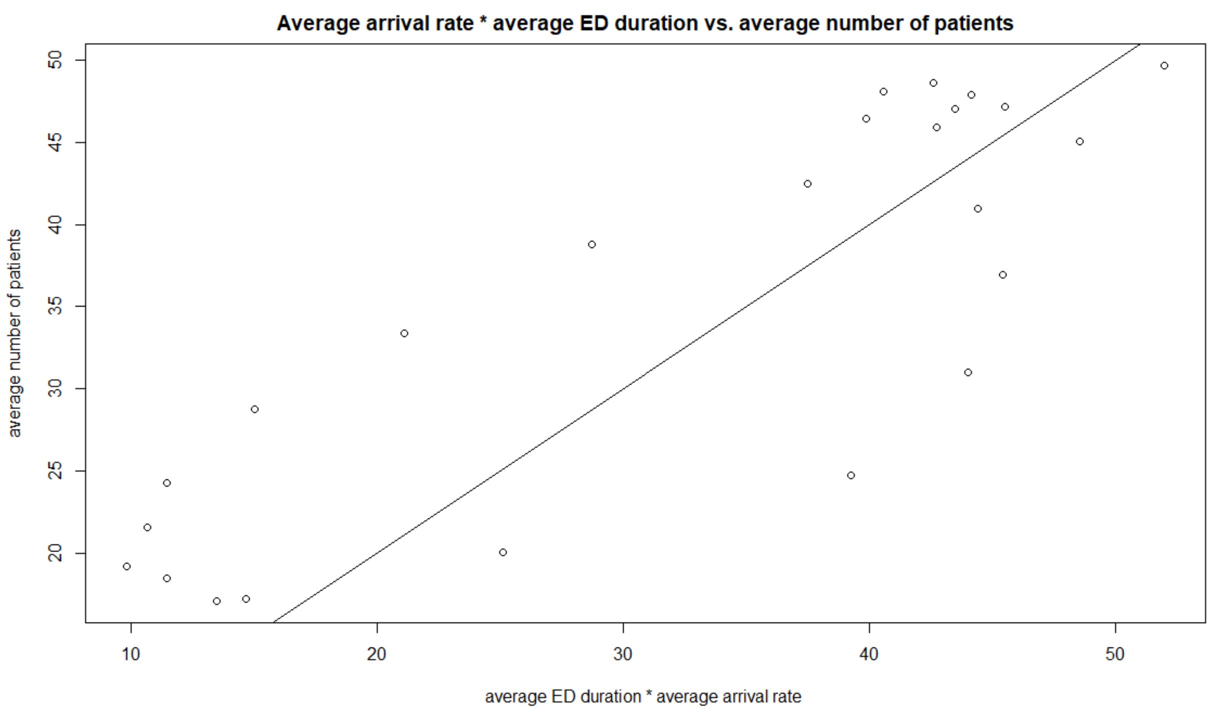

### (b)



### (c)

We can observe that the plot does not show a pattern of line y=x, so the Little's law does not hold. The possible explanation might be that LOS value is too large. If we use `ED_DURATION`, we can observe that the Little's Law holds as the data show some pattern of line y=x.

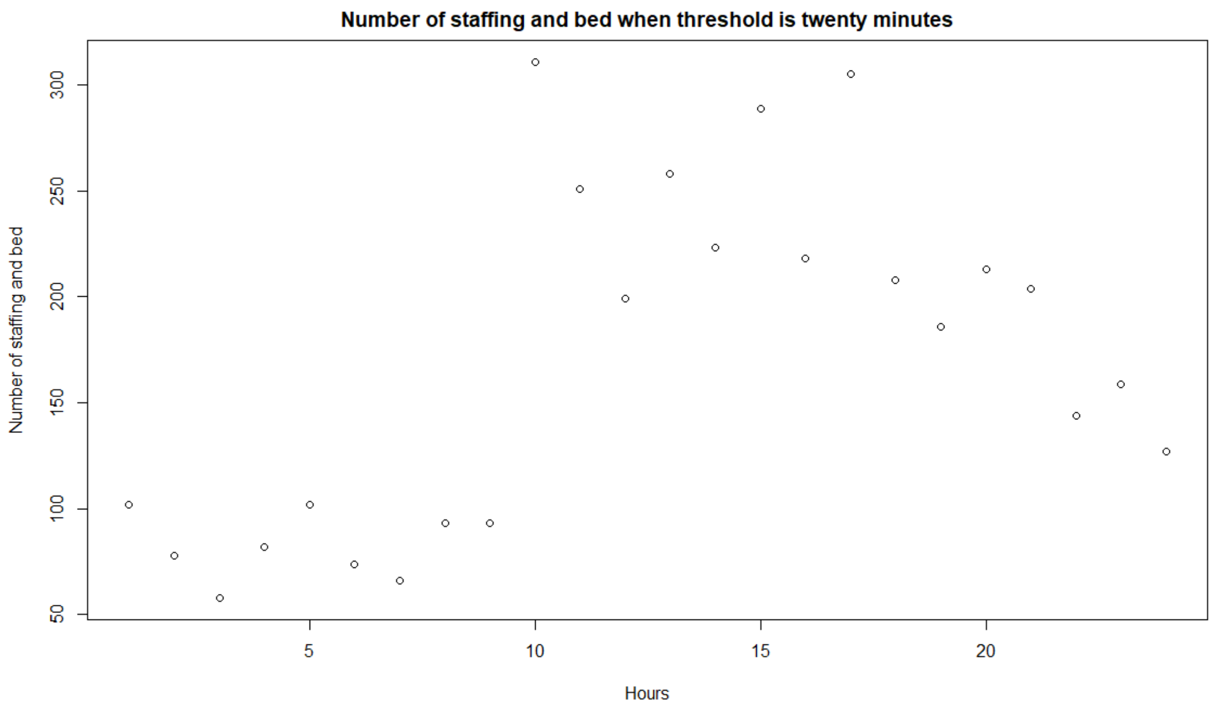


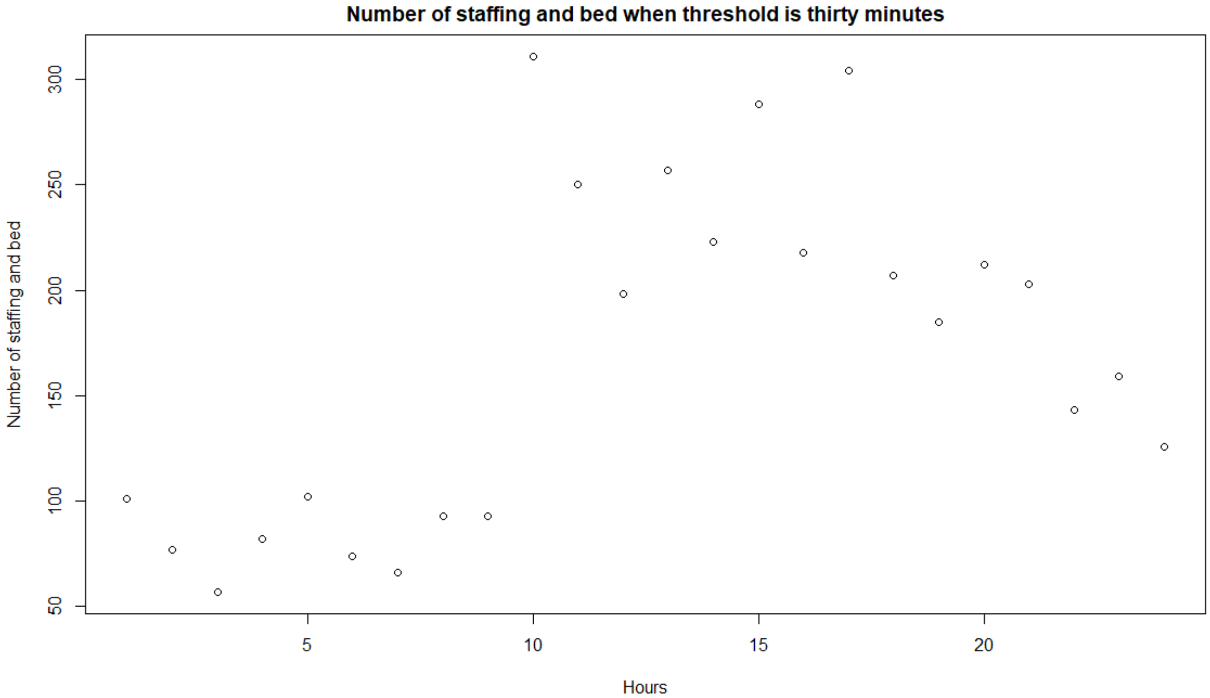



# Part VI. Staffing and capacity decisions

## (a)

Applying Erlang C formula from the website, the staffing and bed capacity strategy is printed below. The first row is if the standard is 20 minutes, and the second row is if the standard is 30 minutes. The data is different from the hourly average number of patients in the ED.






```
##  [1] 102  78  58  82 102  74  66  93  93 311 251 199 258 223 289 218 305 208 186
## [20] 213 204 144 159 127
```

```
##  [1] 101  77  57  82 102  74  66  93  93 311 250 198 257 223 288 218 304 207 185
## [20] 212 203 143 159 126
```

## (b)

The staffing decision of nurses hourly is printed below.

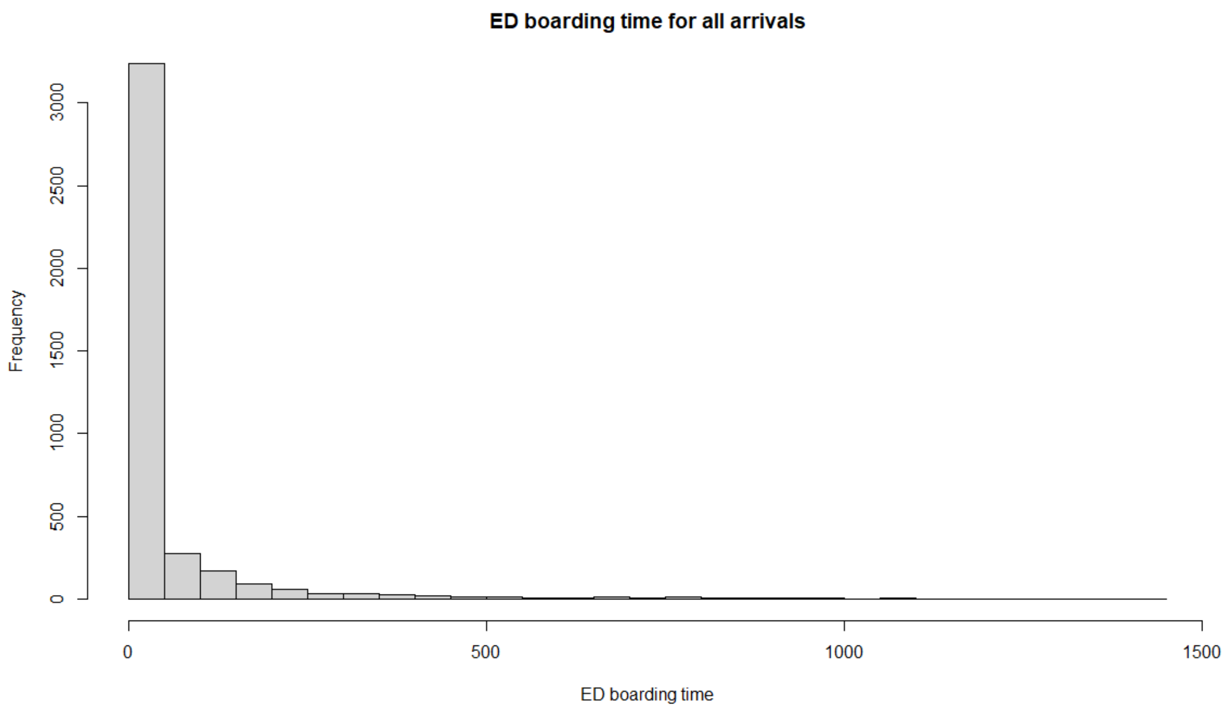
```
##  [1] 41 40 39 39 39 39 39 39 40 41 42 43 43 43 43 42 43 43 43 43 43 43 43 43
```

## (c)

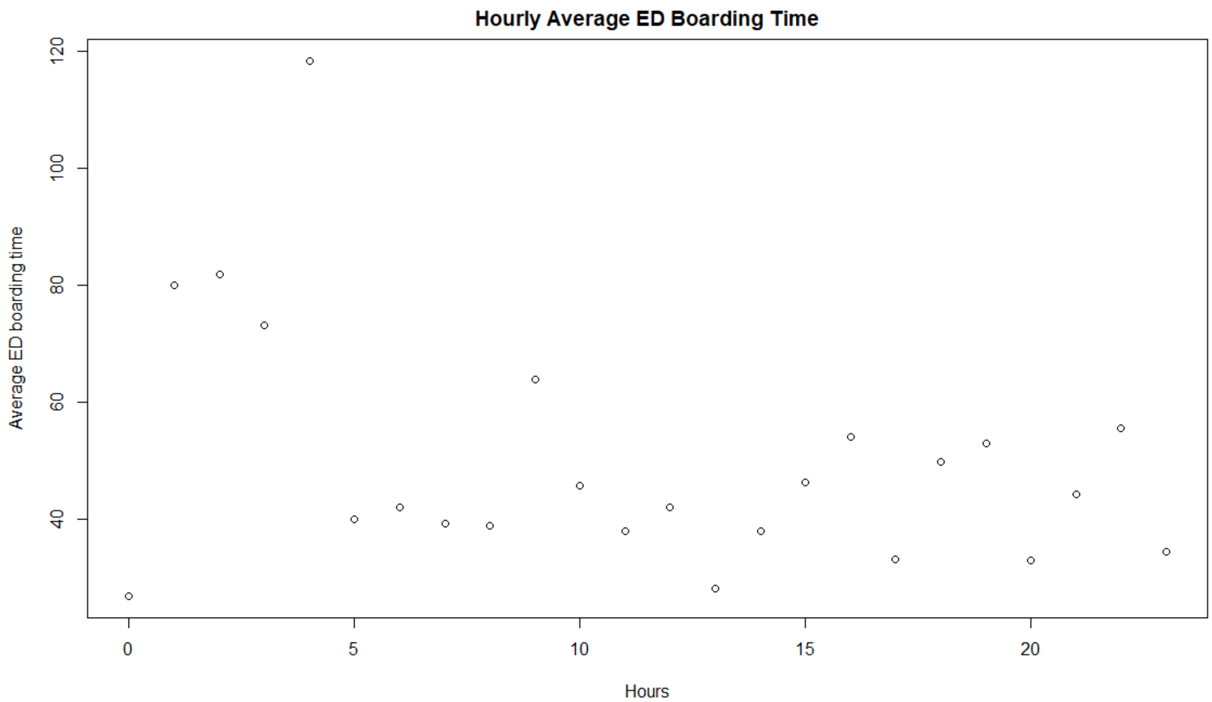
To make more realistic model, I think we need to consider the differentiation of cares needed for the ED arrivals. Maybe it is a good idea to set up different models for each triage because different degrees of urgency need different expected waiting time. For each arrival, first classify his or her degree of urgency. Then, put them into different subsets of the data. Applying Erlang C formula for each subset, and we then get the hourly number of agents needed for different levels of wounds or illnesses.

# Part VII. ED-Ward network

## (a)




```
## [1] "mean is 45.2006440426059"
## [1] "variance is 17640.379454491"
```




## (b)

Many variables are needed for this two-time-scale Markov Model. In this model, LOS and time-of-day that patients are discharged are two i.i.d sequences. The service hour is no longer `LOS=ED_Duration+InpatientLOS` from our dataset. It depends on time-of-day that patients are admitted, LOS, and the time-of-day that patients are discharged. The hourly mean waiting time is now assumed to be geometrically distributed. We can use geometric distribution to estimate the hourly mean waiting time. In this model, discharge has its own independent distribution. For each hour now from 1 to 24, take the hour over 24 (i/24), which is our discharge distribution. We can put this baseline discharge distribution earlier or later to test the model and check system congestion.

## (c)

Because the wards are in days, but the ED is in hours, after each hour, there will be some hospitalized patients that can be transferred from ED beds to the wards. If keeping those patients in the ED, there will be congestion. Therefore, the transition unit is designed hourly for those patients who should be transformed from the ED, but they can only stay in the transition unit because discharge decisions in the wards are made daily. We will the need the average handling time in the transition unit, which would be the difference in hours between discharge time from the ED beds and arrival time to the wards. Besides the average handling time, we also need the average arrival rate to the transition unit, which is the average discharge rate from the ED. Then we can apply Erlang C formula to find the hourly average capacity for the unit. The performance metrics would be the level of the waits for the unit less than certain time according to Erlang C formula.