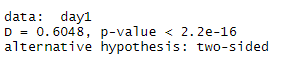
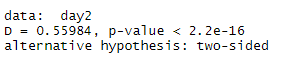
**Distribution Fitting**

1. **Kolmogorov-Smirnov Test**

The data for customer inter-arrival times was examined for 0 to 400 seconds in order to test the validity of the manager’s claim that the data is uniformly distributed. The significant level of 0.05 was used to select the right factor from the Kolmogorov-Smirnov Table. The test is such that the maximum difference in probability distribution functions of the distribution and that of a uniform distribution is compared to the critical value. If the difference is less than the critical, then it is a good fit. Otherwise, it is a bad fit. The probability density distribution was computed in excel for the data provided. A uniform distribution centered around 0.5 was also computed. After comparisons, the result revealed that the distribution is a bad fit. Hence, it is not distributed uniformly, and the manager’s claim is wrong.





1. **Descriptive statistics**

Important descriptive statistics such as mean, median, mode, standard deviation, variance, etc that are important in distribution fitting were computed. The result table is shown below:

|  |  |  |  |
| --- | --- | --- | --- |
|  | Day 1 | Day 2 | Day 1 & 2 |
| MEAN | 55.16803 | 64.22382 | 59.69128 |
| STANDARD DEVIATION | 56.10851 | 58.57514 | 57.50309 |
| MEDIAN | 34 | 45 | 39 |
| MODE | 16 | 14 | 16 |
| SAMPLE SIZE | 488 | 487 | 975 |
| VARIANCE | 3148.165 | 3431.047 | 3306.606 |
| KURTOSIS | 6.364401 | 3.121416 | 4.487857 |
| SKEWNESS | 2.250756 | 1.708377 | 1.952589 |

1. **Frequency Histograms**

Frequency histograms were plotted for 5, 10 and 20 second intervals. The diagrams show that customer arrival increases with inter-arrival time.

For the entire data,

1. **Chi-Square Test**

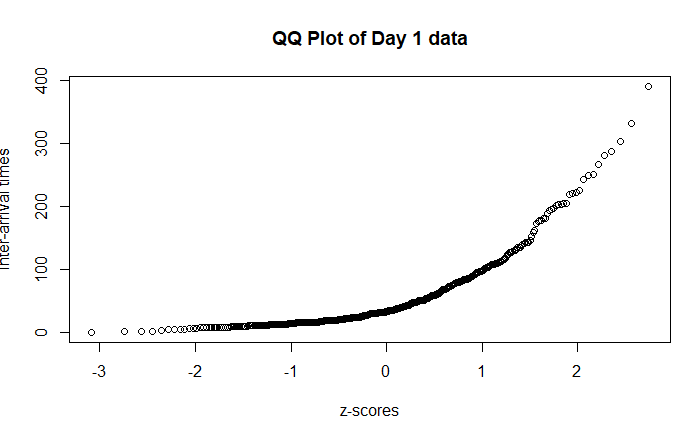
The parameters for the Chi-Square test are:

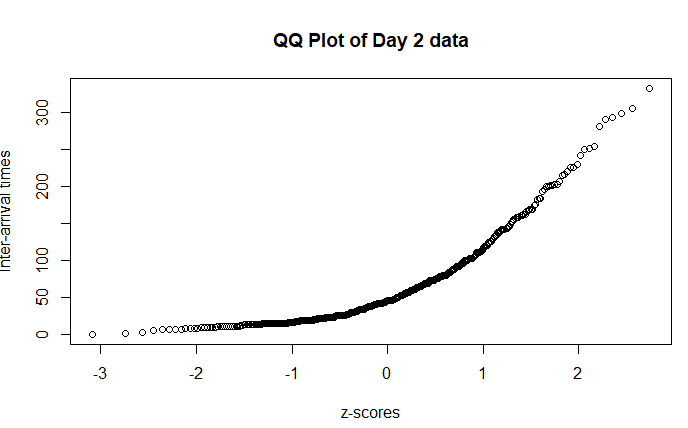
1. The significance level is 0.05
2. The actual data is for day 1 and 2 inter-arrival times
3. The estimated/predicted data is 10 second interval

The Chi-square value was zero which is less than the significant level of 0.05. Thus, the data is statistically significant and made up of values that are independent of each other.

1. **QQ Plot**

The shapes of the QQ plots show exponential behaviour. If it is a normal distribution, the shape will be close to a straight line. The data for the plots were computed in Excel and imported into R-studio for plotting. The plot file is included with this submission. The QQ plots for Day 1 and 2 are shown below:





1. **Plot of inter-arrival times with respect to observation times**

The plot for Day 1 and 2 were done in Excel and examined carefully for stationarity. There was no obvious cyclic pattern, trend or regularity. The data can be said to be stationary.

1. **Autocorrelation**

The existence of autocorrelation in the data is a sign that the model may be unsound. Lag 1 and lag 2 differences can help to determine if the data is autocorrelated or not. If the data is correlated with itself, we will get a value of unity. So, values close to zero means that the correlation of the values are low. Lag 1 difference is 0.008768 while the one for lag 2 is 0.023274. It can be concluded that autocorrelation in the data is low or almost non-existent. The lag 1 and lag 2 differences are plotted below:

