CZ4015

Final Report

Author:

Pavel Jahoda (N1800740K)
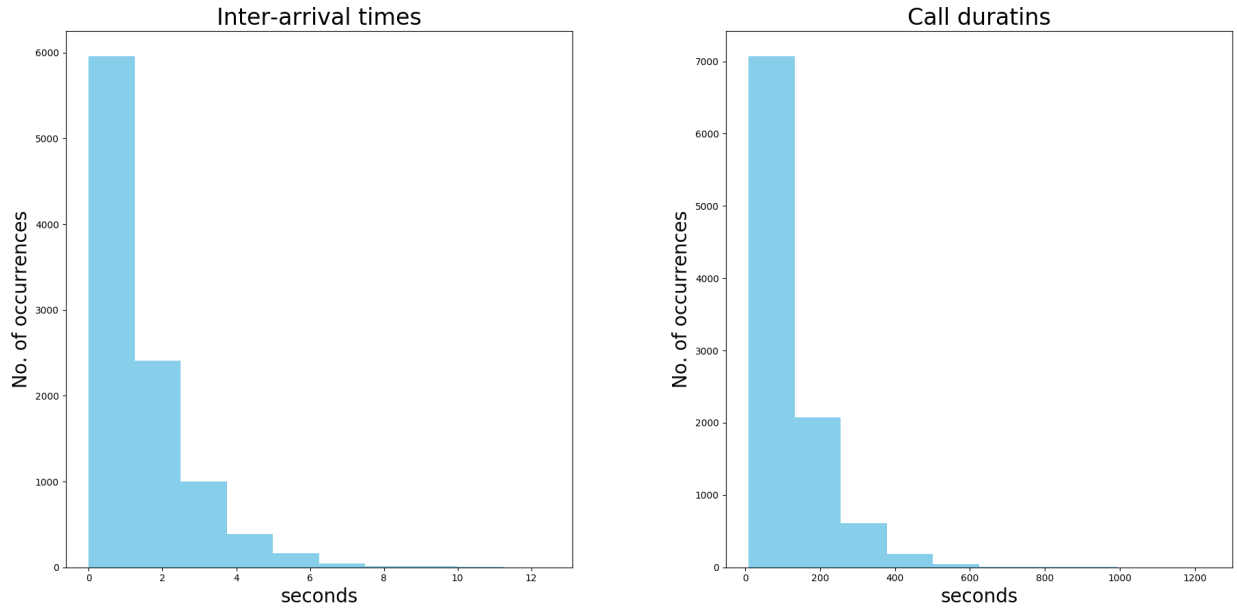
Input Analysis

From the histograms below we can see similary between the distributions of inter-arrival times and call durations. Both of these histograms resemble probability density function of exponential distribution.

First, we will estimate the parameter of the exponential functions using maximum likelihood estimation. The exponential function is denoted as:

$$\lambda \cdot \mathrm{e}^{-\lambda \cdot x} \tag{1}$$

The likelihood function for the parameter lambda given x_1, x_2,..,x_n is denoted as:

$$\mathcal{L}(\lambda | x_1, x_2, .., x_n) = \lambda^n \cdot \mathrm{e}^{-\lambda \cdot \sum_{i=1}^{n} x_i} \tag{2}$$

To find the lambda for which the likelihood function is maximal, we differentiate by lambda and solve for lambda when the derivate is equal to 0, which results in the following formula:

$$\lambda = \frac{n}{\sum_{i=1}^{n} x_i} \tag{3}$$

Using maximum likelihood function of an exponential distribution we have calculated lambda for the inter-arrival times to be 0.73 and lambda for the call duration times to be 0.009.

On the other hand, the two histograms below show two different distributions. The histogram on the left shows that the stations where the cars are located when the call begins are uniformly distributed from 1 to 20. The histogram of the car velocities (on the right) resembles normal distribution.
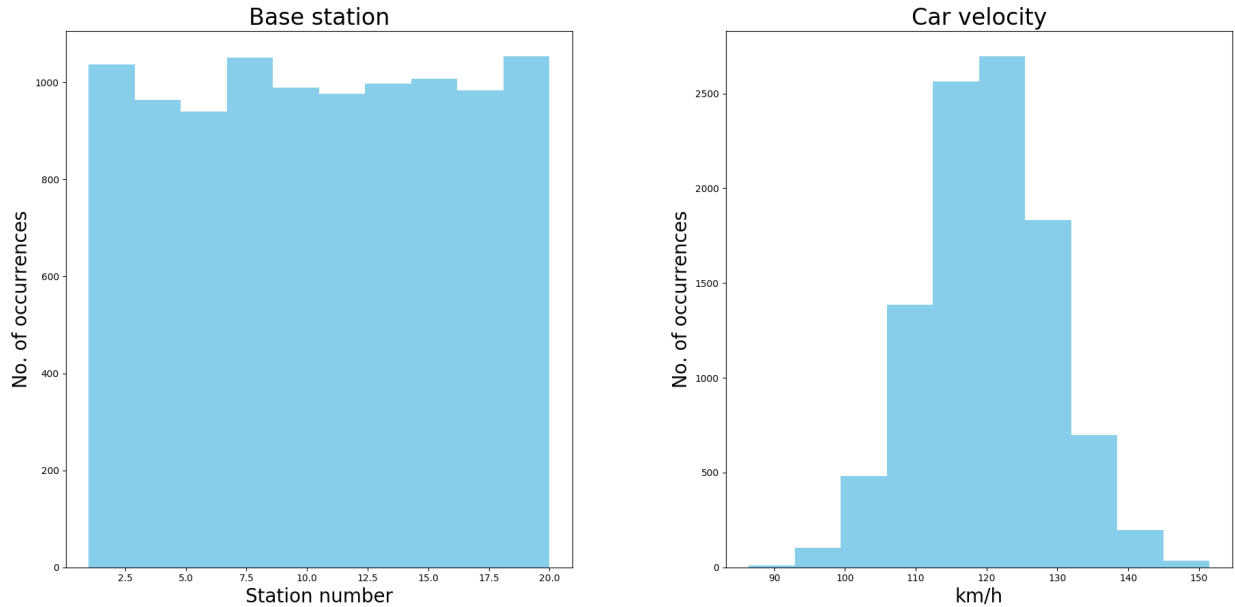
Similarly as with the exponential distributions above, we will use maximum likelihood method to calculate the parameters of the normal distribution (car velocities). We get $\mu = 120.07$ and $\sigma^2 = 81.33$.