To modeling the performance of the service unit, Here, we are using 3 different models to simulate the performance of our System. The first model we use is an M/M/1 model, this is because our system only has one service. The second model we use is the best fit distribution model, use the inter-arrival times best fit ( with exponential distribution) and the service times best fit (with gamma distribution). The third, M3 model uses the empirical distribution of the inter-arrival time and the service time to simulate the performance of our System.

For M1 and M2, I did write a Simpy code to simulate those two models. Where lamb and mu I set is 1/ E(inter-arrive time) and 1/E(service time). The time unit I use is second because this is easy for me to clarify the difference between two models. The number of customers and maximum time I set is a very larger number like 10000, it should be given enough sample to do the simulation.

On the basis of M1 model I build up the M3 model, I was added load data functions and empirical distribution functions. Changed the distribution to empirical distribution of the inter-arrival time and the service time. To build empirical distribution function I have to get a list of independent, identically distributed real random variables(X) and the common cumulative distribution function (F), what I did is remove the duplicate values in queue time list and service time list, Which will become a new Service-timeX list and inter-arrive timeX without the duplicates. Such as (X1…Xn). Then the common cumulative distribution function F(x) defines a number of elements in the sample < x / total number of elements. Once we got the X variables and F variables we can use the formula W[k-1]+((r-F[k-1])/(F[k]-F[k-1]))\*(W[k]-W[k-1]) to get the simulate service time or queue time for each customer. Where r is a random number between 0 to 1. And the k value is the index of F list that larger than r.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | M1 | M2 | M3 | **observation system** |
| W | 52.370062977183 | 47.832741165061 | 65.012917049309 | 67.86803475 |
| Wq | 16.971467155669 | 12.396526048179 | 29.676469406179 | 32.73492416 |
| WS | 35.398595821513 | 35.436215116882 | 29.676469406179 | 35.13311059 |
|  |  |  |  |  |
| L | 0.4794858980176 | 0.4384528303535 | 0.5886484886771 | 0.6213550757 |
| Lq | 0.1554012774302 | 0.1136429753208 | 0.2687349171198 | 0.2996994292 |
| LS | 0.3240846205874 | 0.3248098550327 | 0.3199135715573 | 0.3216556465 |

Compare with M1and M2 model, there are quite similar. The only difference is the service times best fit gamma distribution, the inter-arrival distributions are the same. So, their output is quite similar as well. The waiting time in the queue between two models is just a few seconds. But to compare with the W in the system(67 seconds), which has a larger gap between.

I think this is because during the simulation we are using the exponential distributions such as assume all the customer arrive continuously and independently at a constant average rate. On the other hand, our observation system will have a rush time, normally it occurs during the break time between the lectures. In this time will have a lot of students come to collect their homework, that leads to suddenly increase the waiting time in the queue. The evidence is shown in the data, there have 25 people wait more than two min, it all occurs during the rush time. It also shows from the LQ, which is our observation system has the more average number of people in the queue.

I think we can not avoid this rush time problem, But we can improve by collect data in different time, especially when the service is not that busy. Like, we can collect the data from the system start server unlit the system end of the server, for all business hours. Not just the collecting the data during the middle period.

Overall, We find out the M3 model have the best performance overall. That is because empirical distribution is to do the simulation base on the observation data. In the above image, the green line is empirical distribution function, the blue line is the corresponding sample. As long as we have collected enough data, and do the data cleaning properly. The more standard data will have, the empirical distribution function we get will more close to the corresponding sample. Therefore the result we get will be more accurate.