IE 580: Homework #1

Omar Zu’bi
Due Date: 11/12/2020

Problem 1: Evaluate the four approximation methods of G/G/1 model using simulation. Please define your own evaluation approach.

In order to evaluate the four approximation methods of G/G/1 models, an arena simulation (Q1.HW1.doe) has been created using multiple . The model utilizes a single server, an infinite maximum number of jobs with a queuing discipline of first come first served (FCFS). The simulation is then run for a number of repetitions for different service and arrival distributions. The simulation is then compared to the approximations formulas calculations shown in *table 1*. The simulation was allowed to run for ten replications for 150 intervals. For the simulation, $I(x)$ is the parameter for interarrival times and $S(x)$ as a parameter for service time. Both $I(x)$ and $S(x)$ are independent inputs to test the approximations. For this example, the server's service time was set to norm(8,9) and the customer arrival times were set to norm(10,11).

The following formulas are estimates for the value of L_q for G/G/1 models:

Approximation 1: $L_q(G/G/1) \approx \left(\frac{C_a^2 + C_s^2}{2}\right) \bullet L_q(M/M/1) = \left(\frac{C_a^2 + C_s^2}{2}\right)\left(\frac{\lambda}{\mu-\lambda} - \frac{\lambda}{\mu}\right)$

Approximation 2: $L_q(G/G/1) \approx \left[\frac{\rho^2(1+C_s^2)}{1+\rho^2 C_s^2}\right]\left[\frac{C_a^2 + \rho^2 C_s^2}{2(1-\rho)}\right]$

Approximation 3: $L_q(G/G/1) \approx \left[\frac{\rho^2(1+C_s^2)}{2-\rho+\rho C_s^2}\right]\left[\frac{\rho(2-\rho)C_a^2 + \rho^2 C_s^2}{2(1-\rho)}\right]$

Approximation 4: $L_q(G/G/1) \approx \left[\frac{\rho^2(C_a^2 + C_s^2)}{2(1-\rho)}\right] + \left[\frac{(1-C_a^2)C_a^2\rho}{2}\right]$

Where, C_a^2 : *squared coefficient of variation of interarrival times.*

C_s^2 : *squared coefficient of variation of service times.*

ρ : *is the traffic intensity.*

C_a , C_s , and ρ can all be calculated using the following formulas:

$$\rho = \frac{\lambda}{s\mu}$$

$$C_x^2 = \frac{\sigma_x^2}{(E[X])^2}$$

Where λ is the average arrival rate, μ is the average service rate, and s is the number of servers.

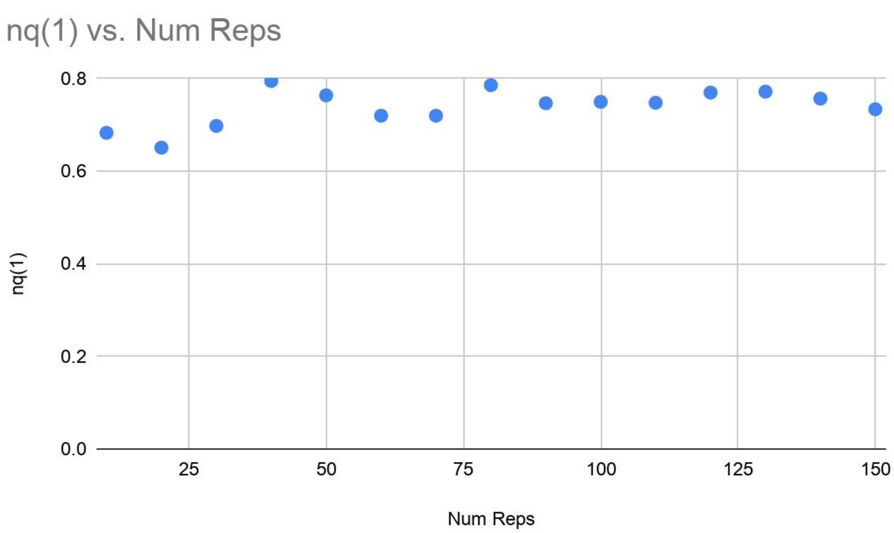
Using the following random parameters given, the calculations for the approximations are as follows:

Arrival Model $I(x)$	**Service Model** $S(x)$	C_a^2	C^2	ρ	**Approximation 1**	**Approximation 2**	**Approximation 3**	**Approximation 4**
NORM(10,11)	NORM(8,9)	1.21	1.27	0.8	3.961	4.046	3.230	3.859

Using the process analyzer, we get the following values for nq(1) for the different number of repetitions. This can be seen in *Table 1.*

Num Reps	nq(1)
10	0.683
20	0.651
30	0.698
40	0.795
50	0.764
60	0.72
70	0.72
80	0.786
90	0.747
100	0.75
110	0.748
120	0.77
130	0.772
140	0.757
150	0.734

Table 1: Reps vs. nq(1)



As seen on the scatter plot, the value for nq(1) stabilizes and eventually converges as the number of repetitions increases.

Problem 2: Consider a finite population of N customers. There is a single server who serves customers in a queue (FCFS) with EXPO(0.5) amount of time. Each customer after completion of service spends EXPO(1) time outside and comes back to the queue. Assume that all customers are initially arriving to the queue following EXPO(1). Plot average waiting time (waiting time in queue + service time) against N (10, 20, 30, 40, 50, 60, 70, 80, 90, 100).

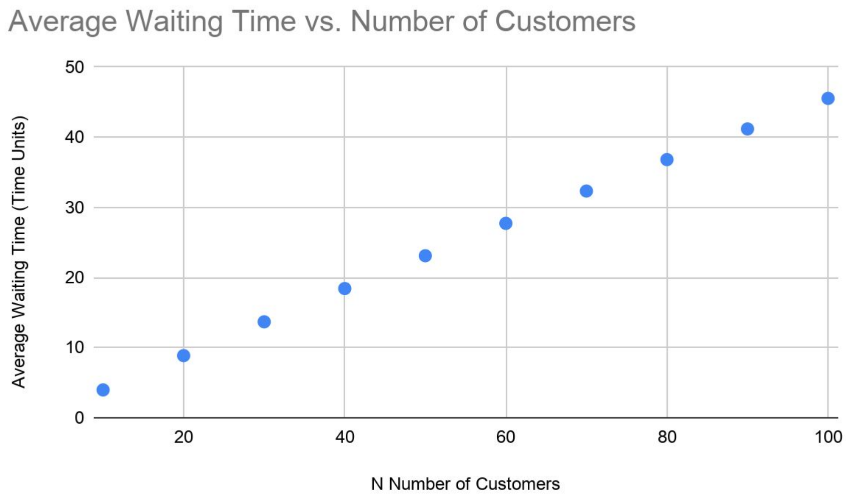
The simulation (Q2.HW1.doe) was created to have a single server and a queuing priority of first come first serve with a service time of EXPO(0.5). The simulation was configured to replicate ten times and have a runtime of 1000 time units. *Table 2* shows the average waiting times for the different N customer sizes. The data was collected using the process analyzer tool.

N Number of Customers	**Average Total Waiting Time (Time units)**
10	3.974
20	8.864
30	13.685
40	18.428
50	23.103
60	27.721
70	32.326
80	36.810
90	41.181
100	45.538

Table 2: Number of Customers and Corresponding Average Total Waiting Times

The data was also plotted in *figure 1*, which shows a linear relationship between the average waiting time and N customers; as the number of customers N increases, the average waiting time also increases:

Figure 1: Scatter Plot of Average Waiting Time Vs. Number of Customers



Problem 3: Consider a call center where customers arrive one by one according to EXPO(1). We have two design alternatives as follows:

- Design A (One skilled operator)
  M/M/1, service time ~ EXPO(0.5)
- Design B (Two novice operators)
  M/M/2, service time of each server ~ EXPO(1)

Compare traffic intensity of the two alternatives, and evaluate the two alternatives in terms of "average time in system" and "average number of jobs in system" based on the simulation parameters as shown below.

Given the parameters of the problems, a simulation was created using arena (Q3.HW1.doe). The configurations for designs A and B are very similar, except for the number of servers and their service times. For both designs, the runtime was controlled by creating 150 jobs, and the simulation was replicated 10 times. The traffic intensity for design A and B can be calculated using the following formula:

$$\rho = \frac{\lambda}{s\mu}$$

Where λ is the average arrival rate and μ is the average service rate. Using the given parameters for the service and arrival rates for designs A and B, the following traffic intensities can be calculated:

$$\rho_a = \frac{1/1}{1(1/0.5)} = 0.5\ ,\ \rho_b = \frac{1/1}{2(1/1)} = 0.5$$

*Note the following equations are used to find the service and arrival rates.

$$\mu = 1/h, \lambda = 1/g$$

Where, h : *the average time to serve a customer.*

g : *the average time between 2 arrivals.*

From the calculations, designs A and B have the same traffic intensities. Despite this, the average time a customer spends in the system is different for both designs. In addition, the simulation tallied the number of jobs in the system throughout the simulation run and found some differences. These differences are reflected in *table 3 below:*

	Average Time in System (time unit)	**Average number of jobs in system**
Design A	1.045	1.099
Design B	1.358	1.391

Table 3: Data collected for Designs A and B

According to the table, customers in design A spend less time in the system, on average, than customers in design B. In addition, design A has, on average, a lower number of jobs in the system than design B. Based on these results, design A is the better of the two models.

Problem 4: There are three types of jobs and each type of jobs enters a queue specifically designated for that job type (so we have 3 queues). All the three queues operate in FCFS. Two machines serve all the three types of jobs and jobs in different queues are prioritized according to:

- Lowest average service time First
- Lowest squared coefficient of variation of service time First
- First Come First Served

Job Type	**Interarrival Time**	**Processing Time**
1	EXPO(1)	NORM(0.2,2)
2	EXPO(1)	NORM(0.1,3)
3	EXPO(1)	NORM(0.4,5)

Evaluate the three priority rules in “average time in system” and “average number of jobs in system”, based on the simulation parameters specified below.

The simulation was built using Arena for all different queuing priorities (Q4.Case1.doe, Q4.Case2.doe, Q4.Case3.doe). Given the parameters from the problem, we are able to assign different priorities for each job type based on the processing times; this is reflected in the simulation .doe files where the assignment of A(3) in the Assign block for the first two cases. For the first case, given that average service times, the queue priority from highest to lowest is: job type 2, job type 1, followed by job type 3.

For the second cases, the queue priority was assigned for the lowest squared coefficient of variance of the different service times. The squared coefficient of variance can be calculated using the following formula:

$$C_s^{\,2} = \frac{\sigma_s^{\,2}}{(E[S])^2}$$

From the formula, the squared coefficient of variance for each job type is as follows:

$$C_s^{\,2}{}_{,\ type\ 1} = 2^2/(0.2)^2 = 100$$
$$C_s^{\,2}{}_{,\ type\ 2} = 3^2/(0.1)^2 = 900$$
$$C_s^{\,2}{}_{,\ type\ 3} = 5^2/(0.4)^2 = 156.25$$

From these calculations, the queue priority for the second case from highest to lowest is job type 1, job type 3, followed by job type 2. For the third case, the queue prioritization is first come first served. To accomplish this in Arena, the seize block was assigned to have the corresponding attribute for FCFS in its priority field; this is different for what was done in the first two queue prioritization cases.

The simulation run time was controlled by the number of jobs the simulation had to run; 50 jobs each for the different job types, a total of 150 jobs. The simulation was replicated ten times. From the simulation, data was gathered on the average time in the system and the average number of jobs in the system for all three cases using the process analyzer tool. This is reflected in *table 4* below.

	Average Time in System (Time units)	**Average Number of Jobs in the System**
Case 1: Lowest average service time	24.433	32.311
Case 2: Lowest squared coefficient of variation of service time first	32.394	42.928
Case 3: First Come First Served	32.709	43.308

Table 4: Data gathered for the three different priority cases for Average time and Jobs

For all three cases, case one prioritization has the lowest average time in the system and average number of jobs in the system.