International Journal of Statistics and Applied Mathematics 2018; 3(2): 21-26

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2018; 3(2): 21-26

www.mathsjournal.com
Received: 12-01-2018
Accepted: 15-02-2018

R Anbarasi
Assistant Professor,
Kg College Of Arts and Science,
Coimbatore, Tamil Nadu, India

Correspondence
R Anbarasi
Assistant Professor,
Kg College Of Arts and Science,
Coimbatore, Tamil Nadu, India

A study on M/M/3 queuing model on waiting time reduction in a local health care centre

R Anbarasi

Abstract
Queuing theory is a mathematical approach to the study of waiting lines. Long waiting time in a health care system indicates the lack in management of the system. This project is an attempt to analyze the use of queuing theory in a local health care clinic. Also considers the clinic as a multi server queuing system following Poisson arrival and Exponential service based system. A comparison of the results using MATLAB pertaining to single server and multiple server queuing system is also provided.

Keywords: M/M/3, local health care centre, Queuing theory

Introduction

Queuing theory has its origin in research by a Danish telephone engineer Agner Krarup Erlang when he generated models to describe the Telephone Exchange of Copenhagen." The theory of Probability and telephone conversations" is the first paper on queuing theory published in 1909. He pondered the problems of determining the number of telephone circuits necessary to provide phone services that would prevent customers from waiting too long for an available circuit. In developing a solution to this problem he began to realize that the problem of minimizing waiting time was applicable to many fields and began developing the theory further. Therefore, A.K. Erlang is considered the father of queuing theory.

The queuing theory is one of the most celebrated problems of operation research which has attracted the attention of researchers, scientists, mathematics and social scientists. A lot of research work has been dedicated to the application of this theory in health care systems, construction industries, human resource management, transportation, traffic and many other such systems.

A lot of contribution and application of queuing theory in the field of health care are found in the literature. In an era of health care reform, queuing theory is applied in improving quality, safety and decreasing health care cost. Whenever it is used appropriately, the results are often remarkable in saving time, increasing revenue and increasing staff and patient satisfaction. Therefore, applying queuing theory to health care sector is a necessary step towards improving quality of care and enhancement of the systems. Queuing theory has been studied in the health care settings. A considerable body of the research has shown the use of queuing theory in the real world health care situation. Research on models for evaluating the impact of bed assignments policies on utilization, wailing time and the probability of turning away patients. Review the use of queuing theory in pharmacy application with particular attention to improving customer's satisfaction. Customer satisfaction is improved by predicting and reducing waiting times and adjusting staffing. Preater presents a brief history of the use of queuing theory in health care. Green applied the queuing theory in health care. She discusses the relationship among delays, utilization and the number of servers; the basic M/M/s model, its assumptions and extensions; and the application of the queuing theory to determine the required number of servers. System design, appointment systems, out-patient appointment systems, the emergency cardiac in-patient flow and others. They successfully established the applicability of queuing theory in the field of health care.

Thus, queuing theory is a mathematical approach to study of the waiting lines. Long waiting time in any health care centre affects the improvement of the centre as well as the nation's economy. Therefore, to reduce the waiting time of arriving patients is a major challenge for services not only in India but all over the world especially in developing countries. While considering improvement in services, centre must measures the cost of providing a given level of service against the potential costs from having patient wait. Queuing theory has increasingly become a universal tool of management for decision making in a local health care clinic.

Description of the Models

Single Channel Queuing System

Consider a single server queuing system, (M/M/1) in which arriving customers is following Poisson's process with the arrival rate λ and the service process is following the exponential distribution with the service rate μ. Here customers are identified as arriving patients. The services in all phases are independent and identical and only one patient at a time is in the service mechanism.

When a patient enters the system and at a time if the system is free, his/her service time starts at once and when the system is not free, the patient joins the queue and wait for their turn/number for service. After completion of services, the patient is free from queue if there is not any further extended service facility. If the server is busy then the arriving patients goes to orbit and becomes of repeated calls. This pool of source of repeated calls may be viewed as a sort of queue. The time it takes to service every patient is an exponential random variable with parameter μ.

A pictorial representation of a single server queuing system, (M/M/1) is given below in which patients are standing in queue, waiting for the server to be free for providing service. In the situation of congestion of patients in a health care system, there is a very less probability for the patients arriving in end to get treatment as there is a single server rendering services. In worst situation, patients may leave the system without being serving.

Multichannel Queuing System

The multichannel queuing model is known in the Kendall's notation as the M/M/m model, where M signifies a Poisson distribution and m is the number of parallel service channel in the system. This is commonly used to analyze the queuing problem. This model commutes the average wait time and queue lengths, given arrival rate, number of servers and service rates. This particular model applies, in which there is multiple channel served by a single queue at a bank teller or many airline ticket counters. The output of the model is as follows:

1. Expected waiting time per patient in the system (health care).
2. Expected waiting time patients in the queue.
3. Expected number of patient in the system (health care).
4. Expected number of patient in the queue.

The exact calculation of these measures requires knowledge of the probability distribution of the arrival rate and service times. Moreover, successive inter-arrival times and service times are assumed to be statistically independent of each other. In this system, there are multiple servers with all sharing common waiting line a waiting line is crated when all the servers are busy in rendering service. As soon as one server becomes free, a customer (patient) is dispatched from the waiting line using dispatching discipline in force for being served. There are obvious from the pictorial representation of the multichannel queuing system which is given below.

Model Assumptions

This research is based on the following assumptions.

a) The finding obtained after investigate from one unit of the medical health centre should be valid in the other units.
b) The patients are almost well familiar with the organization system of the medical centre.
c) The arrival rate of the patients to queue and service rate are compatible to poison distribution or in other words the time interval between two consecutive arrivals and time services both follow exponential distribution.
d) The queuing theory discipline is such that the first patient goes to the server which is ready for service
e) In case of multichannel queuing system, it is assumed that none of the servers are unattended.

Parameters in Queuing Model

- n = Number of customers (units) in the system.
- c = number of parallel servers.
- λ = It is the mean rate of arrivals per unit of time in the system.
- μ = It is the average number of customers served per unit time in the system.
- cμ = Serving rate when c>1 in a system.
- ρ = utilization factor.
- P_0=Steady state probability of all idle servers in the system.
- P_n= Steady satiate probability exactly n patients in the system.
- L_q= Average number of patients in the queue.
- L_s= Expected number of patients in the system.
- W_q = Average waiting time a patients spends waiting in line excluding the service time.
- W_s = The expected time a customer spends in the system.

Data Analysis

The memory less property is utilized to define the state of the queuing system. To determine the performance measures, first find the probability of having n number of customers in the queuing system.

Probability of having 1 customer (i.e. n = 1) in the service system is:

$$P_1 = \rho\, P_0$$

Similarly,

$$P_2 = \rho\, P_1$$

$P_2 = \rho^2\, P_0$………..

$$P_n = \rho^n\, P_0$$

The probability value is 1. i.e.

$$\sum_{n=0}^{\infty} P_n = 1$$

$$P_0 + P_1 + P_2 + \cdots = 1$$

$$P_0 + \rho\, P_0 + \rho^2\, P_0 + \cdots = 1$$

$$(1 + \rho + \rho^2 + \cdots)P_0 = 1$$

Where $(1 + \rho + \rho^2 + \cdots)$is an infinite series,

Sum of infinite series can be written as,

$$(1|1 - \rho)$$

Hence,

$$P_0 = 1 - \rho$$

i.e., the probability of no patient in the system.

To determine performance measures L_s, L_q, W_S, W_q in the queuing system determine the average number of customers in the system.
The average number of customers L_s in the system can be written as,

$$L_s = \sum_{n=0}^{\infty} n \times P_n$$

Where $P_n = \rho^n P_0$ and $P_0 = 1 - \rho$

$$L_s = (1 - \rho) \sum_{n=0}^{\infty} n\rho^n$$

$$= \rho(1 - \rho) \sum_{n=0}^{\infty} n\rho^{n-1}$$

Then,

$$\frac{\partial \rho^n}{\partial \rho} = n\rho^{n-1}$$

$$L_s = \rho(1 - \rho) \frac{\partial}{\partial \rho} \sum_{n=0}^{\infty} \rho^n$$

$$= \rho(1 - \rho) \frac{\partial}{\partial \rho} (1|1 - \rho)$$

$$L_s = \frac{\lambda}{\mu - \lambda}$$

Using little's law according to which the average number of patients in the service system is the product of arrival rate and average time a customer's spends in the system
Average time a patient spends in the system, W_s can be written using little's law as given below

$$W_s = \frac{L_s}{\lambda}$$

Determine L_s and know λ, hence

$$W_s = \frac{\lambda}{\mu - \lambda} . \frac{1}{\lambda}$$

$$W_s = \frac{1}{\mu - \lambda}$$

Average time a customer spends in the queue W_q , can be determined by subtracting expected service time or average service time from average time a customer spends in the system W_s

$$W_q = W_s - \text{expected service time}$$
$$= W_s - \mu$$
$$= \frac{1}{\mu - \lambda} - \mu$$
$$W_q = \frac{\lambda}{\mu(\mu - \lambda)}$$

Using little's law to determine the average number of patients in the queue

$L_q = W_q \lambda$

$$L_q = \frac{\lambda^2}{\mu(\mu - \lambda)}$$

Model Parameters
Traffic Intensity
It is obtained from dividing the average arrival rate λ (in time) to the average service rate μ.

i.e., $\rho = \lambda/\mu$

Whenever λ is larger, the arrival of patients will increase and the system will work harder and queue will be longer. On the contrary, whenever λ is smaller, the queue will be shorter but in this case the use of system will be low. If the arrival rate of patients in the system were more than service rate. I.e. λ>μ then ρ>1, which means the system capacity is less than the arriving patients; therefore the queue length is increased. In this queuing system the average arrival rate is less than the average service rate i.e. λ >μ.

Average Waiting Time in Queue
The average waiting time in queue (before service is rendered) is equal to the average time which a patient waits in the queue for getting services. Its formula is,

$$\frac{\rho}{\mu(1 - \rho)} = \frac{\lambda}{\mu(\mu - \lambda)}$$

Average Time Spend in the System
The average time spent in a system (on queue and receiving service) is equal to the total time that a patient spends in a system which includes the waiting time and service time. Its formula is,

$$\frac{1}{\mu(1 - \rho)} = \frac{1}{\mu - \lambda}$$

Average Number of Patient in the System
The average number of patients in the system is equal to the average number of patients who are in the line or server. It is defined as

$$\frac{\rho}{1 - \rho} = \frac{\lambda}{\mu - \lambda}$$

Average Queue Length
The average queue length is composed of the average number of patients who are waiting in the queue. It is defined as

$$\frac{\lambda^2}{\mu(\mu - \lambda)}$$

The Probability of not Queuing on the Arrival
$= 1 - \rho$

Results and Discussions
Results of M/M/1
The arrival time as well as the time service began and ended for 125 patients in the local health care clinic. There are two types of services; consultancy and surgery. A total of 22 days were used for the data collection. On the basis of actual observed collected data, find the
Total waiting time of 125 patients for 22 days = 1584 minutes
Total service time of 125 patients for 22 days = 1320 minutes
Using the model parameters for the single channel queuing model, arrive the following results:
The arrival rate,

$$\lambda = \frac{\text{total number of patients}}{\text{Total waiting time}}$$

$= \frac{125}{1584} = 0.0789$ (1)

The service rate,

$$\mu = \frac{\text{total number of patients}}{\text{Total waiting time}}$$

$= \frac{125}{1320} = 0.09469$ (2)

The average waiting time in a queue,

$$= \frac{\lambda}{\mu(\mu - \lambda)}$$

$$= \frac{0.0789}{0.09469(0.09469 - 0.0789)}$$

$= 52.9530 \equiv 52$minutes (3)

The average time spent in a system,

$$\frac{1}{\mu - \lambda} = \frac{1}{0.09469 - 0.0789}$$

$= 63.3312 \equiv 63$minutes (4)

Average number of patient in the system,

$$\frac{\lambda}{\mu - \lambda} = \frac{0.0789}{0.09469 - 0.0789}$$

$= 4.9968 \equiv 5$ (5)

Average queue length (Average number of patient in the queue),

$$\frac{\lambda^2}{\mu(\mu - \lambda)} = \frac{\rho^2}{1 - \rho}$$

$= \frac{0.6942}{0.167} = 4.1570 \equiv 5$ (6)

The probability of queuing on arrival i.e. traffic,

$= \lambda/\mu$

$= \frac{0.0789}{0.09469} = 0.8332$ (7)

The probability of not queuing on the arrival,

$= 1 - \rho$

$= 1 - 0.8332$

$= 0.167$ (8)

Discussions

A single channel queuing system is used to represent the local health care clinic where doctor is treated as a single server and the mode parameters are applied for calculations. The study for this case is on the basis of actual observed data collection in 22 days of service for 125 patients.

The traffic intensity, $\rho = \lambda/\mu = 0.8332$ obtained in (7) shows the probability of patients queuing on arrival. This reveals the congestion of patients waiting for treatment as doctor is engaged in rendering service to the patients that has earlier been given appointment either for surgery or consultancy. This represents the inadequate service system of the clinic.

Also from the results (3) and (4) it is obvious that the average time spent in the clinic (in queue and in receiving treatment) is greater than the average time spent in the queue before providing treatment. Thus, there will always be a queue of patients in the clinic which is also very clear from the results (5) and (6). The result (8) shows that there is a very less possibility services to new arriving patients.

Multi-Channel Queuing Theory Model (M/M/c: FCFS/∞/∞)

Multi- channel queuing theory treats the condition in which there are several service stations in parallel and each element in the waiting line can be served by more than one station. Each service facility is prepared to deliver the same type of service. The new arrival selects one station without any external pressure. When a waiting line is formed, a single line usually breaks down into shorter lines in front of each service station. The arrival rate λ and service rate μ are mean values from poison distribution and exponential distribution respectively. Service discipline is first-come first served and customers are taken from a single queue i.e., any empty channel is filled by the next customer in line.

M/M/c Queuing Model

The first known values in a calculation of performance measure is,

(i) Traffic intensity(ρ)
(ii) Probability of the system should be idle (P_0)

The traffic intensity is,

$\rho = \lambda/c\mu$

Whenever λ is larger, the arrival of patients will increase and the system will work harder and queue will be longer. On the contrary, whenever λ is smaller, the queue will be shorter but in this case the use of system will be low. If the arrival rate of patients in the system were more than service rate. I.e. $\lambda > \mu$ then $\rho > 1$, which means the system capacity is less than the arriving patients; therefore the queue length is increased. In this queuing system the average arrival rate is less than the average service rate i.e. $\lambda > \mu$.

The probability that the system should be idle.

$$P_0 = \left[\sum_{n=0}^{c-1} \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n + \frac{1}{c!}\left(\frac{\lambda}{\mu}\right)^n \frac{c\lambda}{c\mu - \lambda}\right]^{-1}$$

The average number of patients in the system,

$$L_s = \frac{\lambda.\mu\left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\mu - \lambda)^2} P_0 + \frac{\lambda}{\mu}$$

The average number of patients waiting in the queue,

$$L_q = L_s - \text{average number being served}$$

$$= L_s - c\left(\frac{\lambda}{c\mu}\right)$$

$$L_q = \frac{\lambda.\mu\left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\mu - \lambda)^2} P_0$$

Average waiting time a customer spends in the system,

$$W_s = \frac{L_s}{\lambda}$$

$$W_s = \frac{\mu\left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\mu - \lambda)^2} P_0 + \frac{1}{\mu}$$

Average waiting time of a customer in the queue,

$$W_q = \frac{L_q}{\lambda}$$

$$W_q = \frac{\mu\left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\,(c\mu - \lambda)^2} P_0$$

Utilization factor, $\rho = \frac{\lambda}{c\mu}$

Average number of idle servers,
= c – (average number of customers served)
Probability that a customer has to wait,

$$p(n \geq c) = \frac{\mu\left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\ (c\mu - \lambda)} P_0$$

Probability that a customer enters the service without waiting,

$$1 - p(n \geq c) = 1 - \frac{\mu\left(\frac{\lambda}{\mu}\right)^c}{(c-1)!\ (c\mu - \lambda)} P_0$$

Results of M/M/3
The arrival time as well as the time service began and ended for 125 patients in the local health care clinic "Venkatesh Nursing Home". There are two types of services; consultancy and surgery. A total of 22 days were used for the data collection. On the basis of actual observed collected data, find the
Total waiting time of 125 patients for 22 days = 1584 minutes (9)
Total service time of 125 patients for 22 days = 1320 minutes (10)
Consider 3 servers. That is 3 doctors. one doctor is for surgery, 2^{nd} one for consultancy and 3^{rd} one for emergency patient.
Using the model parameters for the single channel queuing model, the following results:
The arrival rate,

$$\boldsymbol{\lambda} = \frac{\text{total number of patients}}{\text{Total waiting time}}$$

$= \frac{\mathbf{125}}{\mathbf{15}84}$ =0.0789 (11)

The service rate,

$$\mu = \frac{\text{total number of patients}}{\text{Total waiting time}}$$

$= \frac{125}{1320}$=0.09469 (12)

The probability that the system should be idle,

$$P_0 = \left[\sum_{n=0}^{2} \frac{1}{n!}\left(\frac{0.0789}{0.09469}\right)^n + \frac{1}{3!}\left(\frac{0.0789}{0.09469}\right)^3 \frac{3(0.09469)}{3(0.09469) - 0.0789}\right]^{-1}$$

=0.4322 (13)

The average number of patients in the system,

$$L_s = \frac{(0.0789).(0.09469)\left(\frac{0.0789}{0.09469}\right)^3}{(3-1)!\,(3(0.09469) - 0.0789)^2}(0.4322) + \frac{0.0789}{0.09469}$$

=0.8484≡1 (14)

The average number of patients waiting in the queue,

$$L_q = \frac{(0.0789).(0.09469)\left(\frac{0.0789}{0.09469}\right)^3}{(3-1)!\,(3(0.09469) - 0.0789)^2}(0.4322)$$

= 0.0152
≡1 (15)

The average waiting time of patients in the system,

$$W_s = \frac{(0.09469)\left(\frac{0.0789}{0.09469}\right)^3}{(3-1)!\,(3(0.09469) - 0.0789)^2}(0.4322) + \frac{1}{0.09469}$$

= 11.5589
≡12minute (16)

The average waiting time of patients in the queue,

$$L_q \frac{(0.09469)\left(\frac{0.0789}{0.09469}\right)^3}{(3-1)!\,(3(0.09469) - 0.0789)^2}(0.4322)$$

= 10.752

≡11minute (17)

The traffic intensity is,

$$\rho = \frac{0.0789}{3(0.09469)}$$

=0.2777. \ (18)

The probability of not queuing on the arrival,
= 1 – ρ
= 1 – 0.2777
=0.7223. (19)

Comparing the Results of M/M/1 and M/M/3

Comparison	M/M/1	M/M/3
Traffic intensity	0.8332	0.2777
average number of patients in the system	5	1
average number of patients in the queue	5	1
average waiting time of patients in the system	63 minutes	12 minutes
average waiting time of patients in the queue	52 minutes	11 minutes
The probability of new patient Arrival	0.167	0.7223

The waiting time in M/M/3 is less than the M/M/1. The average number of patients in the queue is also less. The probability of new patient's arrival is large in the multi server.

Average Waiting Time in a System of Single Server and Multi Server by Using MATLAB

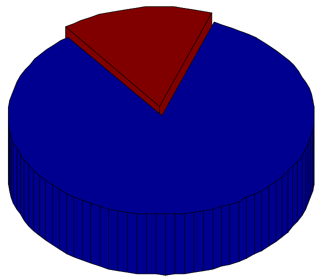

Fig 1

Average Waiting Time in a Queue of Single Server and Multi Server by Using MATLAB

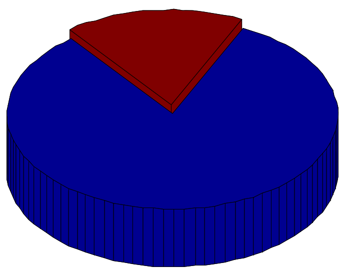

Fig 2

New Patient's Arrivals Probability of Single Server and Multi Server by Using MATLAB

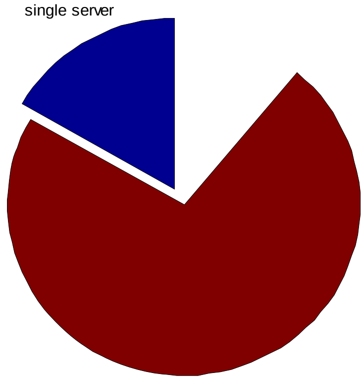

Fig 3

Conclusion

In this paper, multi-server queuing system was discussed. In this case the waiting time in a queuing model (M/M/1) was more than that of queuing model (M/M/3) and observed that new patient's arrival probability of single server queuing system was less than that of multi-server queuing system. The multi-server queuing system increases the efficiency of the hospital and reducing time compared to the single server queuing system.

Reference

1. Adeleke RA. Application of Queuing Theory to Waiting Time of Out-Patients in Hospitals, Pacific Journal of Science and Technology, 2009; 10(2):270-274.
2. Agnihotri SRA, Taylor PF. Staffing a centralized appointment scheduling department in Lourdes Hospital, Interfaces-Journal Online, 1991; 21:1-11.
3. Bailey NTJ. A study of queues and appointment systems in hospital out-patient departments with special reference to waiting times, Journal of Royal Statistical Society, Netherlands, 1952; 14:185-199.
4. Erlang AK. The Theory of Probabilities and Telephone Conversations, NytTidsskrift for Matematik, 1909; 20(B):33-39.
5. Fomundam S, Herrmann JW. A Survey of Queuing Theory Applications in Healthcare, ISR Technical Report 2007, 24.
6. Green LV. Queuing analysis in healthcare, inpatient flow: Reducing delay in healthcare delivery, (Hall, R.W., Springer, New York, 2006a, 281-308.
7. Khan MR, Callahan BB. Planning laboratory staffing with a queuing model, European Journal of Operational Research, 1993; 67(3):321-331.
8. Jagan T. EOQ for deteriorating item with non-instantaneous receipt under trade credits with shortages", International Journal of Scientific Development and Research, 2(5), 2455-2631.
9. McClain JO. Bed planning using queuing theory models of hospital occupancy: a sensitive analysis, Inquiry, 1976; 13:167-176.
10. Nose RA, Wilson JP. Queuing theory and customer satisfaction: A review of terminology, trends and applications to a pharmacy practice, Hospital Pharmacy, 2001; 3:275-279.
11. Premkumar K. Operations Research", 2002; 5:283.