Analysis the Customers' Patience in Restaurant

Ank Phyu Win
Faculty of Computing
University of Computer studies (Sittway)
ankphyuwin@ucssittway.edu.mm

DOI: 10.29322/IJSRP.9.08.2019.p9240
http://dx.doi.org/10.29322/IJSRP.9.08.2019.p9240

Abstract- Queues (waiting lines) are a part of everyday life. However, having to wait is not just a petty personal annoyance. A queuing model is constructed so that queue lengths and waiting time can be predicted. Queuing models represent the queuing systems. This model enables the performance measures and customers' patience in the restaurant. The common problem arises in almost every famous restaurant is that they lose the customers because of a long waiting line and poor quality service. This paper analyze the performance measures and the customers' patience in the system. The data sets are obtained from a Kentucky Fried Chicken shop in Yangon in order to derive the arrival rate, the service rate, utilization factor, waiting time and probability of potential balking customers. The collected data is analyzed by Little's formula and queuing model which is M/M/1 model.

Index Terms- Kendall's Notation, Little's Formula, Queuing Models, Queuing Theory

I.INTRODUCTION

To improve the customers' return, the managers in the restaurant must consider a long waiting line and poor quality services. Although the taste, cleanliness, the restaurant layout and view are the most important factors for the restaurant, other factors such as customers' patience, long waiting line must be considered by managers. These factors will reduce customers' demands and even profit. In mathematics, a random service theory is called the queuing theory to solve the problems and analyze the system. By using queuing theory, the average waiting time in queue, average waiting time in the system, the average queue length, the average number of customers in queue and average number of customers in the system, the busy period and idle time can be calculated To model the restaurant operation, queuing theory have been previously used by many researchers. This study is to measure the performance in the system and analyze the customers' patience on the administrative issues.

II. METHODOLOGY

A. The Conceptual Framework: Queuing Theory

In queuing theory, a discipline within the mathematical theory of probability, an M/M/1 queue represents the queue length in a system having a single server, whose arrivals are determined by a Poisson process and job service times have an exponential distribution. The model name is written in Kendall's notation. An M/M/1 queue is a stochastic process whose state space is the set { 0,1,2,….,n} where the values corresponds to the number of customers in the system, including any currently in service.[1] To make the business decisions providing a service, queuing theory is considered. Queuing theory is based on probability theory and random process. From the fundamental features of the queuing theory, the customer's arrival pattern, the service time of server may be one of the probability distributions. The M/M/1 system is made of a Poisson arrival, one exponential (Poisson) servers, FIFO (or not specified) queue of unlimited capacity and unlimited customer population. The M/M/1 model show clearly the basic ideas and methods of Queuing Theory.[2]

B. Basic Structure of Queuing Models

The basic process assumed by most queuing models is the following: Customers requiring service are generated over time by an input source. At certain times, a member of the queue is selected for service by some rule known as the queue discipline. The required service is then performed for the customer by service mechanism. This process is depicted in Figure 1.[3]

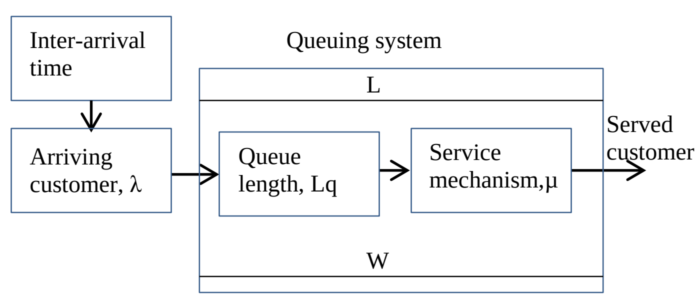

Figure 1. M/M/1 Queuing model

The input source is the total number of customers that might require service from time to time, i.e., the total number of distinct potential customers. The input source is said to be either unlimited or limited.

The queue is where customers wait before being served. The queue length is the maximum permissible number of customers that it can contain.

The queue discipline indicates the manner in which the units are taken for service. The usual queue discipline is first come first served, or FCFS through sometimes there are other service discipline, such as last come, first served or service in random order.[4]

The service mechanism consists of one or more service facilities, each of which contains one or more parallel service channels, called servers.

The time elapsed from the commencement of service to its completion for a customer at a service facility is the service time.

A model of a particular queuing system must specify the probability distribution of service times for each server. The service time distribution that is most frequently assumed in practice is the exponential distribution.

C. The Role of the Exponential Distribution

The operating characteristics of queuing systems are determined largely by two statistical properties, namely, the probability distribution of inter-arrival times and the probability distribution of service times. For real queuing systems, these distributions can take on almost any form. The most important probability distribution in queuing theory is the exponential distribution. The cumulative probabilities are

$$P\{T \ll t\} = 1 - e^{-\alpha t} \quad (1)$$
$$P\{T > t\} = e^{-\alpha t} \quad (2)$$

D. Little's Law

Little's law is a theorem that determines the average number of customers in a business system based on the average waiting time of a customer within a system and the average number of customers arriving at the system per unit time. [5] Mathematically, Little's law is expressed through the following equation:

$$L = \lambda W \quad (3)$$

Although it looks intuitively easy, it is quite a remarkable result as the relationship is not influenced by the arrival process distribution, the service distribution, the service order or practically anything else.

E. Queuing Models and Kendall's Notation

In queuing models, the following factors are characterized:[3]

Arrival time distribution: Inter -arrival times independently and identically distributed according to an exponential distributions (i.e. the input process is Poisson)

Service time distribution: The service time distribution can be exponential distribution. The service time is independent of the inter-arrival time.

Service mechanism: It consists of one or more service facilities, each of which contains one or more parallel service channels, called servers.

Queue lengths: There may be infinite or finite queue length .

Queuing discipline: The queue discipline refers to the order in which members of the queue are selected for service. It may be first-come-first-served, random, according to some priority procedure, or some other order. First-come-first-served usually is assumed by queuing models.

Capacity: The maximum number of customers in a system can be several.

Kendall's notation is a system of notation according to which the various characteristics of a queuing model are identified. Kendall (Kendall, 1951) has introduced a set of notations which have become standard in the literature of queuing models. A general queuing system is denoted by (a/b/c):(d/e) [6] where

a= probability distribution of the inter-arrival time
b=probability distribution of the service time
c=number of servers in the system
d=maximum number of customers allowed in the system
e=queue discipline

III. PROCEDURAL OVERVIEW

From observation, this shop has several waiters but only one chef can serve all of the customers.

A. Procedure with Data

In M/M/1 queuing model, the following formulas will be used[3].

$$\rho = \frac{\lambda}{\mu} \quad (4)$$
$$P_0 = 1 - \rho \quad (5)$$
$$P_n = (1-\rho)\rho^n \quad (6)$$
$$L = \frac{\lambda}{\mu - \lambda} \quad (7)$$
$$L_q = \frac{\rho^2}{1-\rho} \quad (8)$$
$$W = \frac{1}{\mu - \lambda} \quad (9)$$
$$W_q = \frac{L_q}{\lambda} \quad (10)$$

λ = expected number of arrivals per unit time

μ =expected number of customers completing service per unit time

ρ =utilization factor

P_0 = probability of no customer in the system

P_n = probability of exactly n customers in queuing system

L = average number of customers in the system

L_q = average number of customers in queue

W = average waiting time spent in the system

W_q =average waiting time in queue

B. Calculation

From observations, the KFC shop provides 25 tables and 98 chairs in shop 1 and 17 tables and 72 chairs in shop 2. If necessary, the extra chairs are provided. There are several waiters to serve. In this shop, between 400 and 600 customers can be served during weekdays and between 500 and 1000 customers can be served during weekends in shop 1 and between 300 and 500 during weekdays and between 500 and 700 during weekends in shop 2. This paper seeks to analyze the performance measures and customers' patience during weekdays and weekends respectively. The time window is 3 hours.

In shop 1, at the off-peak hours, there are an average 500 people coming to this shop in 3 hours. From this, the arrival rate can be derived as: $\lambda = \frac{500}{180} = 2.78$ customers per minute. From observations, each customer spends 30 minutes an average in it (W), the waiting time (W_q) is about 3 minutes in peak hours.

Next, the average customers in this shop is $L = \lambda W = 2.78 \times 30 = 83.34$ customers. It is Little's formula. There is no customer in queue because the system capacity is greater than the average number of customers. There is also no waiting time in queue in off-peak hours. The service rate can be derived as:

$$\mu = \frac{\lambda(1+L)}{L} = \frac{2.78(1+83.34)}{83.34} = 2.81 \textit{ customers per minute}$$

It means that nearly 3 customers can be served in one minute. The utilization factor or busy period can be calculated as:

$$\rho = \frac{2.78}{2.81} = 0.989$$

It means that the chef is very busy. And the probability that the system is idle can be calculated:

$$P_0 = 1 - 0.989 = 0.01$$

It means that the system may be free in 1%.The probability of exactly 'n' customer in the system:

$$P_n = (1-\rho)\rho^n$$

In off-peak hours, there is no customer in queue to be served. So, there is no customer to balk in this situation.

In peak-hours, the measures of performance can be calculated as follows: there are on average 750 people coming to this shop in 3 hours. From this, the arrival rate is 4.167 customers per minute. As above, W=30 minutes and W_q=3 minutes, so, L=125 customers. In this situation, the capacity is just 98 but the number of customer is 125 people. So, the number of customer is greater than the capacity. So, there is also need to know the average number of customers in queue. L_q= 12.5 customers are in queue. The service rate for this situation is: μ=4.2 customers per minute.

It means that 4.2 customers can be served in one minute. The utilization factor is 0.992 and the idle time is 0.008.

In peak-hours, potential customers will start to balk because of a long waiting line. When the customers see more than 10 people are already in queue and a potential can tolerate is 15 people, the probability of customers going away is:

$$P_{125-140} = \sum_{n=125}^{140} (0.008)\,(0.992)^n = 4.156 \quad \%$$

In shop 2, at the off-peak hours, there are an average 400 people coming to this shop in 3 hours. From this, the arrival rate can be derived as: $\lambda = 2.22$ customers per minute. From observations, each customer spends 30 minutes an average in it (W), the waiting time (W_q) is about 3 minutes in peak hours.

Next, the average number of customers in this shop is 67 customers. There is no customer in queue because the capacity of this shop is greater than the average number of customers in it. There is also no waiting time in queue in off-peak hours. The service rate is 2.25 customers per minute. It means that nearly 3 customers can be served in one minute. The utilization factor or busy period is 0.987. And the probability that the system idle is 0.013. It means that the system may be free in 1.3%.

In peak-hours, the measures of performance can be calculated as follows: there are on average 600 people coming to this shop in 3 hours. From this, the arrival rate is 3.33 customers per minute. As above, W=30 minutes and W_q=3 minutes, so, there are 100 customers in the system. In this situation, the capacity is just 72 but the number of customer is 100 people. So, the number of customer is greater than the capacity. The average number of customers in queue is L_q= 10 customers is in queue. The service rate for this situation is 3.36 customers per minute. The utilization factor is 0.99 and the idle time is 0.01.

In peak-hours, the potential customers who will start to balk because of a long waiting line. When they see more than 10 people are already in queue and a potential can tolerate is 15 people, the probability of customers going away is:

$$P_{100-125} = \sum_{n=100}^{125} (0.01)\,(0.99)^n = 8.42 \quad \%$$

IV. RESULTS

The utilization rate is very high at 0.989 and 0.987 at weekdays and 0.992 and 0.99 at weekends. But the service rate is at 2.81 and 2.25 customers per minute on weekdays and at 4.2 and 3.36 customers per minute on weekends each shop. But the quality of service is needed to consider. Because the more the customer, the less the service of quality. It is main point. The faster the service, the poorer the taste. But, when the service rate is higher, the utilization rate will be lower. It can decrease the probability of the customers going away.

This paper can help to increase the quality of service by anticipating if there are many customers in the queue. By analyzing the huge number of customers coming and going in a day, the manager can determine how to decide to get profit. The summarized result is shown in Table 1 and Table 2 respectively.

Table 1. The Summarized Results in Shop 1

Symbols	On Weekdays	On Weekends
λ	2.78	4.167
μ	2.81	4.2
ρ	0.989	0.992
P_0	0.01	0.008
L	83.34	125
L_q	0	12.5
W	30	30

W_q	0	3
$P_{125-140}$	0	4.156%

Table 2. The Summarized Results in Shop 2

Symbols	On Weekdays	On Weekends
λ	2.22	3.33
μ	2.25	3.36
ρ	0.987	0.99
P_0	0.013	0.01
L	67	100
L_q	0	10
W	30	30
W_q	0	3
$P_{125-140}$	0	8.42%

V. CONCLUSIONS

The overall analysis of the results can be summarized as follows: On weekdays, there are nearly 83 customers and 67 customers in this system but no customer in queue. Nearly 3 customers and 2 customers can be served in one minute each shop. So, the server is very busy in this situation. On weekends, there are 125 customers and 100 customers each shop and nearly 13 customers and 10 customers in queue to be served in peak hours. 4 customers can be served in one minute at shop 1 and 3 customers can be served in one minute at shop 2. The service rates are greater than before. So, the quality of service may be poor. The chef is very busy too. In shop 1, 4% of customers will start to balk and nearly 8% of the customers will balk at shop 2. And the fact that the potential customers who will start to balk must be considered. This theory can also be applied for the restaurant if they want to calculate all the data daily. By analyzing the several number of customers in the system to serve and the busy period, this shop can set a target how to manage.

ACKNOWLEDGEMENT

This paper is supported by U Kyaw Myo Zaw, the manager of KFC shop in Yangon, Myanmar and by Daw Mi Cho Cho, Head of Faculty of Computing in University of Computer Studies (Sittway).I'd like to thank them for their constructive data and supporting.

REFERENCES

[1] https://en.m.wikipedia.org/wiki/M/M/1-queue
[2] http://staff.um.edu.mt/jskl1/Simucb/mm1.htm
[3] Frederick S. Hiller, Introduction to Operations Research, Seventh Edition (Stanford University), January, 2000.
[4] J.MEDHI, in Stochastic Models in Queuing Theory (Second Edition), 2003.
[5]https://corporatefinanceinstitute.com/resources/knowledge/other/Little's- Law
[6] https://www.oreilly.com / Library /view /quantitative-techniques-theory /9789332512085 /xhtml /ch9sec10.xhtml.