

MINIMIZING THE WAITING TIME FOR SERVICE WITH QUEUING MODEL

Dr. Prashant Makwana¹ and Gopalkrushna Patel²

¹Director, SLPT MBA Mahila College, Amreli, Gujarat

Email: prmak1@hotmail.com

²HOD, SLPT MBA Mahila College, Amreli, Gujarat

Email: gopalmeb@gmail.com

ABSTRACT

*Today no body interested to wait for food in restaurant. Some restaurants initially provide more waiting chairs for customers However, waiting chairs alone would not solve a problem, and the service time may need to be improved. This shows a need of a numerical model for the restaurant management to understand the situation better. This paper aims that, using queuing theory satisfies the model when tested with a real-case scenario. We obtained the data from a restaurant in Ahmedabad. We then derive the arrival rate, service rate, utilization rate, waiting time in queue and the probability of potential customers to balk based on the data using Little's Theorem and M/M/1 queuing model. The arrival rate at **Born Vivant** during its busiest period of the day is 2 customers per minute (cpm) while the service rate is 2.05cpm. The average number of customers in the restaurant is 90 and the utilization period is 0.976. We conclude the paper by discussing the benefits of performing queuing analysis to a busy restaurant*

Keywords: Queue, Little's Theorem, Restaurant, Waiting Lines.

INTRODUCTION

There are several determining factors for a restaurant to be considered a good or a bad one. Taste, cleanliness, the restaurant layout and settings are some of the most important factors. These factors, when managed carefully, will be able to attract plenty of customers. However, there is also main factor that needs to be considered especially when the restaurant has already succeeded in attracting customers. This factor is the **customers queuing time**.

Queuing theory is the study of queue or waiting lines. Some of the analysis that can be derived using queuing theory include the expected waiting time in the queue, the average time in the system, the expected queue length, the expected number of customers served at one time, the probability of balking customers, as well as the probability of the system to be in certain states, such as empty or full.

Waiting lines are a common sight in restaurants especially during lunch and dinner time. Hence, queuing theory is suitable to be applied in a restaurant setting since it has an associated queue or waiting line where customers who cannot be served immediately have to

queue (wait) for service. Researchers have previously used queuing theory to model the restaurant operation [2], reduce cycle time in a busy fast food restaurant [3], as well as to increase throughput and efficiency [5].

This paper uses queuing theory to study the waiting lines in **Born Vivant** Restaurant at Ahmedabad city, Gujarat. The restaurant provides 20 tables of 5 people. There are 8 to 9 waiters or waitresses working at any one time. On a daily basis, it serves over 400 customers during weekdays and over 1000 customers during weekends. This paper seeks to illustrate the usefulness of applying queuing theory in a real case situation.

Queuing Theory

In 1908, Copenhagen Telephone Company requested Agner K. Erlang to work on the holding times in a telephone switch. He identified that the number of telephone conversations and telephone holding time fit into Poisson distribution and exponentially distributed. This was the beginning of the study of queuing theory. In this section, we will discuss two common concepts in queuing theory.

Little's Theorem

Little's theorem [7] describes the relationship between through put rate (i.e. arrival and service rate), cycle time and work in process (i.e. number of customers/jobs in the system). This relationship has been shown to be valid for a wide class of queuing models. The theorem states that the expected number of customers (N) for a system in steady state can be determined using the following equation:

$$L = \lambda T \quad --1$$

Here, λ is the average customer arrival rate and T is the average service time for a customer. Consider the example of a restaurant where the customer's arrival rate (λ) doubles but the customers still spend the same amount of time in the restaurant (T). These facts will double the number of customers in the restaurant (L). By the same logic, if the customer arrival rate (λ) remains the same but the customers service time doubles this will also double the total number of customers in the restaurant. This indicates that in order to control the three variables, managerial decisions are only required for any two of the three variables.

Three fundamental relationships can be derived from Little's theorem [6]:

- L increases if λ or T increases
- λ increases if L increases or T decreases
- T increases if L increases or λ decreases

Rust [8] said that the Little's theorem can be useful in quantifying the maximum achievable operational improvements and also to estimate the performance change when the system is modified.

Queuing Models and Kendall's Notation

In most cases, queuing models can be characterized by the following factors:

- **Arrival time distribution.** Inter-arrival times most commonly fall into one of the following distribution patterns: a Poisson distribution, a Deterministic distribution,

or a General distribution. However, inter-arrival times are most often assumed to be independent and memory less, which is the attributes of a Poisson distribution.

- **Service time distribution.** The service time distribution can be constant, exponential, hyper exponential, hypo-exponential or general. The service time is independent of the inter-arrival time.
- **Number of servers.** The queuing calculations change depends on whether there is a single server or multiple servers for the queue. A single server queue has one server for the queue. This is the situation normally found in a grocery store where there is a line for each cashier. A multiple server queue corresponds to the situation in a bank in which a single line waits for the first of several tellers to become available.
- **Queue Lengths (optional).** The queue in a system can be modeled as having infinite or finite queue length.
- **System capacity (optional).** The maximum number of customers in a system can be from 1 up to infinity. This includes the customers waiting in the queue.
- **Queuing discipline (optional).** There are several possibilities in terms of the sequence of customers to be served such as FIFO (First In First Out, i.e. in order of arrival), random order, LIFO (Last In First Out, i.e. the last one to come will be the first to be served), or priorities.

Kendall, in 1953, proposed a notation system to represent the six characteristics discussed above. The notation of a queue is written as :A/B/P/Q/R/Z where A, B, P, Q, R and Z describe the queuing system properties.

- A describes the distribution type of the inter arrival times.
- B describes the distribution type of the service times.
- P describes the number of servers in the system.
- Q (optional) describes the maximum length of the queue.
- R (optional) describes the size of the system population.
- Z (optional) describes the queuing discipline.

SUSHI TEI QUEUING Model

The data were obtained from **Born Vivant** through interview with the restaurant manager as well as data collections through observations at the restaurant.

The daily number of visitors was obtained from the restaurant itself. The restaurant has been recording the data as part of its end of day routine. We also interviewed the restaurant manager to find out about the capacity of the restaurant, the number of waiters and waitresses, as well as the number of chefs in the restaurant. Based on the interview with the restaurant manager, we concluded that the queuing model that best illustrate the operation of **Born Vivant** is M/M/1.

This means that the arrival and service time are exponentially distributed (Poisson process). The restaurant system consists of only one server. In our observation the restaurant has

several waitresses but in the actual waiting queue, they only have one chef to serve all of the customers.

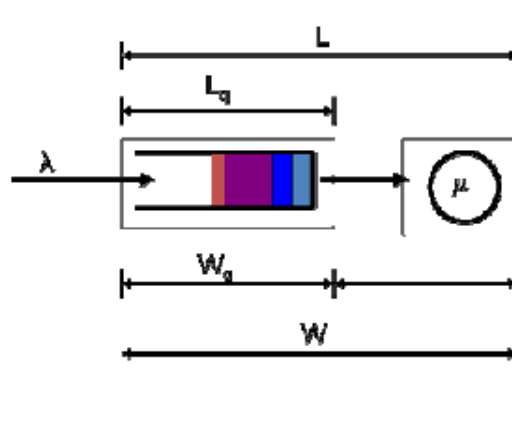


Figure 1. Illustrates the M/M/1 queuing model

For the analysis of the **Born Vivant** M/M/1 queuing model, the following variables will be investigated [6]:

- λ : The mean customers arrival rate
- μ : The mean service rate
- $\rho = \lambda/\mu$: utilization factor
- Probability of zero customers in the restaurant:

$$P_0 = 1 - \rho \quad (2)$$

- P_n : The probability of having n customers in the restaurant.

$$P_n = P_0 \rho^n = (1 - \rho) \rho^n \quad (3)$$

- L : average number of customers dining in the restaurant.

$$L = \frac{\rho}{1 - \rho} = \frac{\lambda}{\mu - \lambda} \quad (4)$$

- L_q : average number in the queue.

$$L_q = L \times \rho = \frac{\rho^2}{1 - \rho} = \frac{\rho \lambda}{\mu - \lambda} \quad (5)$$

- W : average time spent in Sushi Tei, including the waiting time.

$$W = \frac{L}{\lambda} = \frac{1}{\mu - \lambda} \quad (6)$$

- W_q : average waiting time in the queue.

$$W_q = \frac{L_q}{\lambda} = \frac{\rho}{\mu - \lambda} \quad (7)$$

RESULT AND DISCUSSION

The one month daily customer data were shared by the restaurant manager as shown in following table.

Mon	Tue	Wed	Thu	Fri	Sat	Sun
200	220	250	190	225	300	509
310	335	340	344	326	425	564
421	520	500	480	524	600	870
450	570	590	600	697	900	1050

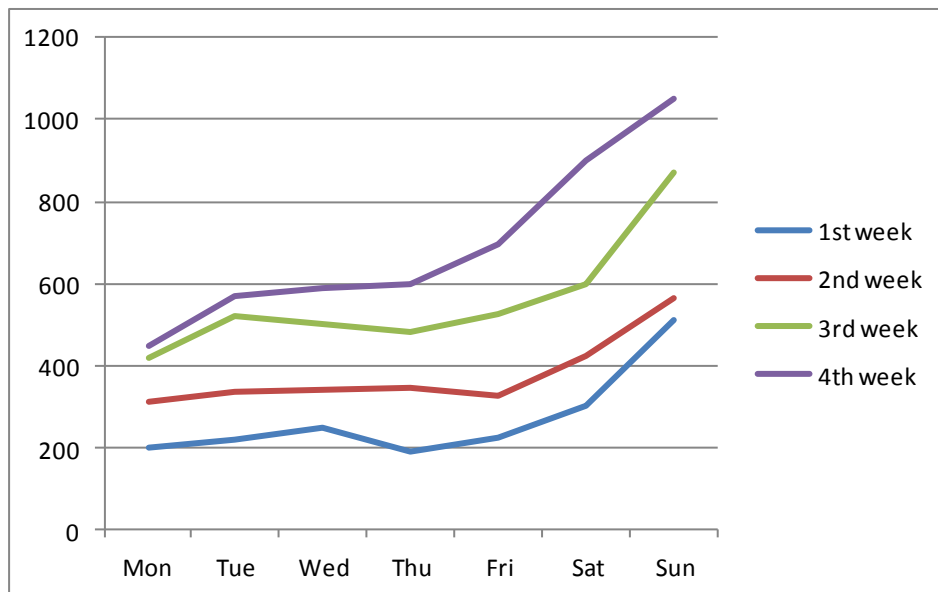


Figure 2. One month daily customer counts

As can be seen in Figure 2, the number of customers on Saturdays and Sundays are double the number of customers during weekdays. The busiest period for the restaurant is on weekend during dinner time. Hence, we will focus our analysis in this time window.

Calculation

Our teams conducted the research at dinner time. There are on average 400 people are coming to the restaurant in 3hours' time window of dinner time. From this we can derive the arrival rate as:

$$\lambda = 400/180 = 2.22 \text{ customer/minute (cpm)}$$

We also found out from observation and discussion with manager that each customer spends 55 minutes on average in the restaurant (W), the queue length is around 36 people (Lq) on average and the waiting time is around 15 minutes.

It can be shown using (7) that the observed actual waiting time does not differ by much when compared to the theoretical waiting time as shown below.

$$W_q = 36 \text{ customer} / 2.22 \text{ cpm} = 16.22 \text{ minutes}$$

Next, we will calculate the average number of people in the restaurant using (1).

$$L = 2.22 \text{ cpm} \times 55 \text{ min} = 122.1 \text{ customers}$$

Having calculated the average number of customers in the restaurant, we can also derive the utilization rate and the service rate using (4).

$$\mu = \frac{\lambda(1+L)}{L} = \frac{2.22(1+122.1)}{122.1} = 2.24 \text{ cpm}$$

With the very high utilization rate of 0.991 during dinnertime, the probability of zero customers in the restaurant is very small as can be derived using (2).

$$P_0 = 1 - \rho = 1 - 0.976 = 0.024$$

The generic formula that can be used to calculate the probability of having n customers in the restaurant is as follows:

$$P_n = P_0 \rho^n = (1 - \rho) \rho^n = (1 - 0.976) 0.976^n = 0.024 (0.976)^n$$

We assume that potential customers will start to balk when they see more than 10 people are already queuing for the restaurant. We also assume that the maximum queue length that a potential customer can tolerate is 40 people. As the capacity of the restaurant when fully occupied is 120 people, we can calculate the probability of 10 people in the queue as the probability when there are 130 people in the system (i.e. 120 in the restaurant and 10 or more queuing) as follows:

Probability of customers going away = $P(\text{more than 15 people in the queue}) = P(\text{more than 130 people in the restaurant})$

$$P_{131-160} = \sum_{n=131}^{160} P_n = 0.02147 = 2.14 \%$$

Evaluation

- The utilization is directly proportional with the mean number of customers. It means that the mean number of customers will increase as the utilization increases.
- The utilization rate at the restaurant is very high at 0.991. This, however, is only the utilization rate during lunch and dinner time on Saturdays and Sundays. On weekday, the utilization rate is almost half of it. This is because the number of visitors on weekdays is only half of the number of visitors on weekends. In addition,

the number of waiters or waitresses remains the same regardless whether it is peak hours or off-peak hours.

- In case the customers waiting time is lower or in other words we waited for less than 15 minutes, the number of customers that are able to be served per minute will increase. When the service rate is higher the utilization will be lower, which makes the probability of the customers going away decreases.

Benefits

- This research can help to Management of **Born Vivant** to increase their profit by QOS (Quality of Service), by anticipating if there are many customers in the queue.
- The result of this paper work may become the reference to analyze the current system and improve the next system. Because the restaurant can now estimate of how many customers will wait in the queue and the number of customers that will go away each day.
- By anticipating the huge number of customers coming and going in a day, the restaurant can set a target profit that should be achieved daily.
- The formulas that were used during the completion of the research are applicable for future research and also could be used to develop more complex theories.
- The formulas provide mechanism to model the restaurant queue that is simpler than the creation of simulation model in [9,4].

CONCLUSION

This research paper has discussed the application of queuing theory of **Born Vivant** Restaurant. Here we have focused on two particularly common decision variables (as a vehicle for introducing and illustrating all the concepts. From the result we have obtained that the rate at which customers arrive in the queuing system is 2.22 customers per minute and the service rate is 2.24 customers per minute. The probability of buffer flow if there are 10 or more customers in the queue is 15 out of 100 potential customers. The probability of buffer overflow is the probability that customers will run away, because may be they are impatient to wait in the queue. This theory is also applicable for the restaurant if they want to calculate all the data daily. It can be concluded that the arrival rate will be lesser and the service rate will be greater if it is on weekdays since the average number of customers is less as compared to those on weekends. The constraints that were faced for the completion of this research were the inaccuracy of result since some of the data that we use was just based on assumption or approximation. We hope that this research can contribute to the betterment of **Born Vivant** restaurant in terms of its way of dealing with customers.

As our future works, we will develop a simulation model for the restaurant. By developing a simulation model we will be able to confirm the results of the analytical model that we develop in this paper. In addition, a simulation model allows us to add more complexity so that the model can mirror the actual operation of the restaurant more closely [1].

REFERENCES

1. T. Altiok and B. Melamed, *Simulation Modeling and Analysis with ARENA*. ISBN 0-12-70523-1. Academic Press, 2007
2. D.M. Brann and B.C. Kulick, "Simulation of restaurant operations using the Restaurant Modeling Studio," Proceedings of the 2002 Winter Simulation Conference, IEEE Press, Dec. 2002, pp. 1448-1453
3. S. A. Curin, J. S. Vosko, E. W. Chan, and O. Tsimhoni, "Reducing Service Time at a Busy Fast Food Restaurant on Campus," Proceedings of the 2005 Winter Simulation Conference, IEEE Press, Dec. 2005.
4. K. Farahmand and A. F. G. Martinez, "Simulation and Animation of the Operation of a Fast Food Restaurant," Proceedings of the 1996 Winter Simulation Conference, IEEE Press, Dec. 1996, pp. 1264-1271
5. K. Kharwat, "Computer Simulation: an Important Tool in The Fast-Food Industry," Proceedings of the 1991 Winter Simulation Conference, IEEE Press, Dec. 1991, pp. 811-815.
6. M. Laguna and J. Marklund, *Business Process Modeling, Simulation and Design*. ISBN 0-13-91519-X. Pearson Prentice Hall, 2005
7. J. D. C. Little, "A Proof for the Queuing Formula: $L = \lambda W$," Operations Research, vol. 9(3), 1961, pp. 383-387, doi:10.2307/167570.
8. K. Rust, "Using Little's Law to Estimate Cycle Time and Cost," Proceedings of the 2008 Winter Simulation Conference, IEEE Press, Dec. 2008, doi: 10.1109/WSC.2008.4736323.
9. T. C. Whyte and D. W. Starks, "ACE: A Decision Tool for Restaurant Managers," Proceedings of the 1996 Winter Simulation Conference, IEEE Press, Dec. 1996, pp. 1257-1263.