Modelling Single Channel Server (M/M/1) to the Problem of Queues in Banking Industry

Bello, A. Hamidu

Department of statistics, school of sciences, Federal university of technology, Akure - Ondo state, Nigeria
e-mail: habello@futa.edu.ng.

Abstract: The research is modelling single channel to the problem of queues in banking industry: A comparative study of Guaranty Trust Bank (GTB) and First Bank of Nigeria (FBN) at Federal University of Technology, Akure (FUTA) branch. Nigeria. The Objectives are to determine the average number of customers in a queue, average time customers spend on queue at the Automated Teller Machine (ATM), average waiting time of customers in the system, compute the probability that server would be idle in a single day. Data were primarily collected and recorded based on the arrival pattern and service pattern of customers. One sample Kolmogorov-Smirnov test was performed on the collected data and this established that the arrival and service time of customers at the ATMs centre are Poisson and exponentially distributed respectively. The M/M/s queuing model is therefore considered to be the best in modelling the queuing system. The results obtained from the estimation of parameters showed that the number of customers that arrive per unit time for GTB is greater that of FBN (0.4885 > 0.3712). It can be seen that the traffic intensity at GTB is more than that of FBN which means there will be more customers on the queue at GTB than FBN which is also confirmed by the values of the expected number of customers in the queue. Also, the probability that a customer will have to wait for service at GTB is 0.7 which is higher than that of FBN which is 0.6. In conclusion the service level increases at an optimal service level, as the waiting time of the customer is also reduced during the off pick period and the finding also shows that using FBN ATM is faster than GTB ATM.

[Bello, A. Hamidu.. **Modelling Single Channel Server (M/M/1) to the Problem of Queues in Banking Industry.** *Researcher* 2018;10(11):25-31]. ISSN 1553-9865 (print); ISSN 2163-8950 (online). http://www.sciencepub.net/researcher. 5. doi:10.7537/marsrsj101118.05.

Keywords: Modelling, Kolmogorov – Smirnov test, M/M/1 Queuing model, ATM.

1.0 Introduction

Queues (or waiting lines) are general phenomenon in everyday life which is usually seen at post offices, bus stops, hospitals, bank counters, petrol filing stations etc. Queues are formed when customers (human or object) demanding service have to wait because their number exceeds the number of servers available; or the facility does not work efficiently or takes more than the time prescribed to service a customer. Queues are an everyday common sight in banks, especially at the bank`s automated teller machines (ATM). Abor (2005) in citing Rose (1999) gives a fitting description of ATM as "… a computer terminal, record-keeping system with cash vault in one unit, permitting customers to enter the bank`s book keeping system with a plastic card containing a Personal Identification Number (PIN) or by punching a special code number into the computer terminal linked to the bank`s computerized records 24hours a day". ATM as an electronic delivery channel has come a long way very much as a result of the success of its forerunners. The first instance of such machine started operations at a branch office of Barclays Bank in 1967 in the United Kingdom. These were initially designed simply as cash dispensers to allow clients to perform cash withdrawals. Their utility has however grown to allow advance operations such as allow payments of utility bills, cash deposits, transfer of money from one bank account to another and much more. The issue of queue control in banking halls via ATM is essentially defeated because ultimately access to these facilities is not limited to customers going to the bank. ATMs themselves have as a result become subjects of large service demands which directly translate to queues for service when demands cannot be quickly satisfied. This situation becomes more compounded and more evident during festive periods and months ends when the demand for cash is high. Queuing theory basically a mathematical approach, used for the analysis of waiting lines. It deals with the problems that involve waiting (or queuing). Queuing theory is used to analyse the congestions and delays of waiting in line. It is used to develop more efficient queuing system that reduce customer waiting time and increase the number of customer that can be served.

The aim of this study is to model single-channel (M/M/1) to the problem of queue in GTB and FBN. The objectives is the comparative study of the two banks to determine the average waiting time and service time of customers on the queue and in the system. The method of data collection is through direct observation, a wrist watch, a pen and a notepad were requirements needed for the recording of relevant information such as; number of customers, the arrival times of customers and service time. The observation was made during the working and closing hours (10am – 12am) and (4pm -6pm).

Queuing models are used to represent the various types of queuing systems that arise in practice, the models enable in finding an appropriate balance between the cost of service and the amount of waiting. Some of the analysis that can be derived using queuing theory include the expected waiting time in the queue, the average waiting time in the system, the expected queue length, the expected number of customers served at one time, as well as the probability of the system to be in certain states, such as empty or full (Allen, 1990).

Queuing theory had its beginning in the research work of a Danish engineer named Erlang around 1909. He conducted an experiment and had observed that the demand in telephone traffic is fluctuating. Latter he published a report addressing the delays in automatic dialling equipment. Soon after his published work and during the end of World War II, Erlang (1918) early work was extended to more general problems and to business applications of waiting lines and to various service industries. Modelling a service industries as a queue system has many advantages such as diagnosing problems, identifying constraints so as to understanding the real systems rather than indicating individual prediction about the system.

A determined attempt was made by Cooper (1990) to show that from the stand point of a queuing model, a waiting line situation is created in the following manner. As the customer arrives at the facility, he joins a waiting line (or a queue). The server chooses a customer from the waiting line to begin service, upon the completion of a service, then process of choosing a new (waiting) customer is repeated.

Chee-Hoch (2002) defined queuing as an aggregation of items awaiting a service function. They viewed queuing as being broad as queue may consist of people, cars, components awaiting Machining telephone calls, aeroplane indeed and discrete items. But in our everyday lives we tend to think of queue as dating to people only. They are some of the factors whose relationship to increase in productivity the researcher sought to establish.

Stevenson (1972) was a steady-state version of the M/M/S model to determine the required number of ambulance, in a regime. Bell and Allen (1969) developed queuing theory for determining the number of ambulances required to respond immediately to 95 or 99 percent of service requests. Smith (1977) argued through his schematic approval to the system that a queuing system contains several components. He listed these components as calling population, calling unit, and service discipline and service facility.

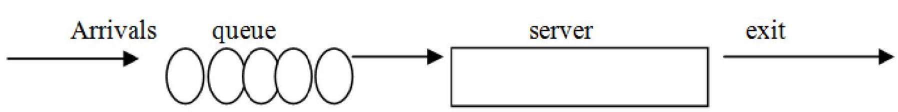
2.0 Methodology

The data for this study was obtained primarily through direct observation at the bank's ATM. The researcher had to relate with the managers of the banks to inform them about the research and how the banks have been chosen as a case study for this project and got the permission to go ahead. The following week, a form that consists the arrival time and service time was printed for taking record of customers as they arrive and join the queue at the ATM until they leave. This was done on Mondays to Fridays between (10:00am to 12:00am) and (4:00pm to 6:00pm). The main reason for choosing the ATM of this bank and not the Banking Hall is because most of their customers are students and not working class, and majority of them prefer using ATM to teller/cheque book. Thus, the researcher recorded the following events as it happened at the ATM using a Digital wrist watch.

The method of analysis for this study is the single-server queuing modeling system which follows (M/M/S): (∞/FCFS) specification. In this case, the performance measure analysis includes the arrival time, waiting time and service time for average customers using the appropriate tools.

Single channel, single phase

The simplest form of queue is a simple line of a customer waiting to be served by a single service point called single channel/single phase. It is applicable in, banks, post offices etc.



Secondly, the graphical representation of the generated performance measure values was done and has effectively employed the following in the analysis:

1. Test whether the arrival and service time follows exponential or poisson distribution
2. Using R-package to determine the following

The average number of customers in a queue:

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

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

$$= \lambda\left[\frac{\mu-\mu+\lambda}{\mu(\mu-\lambda)}\right] = \frac{\lambda}{\mu}\frac{\lambda}{\mu-\lambda},$$

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

Therefore, $L_q = \frac{\rho\lambda}{\mu-\lambda}$

The probability that server would be idle in a single day, P_o:

$$\frac{1}{1+\frac{\lambda/\mu^{\,c}}{c!\left[1-\frac{\lambda/\mu}{c}\right]}+\frac{\lambda/\mu^{\,1}}{1!}+\frac{\lambda/\mu^{\,2}}{2!}+\cdots+\frac{\lambda/\mu^{\,c-1}}{(c-1)!}}$$

Average time customers spend on queue at the ATM

$W_q = \frac{L_q}{\lambda} = \frac{\rho}{\mu-\lambda}$

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

The average waiting time of customers in the system

$W_s = \frac{\text{expected number of units in the system}}{\text{Arival rate}} =$

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

Variance of queue length:

Var (n)$=E(n)^2 - [E(n)]^2 =$

$\sum_{n=0}^{\infty} n^2 P_n - [\sum_{n=0}^{\infty} nP_n]^2$

(for n=0,both terms are zero only)

$= \sum_{n=1}^{\infty} n^2 P_n - [\sum_{n=1}^{\infty} nP_n]^2$

$= \sum_{n=1}^{\infty} n^2.\left(1-\frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n - [L_s]^2$

$= \left[1-\frac{\lambda}{\mu}\right]\left[1.\frac{\lambda}{\mu}+2^2.\left[\frac{\lambda}{\mu}\right]^2+3^2.\left[\frac{\lambda}{\mu}\right]^3+\cdots\right]-\left(\frac{\lambda}{\mu-\lambda}\right)^2$

$= \frac{\lambda}{\mu}\left[1-\frac{\lambda}{\mu}\right]\left[1+2^2.\frac{\lambda}{\mu}+3^2.\left[\frac{\lambda}{\mu}\right]^2+\cdots\right]-\left(\frac{\lambda}{\mu-\lambda}\right)^2$

Let S = $1+2^2.\frac{\lambda}{\mu}+3^2.\left[\frac{\lambda}{\mu}\right]^2$

$= 1+2^2\rho+3^2\rho^2+\ldots\ (\rho=\frac{\lambda}{\mu})$

Integrating both sides with respect to ρ from 0 to ρ, we have,

$=$

$\int_0^{\rho} S.d_p = \int_0^{\rho}(1+2^2\rho+3^2\rho^2+\ldots)d_p = [\rho+2\rho^2+3\rho^3+\cdots]_0^{\rho}$

$= \rho+2\rho^2+3\rho^3+\cdots = \rho[1+2\rho+3\rho^2+\cdots]$

Where, $1+2\rho+3\rho^2+\cdots = \frac{1}{(1+\rho)^2}$

$= \rho.\frac{1}{(1+\rho)^2} = \frac{\rho}{(1+\rho)^2}$

Now, differentiate both sides with respect to ρ, we have

S=

$\frac{1}{(1+\rho)^2}+\rho(-2).(1-\rho)^{-3}(-1) = \frac{1}{(1+\rho)^2}+\frac{2\rho}{(1+\rho)^3}$

$= \frac{1+\rho}{(1-\rho)^3}$, recall $\rho=\frac{\lambda}{\mu}$

$= \frac{1+\frac{\lambda}{\mu}}{\left(1-\frac{\lambda}{\mu}\right)^3}$

Var (n)$= \frac{\lambda}{\mu}\left[1+\frac{\lambda}{\mu}\right]\frac{\left[1+\frac{\lambda}{\mu}\right]}{\left[1-\frac{\lambda}{\mu}\right]^3} - \left[\frac{\lambda}{\mu-\lambda}\right]^2$

Variance of queue length =

$= \frac{\frac{\lambda}{\mu}\left[1+\frac{\lambda}{\mu}\right]}{\left(1-\frac{\lambda}{\mu}\right)^2} - \frac{\frac{\lambda^2}{\mu^2}}{\left(1-\frac{\lambda}{\mu}\right)^2} = \frac{\frac{\lambda}{\mu}}{\left(1-\frac{\lambda}{\mu}\right)^2}$

The probability that an arriving customer has to wait for service

$P_w = \frac{1}{k!}\left(\frac{\lambda}{\mu}\right)^k\frac{k\mu}{k\mu-\lambda}P_o$

Since k=1, $P_w = \frac{\lambda}{\mu}\frac{\mu}{\mu-\lambda}P_o$

3.0 Data Analysis

In other to carry out the necessary estimation on the data and also to know if the distribution is Poisson, Three main assumptions must hold:

1. Arrivals are random.
2. The number of arrivals is non-overlapping intervals (i.e. two customers cannot stand side by side on a queue except one stand in front of the other).
3. The Inter arrival times are independent and obey the exponential distribution.

The first two assumptions holds for the distribution in the data gotten (i.e. the arrivals are random and the number of arrivals is non-overlapping intervals). To continue with the estimation, It must be shown if the inter arrivals is exponentially distributed. **A Kolmogorov– Smirnov test** was carried out with the aid of a statistical package for social sciences (SPSS) and the output was:

Table 1: Descriptive Statistics

DAY/WEEK	N	Mean	Std. Deviation	Minimum	maximum
GTB WEEK1	528	2.15	2.624	0	22
GTB WEEK2	600	1.96	2.229	0	26
FBNWEEK1	406	2.83	2.886	0	17
FBN WEEK2	446	2.57	2.426	0	20

The table above shows the descriptive statistics of the test distribution on a weekly basis, for instance it shows the mean (total inter arrival divided by the total number of customers recorded for each day) and also shows the standard deviation (a measure of how spread out data values are around the mean for each week or the square root of the variance), together with the minimum (smallest) and maximum (largest) of the test data.

Table 2: Kolmogorov Smirnov Test

METHOD	GTB WK1	GTB WK2	FBN WK1	FBN WK2
Kolmogorov-smirnov Z	9.942	7.601	5.425	4.930
Asymp. Sig. (2-tailed)	.000	.000	.000	.000

METHOD	GTB	FBN
Kolmogorov-smirnov Z	12.201	7.333
Asymp. Sig. (2-tailed)	.000	.000

The above table shows the output of one sample kolmogorov-sminov test that was carried out for the test distribution.

Test of Hypothesis

Ho = the distribution does not follows an exponential distribution

H_1 = the distribution follows an exponential distribution

Significance level = 0.05

From the output, the p-values is less than the significance level (say 0.05) which implies that the test distribution (inter arrival) is exponential hence the number of arrival is Poisson distributed.

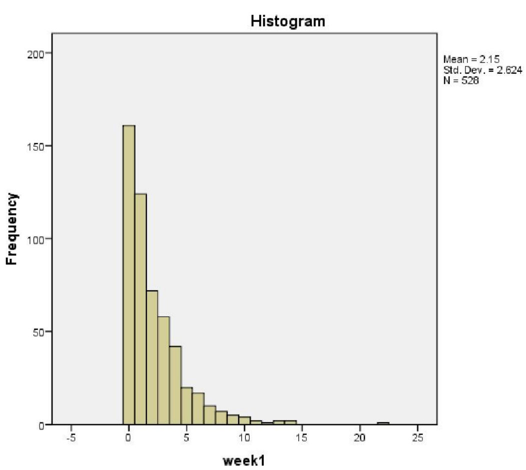


Fig 1: histogram of arrival time for GTB week1

This is the graph of inter arrival time.

This graph is positively skewed which shows that the mean is greater than the median. Thus, the arrival time follows a Poisson distribution.

The histogram above shows the difference between arrival of customer and it can be observed that the chance that a customer will arrive a minute after a customer is very high and with such the queue will be long at that period.

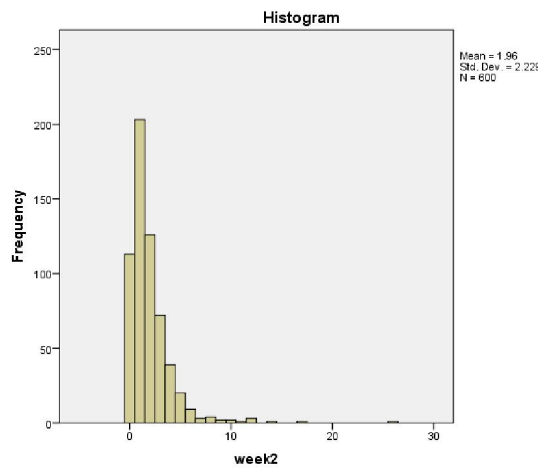


Fig 2: Histogram showing arrival time for GTB week 2

This is the graph of inter arrival time showing that the inter arrival time for the second week in GTB follows a Poisson distribution.

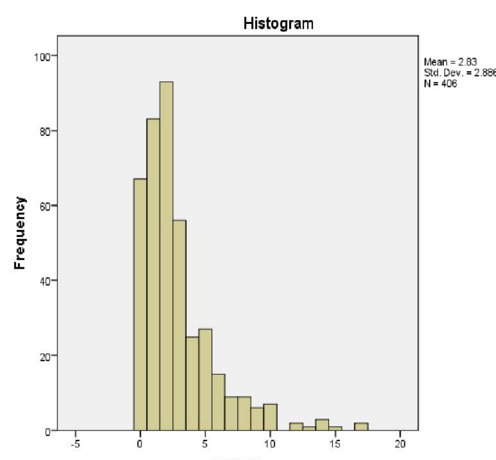


Fig 3: histogram showing inter-arrival time for FBN week 1

This is the graph of inter arrival time showing that the arrival time for the first week in FBN follows a Poisson distribution.

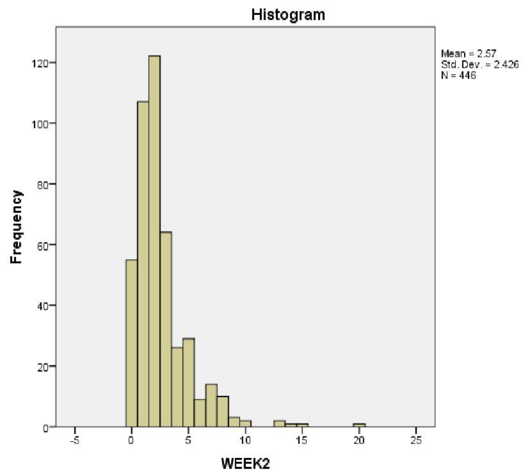


Fig 4: Histogram showing arrival time for FBN week 2

This is the graph of arrival time showing that the inter arrival time for the second week in FBN follows a Poisson distribution.

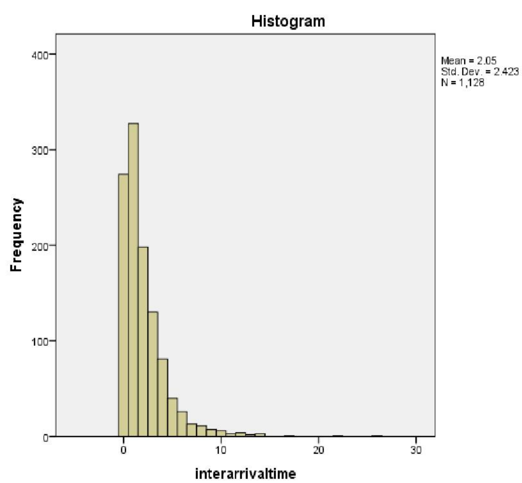


Fig 5: Histogram showing total arrival time for GTB

This is the graph of arrival time, showing that the arrival time for GTB follows a Poisson distribution.

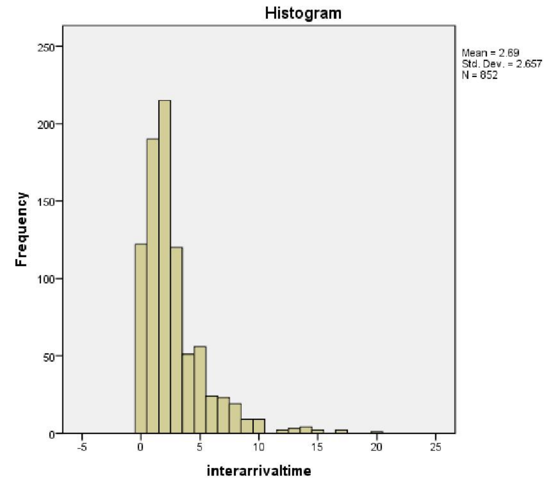


Fig 6: histogram showing total arrival time for FBN

This is the graph of arrival time showing that the inter arrival time for FBN follows a Poisson distribution.

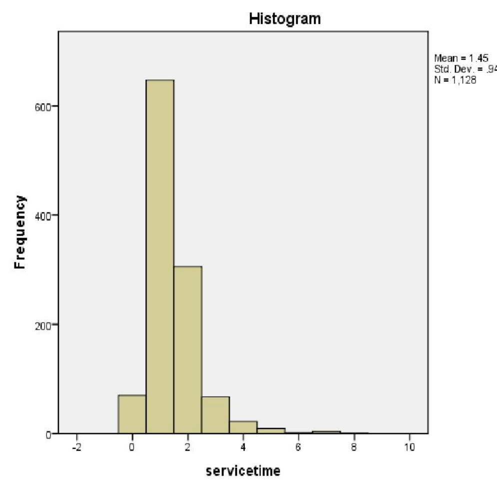


Fig 7: histogram showing service time for GTB

This is the graph of service time showing that the service time in GTB follows an exponential distribution.

3.1 Parameter Estimation

From the data collected for GTB,
The total number of customers, N: 1128
The total arrival time for customers: 2309
Time taken by the customers to be served, S: 1641
Number of server, = 1
From the data collected for FBN,
The total number of customers, N: 852
The total arrival time for customers: 2295

Time taken by the customer to be served, S; 1396

Number of server: 1

From the estimated parameters above, using a statistical package (R-package) the values for mean arrival time (λ), Mean service time (μ), Intensity (ρ), Expected number of customers in the queue (L_q), probability that the server will be idle (P_o), Average waiting time of a customer in the queue (W_q), average waiting time of a customer in the system (W_s) and the probability that an arriving customer has to wait for service (P_w) were obtained.

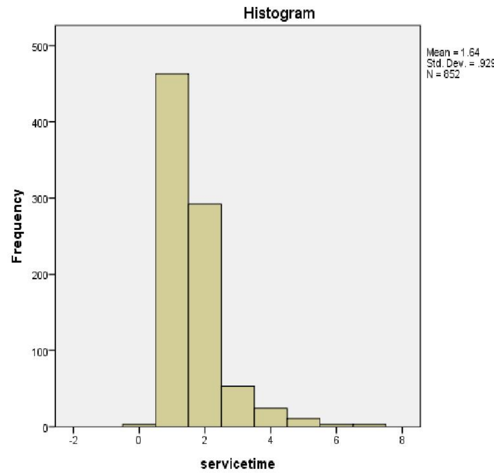


Fig 8: histogram showing service time for FBN

Table 3: Result of the Analysis

	Mean Arrival Time	Service Rate	Intensity	L_q	P_o	Wq	W_s	P_w
GTB	0.4885	0.6874	0.7107	1.7459	0.2893	3.5738	5.0286	0.7107
FBN	0.3712	0.6103	0.6083	0.9446	0.3917	2.5443	4.1828	0.6083

From Table 3, it can be observed that:

GTB customers wait longer on the queue than FBN customers.

The actual time of service delivery for GTB is 0.6874.

The actual time of service delivery for FBN is 0.6103.

This implies that the service rate shows that more customers are served per unit time at GTB than at FBN. Although GTB customers and FBN customers spend almost the same time to get service.

The mean arrival time shows that the number of customers that arrive per unit time for GTB is greater that of FBN (0.4885 > 0.3712).

The expected number of customers in a queue at GTB is 1.74 which means that at least 2 people will be on the queue and the expected number of customers at FBN is 0.94 which means at least 1person will be on the queue.

The average waiting time a customer spends at the queue for GTB and FBN is 3.5738 and 2.5443 respectively which implies that GTB customers spend more time on the queue.

The average waiting time a customer spends in the system for GTB and FBN is 5.0286 and 4.1828 respectively which implies that GTB customers spend more time in the system i.e. a customer has to wait about a minute longer at the system at GTB than at FBN.

It can also be observed that the probability that a server will be idle is higher for FBN than for GTB (0.3917 > 0.2893). The probability that a customer has to wait for service at GTB is 0.7 which is higher than that of FBN (0.6). It can be seen that the traffic intensity of GTB is more than that of FBN which means there will be more customers on the queue for GTB than for FBN; this is confirmed by the values of the expected number of customers in the queue.

4. Conclusions

The traffic intensity of GTB is more than that of FBN which means there will be more customers on the queue for GTB than for FBN; this is confirmed by the values of the expected number of customers in the queue. Also, the probability that a customer will have to wait for service at GTB is 0.7 which is higher than that of FBN which is 0.6. Therefore for quick service, one should go to First Bank Nigeria (FBN).

References

1. Allen, A. O. (1990). Probability, statistics, and queueing theory with computer science applications, 2nd edition. In *academic Press, Inc., Boston, MA.*
2. Anisimov, V., Zakusilo, O., and Donchenko, V. (1987). Elements of queueing theory and asymptotic analysis of systems. In K. Vishola Skola.
3. Chee-Hock, N., and Boon-He, S.. (2002). Queueing modelling fundamentals, 2nd ed. In W. A. SON. chichester.
4. Cooper, R.. (1990). Introduction to Queueing Theory, 3-rd Edition. CEE Press. Washington.

5. Erlang, A. (1909). *The theory of probabilities and telephone conversations. Nyt Tidsskrift for matematik B20, pp 33-39.*
6. Erlang, A. (1918). Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges. *The Post Office Electrical Engineers' Journal 10*, 189-197.
7. Hillier F. S. and Lieberman G. J. (2007). *"Introduction to Operations Research," Tata Mc-Graw Hill, 8th Edition.*
8. Gross, D., Shortle, J., Thompson, J., and Harris, C. (2008). Fundamentals of queuing theory, 4th edition. In *John Wiley and sns.* New York.
9. H. M. Xiao, Z. C. Zang, Y. X. Zhang. (2003). "Operational Research,"Chengdu, China, Electronic Science and Technology University Press,.
10. H. Zeng. (2007). Queue based on the cost of distribution systems,, Vol. 30, No. 3, pp. 330-332. *Journal of Wuhan University of Science and Technology*.
11. Haghighi, A., and Mishev, D.. (2008). Queueing models in industry and business. Nova Science Publishers, Inc.,. New York.
12. Khintchine, A.. (1969). Mathematical methods in the theory of queueing. Hafner. New York.
13. Kleinrock, L.. (1975). *Queuing systems. Vol. I. Theory (in Hungarian). Muszaki Kiadó, Budapest.*
14. Kleinrock, L.. (1976). *Queuing systems. Vol. II: Computer applications. John Wiley & sons.* New York.
15. Kleinrock, L.. (n.d.). *Queuing systems. Vol. I. Theory. John Wiley & Sons.* New York.

11/17/2018