# MA022

# Queueing Theory: Optimisation of Queue Times at Grocery Stores

# Report

The report describes the actual laboratory, field or theoretical research (not library research) for the project. It should involve ideas and preferably, data generated by the student.

Credit must be given to the appropriate source and reference, including, in the event that generative artificial intelligence was used, how it was utilised in the authoring of the project report, beyond basic spelling and grammar checks.

It should include:

(a) Background and Purpose of the research area,
(b) Hypothesis of the research,
(c) an Experimental section including Methods and Results,
(d) Conclusion including a Discussion of the Results and Implications and
(e) Bibliography of References.

Students are advised not to exceed seven pages, inclusive of data in tables, figures and diagrams. References and appendices do not count towards the seven-page limit. To ensure fairness across submission, judges have been advised that they can ignore any part of a report that exceeds the seven-page limit.

# 1. Introduction

Queues play an integral role in many parts of our life, being integrated into a wide range of services that we use daily such as banks, restaurants, public transport and retail. Despite this, we frequently encounter scenarios in which a service that usually require a minute to use taking much longer due to unreasonable queueing times and delays. Naturally, such situations are undesirable and should be minimised.

At first glance, the obvious solution would be to simply increase the number of service counters, or in more general terms, increase the level of service. Unfortunately, this also results in additional operation costs for companies. If too much is spent for a level of service that is unnecessarily high, it can become counterproductive overall. Thus, our goal is to find the best possible level of service for a reasonable cost, which we can define as the efficiency of the queue (an arbitrary function taking both level of service and cost of operation as variables). We can achieve this using Queueing Theory.

Queueing Theory is the mathematical study of queues and is used to characterise queue systems, as well as analyse certain measures of performance such as the expected length and waiting time of these queue systems. In this study, we applied Queueing Theory and simulations to identify the optimal queueing model among the various models observed at separate checkout systems present at a FairPrice supermarket outlet.

We hypothesised that we could model the queues as Poisson queues, and that the optimised system will be able to reduce queue times while minimising cost of running the service.

# 2. Methodology

## *2.1. Data Collection*

Data was collected in real-time from a FairPrice supermarket outlet in Singapore. We categorised the checkout systems observed into 2 types: Self-checkout and Cashier-operated checkout, modelled as $(\boldsymbol{G}/\boldsymbol{G}/\boldsymbol{c_1}):(\boldsymbol{FIFO}/\infty/\infty)$ and $c_2 \times (\boldsymbol{G}/\boldsymbol{G}/\boldsymbol{1}):(\boldsymbol{FIFO}/\infty/\infty)$ queues respectively following Kendall's notation, where $c_1$ is the number of cashier-checkout counters and $c_2$ is the number of self-checkout counters. We define a "customer" as the entity/entities using a single service/self-checkout counter at any given time. We observed that $c_1 = 14$, and $c_2$ was dependent on the no. of counters open that day, which ranged from 4 to 5.

Our study was based on the following assumptions:

1. Capacity of both systems is infinite as the queues do not have a specified length and the outlet can hold a sufficiently large number of people.
2. Size of the source is infinite as our source (population of Singapore) is an overwhelmingly large number as compared to the size of the system.
3. For cashier-checkout counters, all other counters operate at the same rate as the recorded counter.

The following data was collected for each day over the course of 11 weekdays at 1-hour intervals between 3pm-4pm:

- No. of arrivals (no. of customers entering the queue system)
- No. of departures (no. of customers served)
- Time taken to serve each customer

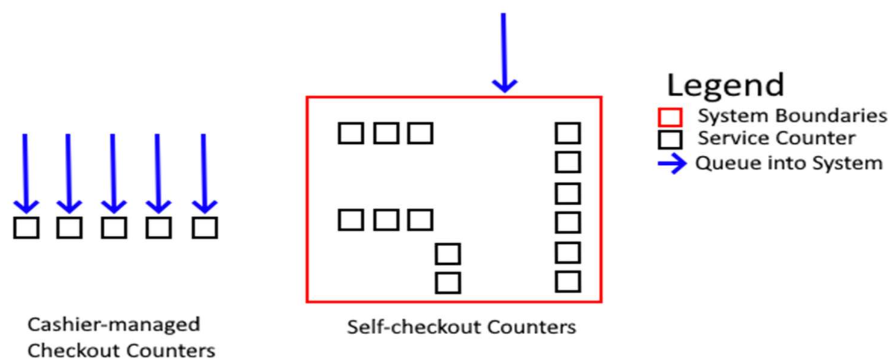



*Figure 1: Diagram representing the queue systems observed in the supermarket outlet*

For the cashier-operated checkouts, we recorded data for one queue that led into one counter, However, for the self-checkout counters, we recorded data based on the entire system of one queue leading to many counters. Recording "Time taken to serve each customer" for the self-checkout counters was done by arbitrarily choosing a counter at the start of each round of data collection and timing each customer that used that specific counter. This decision was made based on the 3rd aforementioned assumption, while also considering the infeasibility for a singular person to simultaneously record more than 5 timings at once.

*2.2. Analysis of Different Queueing Models*

While the proposed models are of the characteristics $(G/G/1)$ and $(G/G/c)$, it is complicated to obtain an accurate prediction of performance, and we typically resort to approximation or bounding. In this case, we used Kingman's equation to approximate the waiting time of

$(G/G/1)$ queues. It should be noted that these bounds are rather "loose and may not be very useful in practice". [3] Comparatively, fewer results have been discovered for $(G/G/c)$ queues.

In Queueing Theory, an $(M/M/1)$ queueing system is the "simplest non-trivial queue" [4] that we can analyse, where both arrivals and departures follow a Markovian process characterised by an arrival rate following a Poisson distribution with parameter $\lambda$ and independent service times exponentially distributed with parameter $\mu$. [4]

Following the assumption that both arrivals and departures were Markovian, we calculated the performance of the selected queues and branched out from there into considering other queue types.

*2.3 Simulation of Performances for Optimisation*

Assuming our queues can be modelled as $M/M/1$ and $M/M/C$ queueing systems, we attempted to find the best parameters possible. We also assumed that $\lambda_{\text{total}}$ of both queue systems combined remained constant due to the total number of customers using both queueing systems will remain the same, regardless of which queue system they choose. Thus, an increase in no. of $M/M/1$ queues $c_1$ would lead to $\lambda_{\text{new M/M/1}} = c_1 \times \lambda_{\text{M/M/1}}$, while $\lambda_{\text{M/M/}c_2} = \lambda_{total} - \lambda_{\text{new M/M/1}}$, subsequently affecting performance of both queues.

We modelled the following cost functions and summed them up to form total cost:

- **Operation cost** = $\boldsymbol{C_{\text{cashier}} \times c_1 + C_{\text{self}} \times c_2}$
- **Customer cost** = $\boldsymbol{(sum\ of\ W_s\ for\ both\ queues) \times C_{\text{wait}}}$

We then plotted graphs of the 3 cost functions to find the best parameters for this system.

## 3. Results and Discussion

*3.1 Comparing the Performance of M/M/1 vs M/M/c*

Using our obtained data, we calculated the following variables for both queue types:

- $\lambda$ = Average no. of arrivals
- $\mu$ = Average time taken to serve one customer

We obtained $\boldsymbol{\lambda = 37.091, \mu = 40.717}$ **for the** $(\boldsymbol{M/M/1})$ cashier-operated system and $\boldsymbol{\lambda = 257.100, \mu = 664.622}$ **for the** $(\boldsymbol{M/M/c})$ self-checkout system respectively.

We used these variables to calculate the following relevant steady-state measures of performance. The formulae used can be found in the *Appendix*.

- $L_s$: Expected no. of customers in system
- $L_q$: Expected no. of customers in queue
- $W_s$: Expected waiting time in system (h: hours | min: minutes | s: seconds)
- $W_q$: Expected waiting time in queue (h: hours | min: minutes | s: seconds)
- $\bar{c}$: Expected no. of busy servers

| **Queue Type** | $\boldsymbol{\rho}$ | $\boldsymbol{L_s}$ | $\boldsymbol{L_q}$ | $\boldsymbol{W_s}$ | $\boldsymbol{W_q}$ | $\boldsymbol{\bar{c}}$ |
|---|---|---|---|---|---|---|
| M/M/1 | 0.911 | 10.228 | 9.317 | 992.754 s | 904.339 s | 0.911 |
| M/M/c | 0.387 | 5.416 | 0.001 | 75.841 s | 0.008 s | 0.387 |

**The calculated values for the performance measures indicated that the self-checkout queues are a better system to use compared to the cashier-operated queues, being able to process customers faster due to shorter wait times overall.**

Additionally, $\rho = 0.911$ for the $(M/M/1)$ cashier-operated queues indicated that the system likely struggled to handle the influx of customers, even during non-peak hours of 3-4pm on weekdays. The queue in its steady state would have a new arrival almost at the same time a customer is finished checking out, causing ridiculously long queue times. A small increase in $\lambda$ or decrease in $\mu$ could also lead to the queue growing unstable and have continuously growing queue size, meaning new customers would "never get served" under steady state.

We noted that the calculated values are rather extreme in context and did not match up with the observed data. For the $M/M/1$ queue, while it was observed to almost always be in operation, the calculated waiting time of ~16.5min was nowhere close to what was observed. Thus, we revisited the initial assumption that our arrivals/departures rate followed a Poisson distribution as it was the most feasible reason for why the formulae did not line up with our observations.

### *3.2. Attempting to Prove Departures Rates follow a Poisson Distribution*

We plotted a histogram of service times for both queue types *(refer to Appendix)* and utilized Chi-square goodness-of-fit tests on both sets of data to determine if our departures rates followed a Poisson distribution. Bins with frequency < 5 were combined into larger bins before proceeding. We are unable to conduct the test on arrival rates due to lack of data.

**For cashier-operated checkout (G/G/1) departures,** we tested for the exponential distribution of service times. We obtained $\chi^2 = 118.204$, with a critical value of 42.557 for a test at the 5% significance level with $df = 29$, leading to the rejection of the null hypothesis.

**For self-operated checkout (G/G/c) departures,** we tested for the Poisson distribution of rate of service. We obtained $\chi^2 = 2432.103$, with a critical value of 21.666 for a test at the 1% significance level with $df = 9$, leading to a confident rejection of the null hypothesis.

Evidently, our departure rates did not follow a Poisson distribution and the same could likely be said for our arrival rates, meaning the initial assumption had been incorrect. Thus, we decided to further analyse the queues using the approximations known for the **expected waiting time** of $(G/G/1)$ queue, keeping the previous results from Section 3.1 in mind.

*(Note: We determined that the rate of departure does not follow a Poisson distribution, despite performing a chi-square test for an exponential distribution, because a Poisson distribution of departure rate would imply an exponential distribution of departure times as shown in the **Appendix**, and vice versa).*

*3.3. Approximating the Performance of the G/G/1Queues*

As mentioned in Section 2.2, the performance of $(G/G/1)$ queues can only be approximated in most cases. Kingman's formula was applied under the assumption of steady state:

$$E(W_q) \approx \left(\frac{\rho}{1-\rho}\right)\left(\frac{c_a^2 + c_s^2}{2}\right)\left(\frac{1}{\mu}\right)$$

In the equation above, $c_a$ and $c_s$ refer to the coefficients of variation for the interarrival and service times respectively. We found that $c_a = 0.185$ and $c_s = 0.530$. Thus,

$$\boldsymbol{E(W_q) \approx 0.040\ \textbf{hours}\ (2.371\ \textbf{min})}$$

the approximation shows that in general, the cashier-checkout queues should be able to handle customers to an adequate extent, which lines up with our observations since the average time spent in the queue observed was 88.415 seconds or 1.474 minutes.

*3.4. Simulating the Best Parameters*

Our simulation revealed that the cost of having customers wait follows a decreasing, concave up curve, meaning that as the no. of cashier-operated counters increases, the effect on decrease in customer cost decreases in significance. The simulation also showed that the lowest total cost was usually achieved when $c_1 = 4$, which matches up with the system at the FairPrice outlet where there were usually 4 open cashier counters and 14 self-checkout counters.

However, the simulation indicates that the best system out of the ones we tested were achieved using 5 cashier-operated and 3 self-checkout systems, with $W_s = 185.455s$ and $223.381s$ for the cashier-operated and self-checkout queue systems respectively with $Total\ Cost = 105.721$ as compared to the original $W_s = 992.754s$ and $76.014s$, with $Total\ Cost = 43.555$ which is a significant reduction in waiting time overall and cost of the system despite the increase in waiting time at the self-checkout system.

This result is suspicious as it is in stark contrast to the original system with 4 cashier-operated and 14 self-checkout counters, which was observed to be in use almost all the time. It is unlikely that this result is accurate considering the prior context, and thus we will not treat it as a notable result.

The graphs for the simulation can be found in the *Appendix*. The code used for the simulation, along with the cost functions used and the simulation results, can be found on the Github repository at https://github.com/appventuremoment/Queueing-Theory-Project.

## 4. Conclusion & Further Studies

To summarise, the $(G/G/c)$ self-checkout system is of better operating efficiency compared to the $(G/G/1)$ cashier-operated checkout system, under the assumption that they operate at steady-state conditions. The calculated measures of performance, namely $W_s, W_q, L_s$ and $L_q$ match up adequately with the observed statistics for both queues, except for the cashier checkout queues where the observed queue usually had $L_s$ and $L_q \approx 4$ to 6 customers at a time compared to the calculated $L_s = 10.228$ and $L_q = 9.317$.

We also discovered through simulations, that on average the better queueing system is obtained when 4 cashier-operated counters are used, which matches up with the current system in place at the FairPrice outlet.

One notable reason for the departure rates not following a Poisson distribution could be customers opting to use the self-checkout counters after noticing long queues at cashier-serviced checkout counters. This is further supported by the observation that customers seem to occur in waves as self-checkout counters, rather than at random.

**Applications**

The results from this project can be used to further optimize the queueing systems present in the FairPrice outlet, allowing for shorter queueing times and preventing a system where customers need to wait an unreasonable amount of time in order to be serviced, while keeping the operating costs at a minimum.

The code found in the Github repository can be utilized for general usage. It allows companies to input data regarding their arrival/departure rates for multiple queueing systems and operation costs to obtain the performance metrics of both queueing systems. The tool can also simulate performances for different combinations of queue systems to optimise the cost efficiency of the systems in place for these services.

**Further Studies**

There are many possible routes for further study, however we believe one of the most important ones would be to increase the quality of data collected, as there were multiple sources of disparity that could have led to inaccurate data.

One possible route would be to employ the use of more manpower or technology when collecting data to reduce the chances of error when collecting data. Additionally, this could open opportunities to measure data from each counter individually instead of assuming similar performance since there was a significant variance in queue length and rate of service at the different counters observed.

Another route for study is to increase the period of data collection, and over more time intervals. This would provide sufficient data to perform the chi-square test on interarrival time and allow for consideration of the comparison of queue performance over different times (peak vs. non-peak hours), and different days (weekdays vs. weekends).

Outside of increasing quality of data, queues in the transient state could also be accounted for, as the performance of the $(M/M/1)$ queue is likely to change over time and there would have been times where $\lambda > \mu$, breaking the rule of steady state.

Different types of queues outside of $(M/M/1)$ and $(M/M/c)$ could be analysed to allow for more extensive usage.

## 5. References

[1] Taha, Hamdy A. (1992). Operations Research: An Introduction (5th ed.). Macmillan.

[2] Banks, J., Carson, J. S., Nelson, B. L., & Nicol, D. M. (2004). Discrete-event system simulation (4th ed.). Pearson Education.

[3] Karmakar, S. (n.d.). Queueing theory: Slide set 12 [PDF document]. Retrieved from http://home.iitk.ac.in/~skb/qbook/Slide_Set_12.PDF

[4] Sztrik, J. (2012). Basic Queueing Theory. University of Debrecen.

[5] Bose, Sanjay K. (2002). An Introduction to Queueing Systems. Springer Science & Business Media.

[6] Romeu, J.L. (2003/2004). ***The Chi-Square: A Large-Sample Goodness of Fit Test.*** RAC START, Volume 10, Number 4. https://web.cortland.edu/matresearch/ChSqSTART.pdf

[7] Fiveable. (n.d.). *Formulae for M/M/c queue*. Retrieved January 15, 2025, from https://library.fiveable.me/stochastic-processes/unit-8/mm1-mmc-queues/study-guide/0OzdE5UrvJQbLoYX

[8] Picodi. (2023). *Average amount Singaporeans spend on groceries per month*. Retrieved January 15, 2025, from https://www.picodi.com/sg/bargain-hunting/spendings-on-food-2023

[9] Korona POS. (n.d.). *How much do self-checkout machines cost?*. Retrieved January 15, 2025, from https://koronapos.com/blog/how-much-do-self-checkout-machines-cost/

[10] Magestore. (n.d.). *Self-checkout systems in retail*. Retrieved January 15, 2025, from https://www.magestore.com/features/self-checkout-systems-in-retail/

[11] Glassdoor. (n.d.). *Part-time cashier salary in Singapore*. Retrieved January 15, 2025, from https://www.glassdoor.sg/Salaries/part-time-cashier-salary-SRCH_KO0,17.htm

## 6. Appendix

*Proving that if interarrival time follows an exponential distribution, then no. of arrivals follows a Poisson distribution*

A pure birth queue is a specific type of queue where only arrivals take place. In this case, we first define λ as the number of customers entering the queue per unit time. Given that the interarrival time is exponential, then

$$\begin{aligned} p_0(t) &= P\{interarrival\ time \geq t\} \\ &= 1 - P\{interarrival\ time \leq t\} \\ &= 1 - \left(1 - e^{-\lambda t}\right) = e^{-\lambda t} \end{aligned}$$

For a sufficiently small time interval $h > 0$, we have

$$p_0(h) = e^{-\lambda h} = 1 - \lambda h + \frac{(\lambda h)^2}{2!} - \cdots = 1 - \lambda h + O(h^2) - \cdots$$

The exponential distribution is built on the assumption that that for any sufficiently small time interval $x$, the probability of more than one arrival occurring is negligible compared to the probability of zero or one arrival. Thus, as $h \to 0$,

$$p_1(h) = 1 - p_0(h) \approx \lambda h$$

This result implies that the probability of an arrival during $h$ is almost directly proportional to h, with the arrival rate, λ, being the constant of proportionality.

We want to derive the distribution of *the number of arrivals* during a period $t$ when the interarrival time is exponential with the mean $\frac{1}{\lambda}$, we thus define $p_n(t)$ as the probability of $n$ arrivals during time $t$.

For a sufficiently small $h > 0$,

$$p_n(t + h) \approx p_n(t)(1 - \lambda h) + p_{n-1}(t)\lambda h, \qquad n > 0$$

$$p_0(t+h) \approx p_0(t)(1-\lambda h), \qquad n = 0$$

In the case of $n$ occurrences happening in the time $t+h$, it would be equal to the chance of $n$ occurrences happening during time $t$ multiplied by the chance of 0 occurrences happening during the time $h$ plus the chance of $n-1$ occurrences during time $t$ multiplied by 1 occurrence happening during time $h$. These are the only 2 cases due to the aforementioned way that exponential distribution is defined, where only 1 occurrence is allowed to happen in the very small duration of $h$.

For the special case of 0 occurrences in the time $t+h$, then it would be equal to the chance of 0 occurrences during the time t multiplied by the chance of 0 occurrences during the time h.

The property of the product of the chances of the number of occurrences in time $t$ and the chances of the number of occurrences in time $h$ being equivalent to the chances of the number of occurrences in time $t+h$ is due to the memoryless property exponential distributions. This makes arrivals independent of how much time has passed since the most recent arrival.

Rearranging the terms above and taking the limit as $h \to 0$ to obtain the first derivative of $p_n(t)$ with respect to $t$ gives us,

$$p_n(t+h) \approx p_n(t)(1-\lambda h) + p_{n-1}(t)\lambda h, \qquad n > 0$$

$$p_0(t+h) \approx p_0(t)(1-\lambda h), \qquad n = 0$$

$$p_n(t+h) - p_n(t) \approx -p_n(t)\lambda h + p_{n-1}(t)\lambda h, \qquad n > 0$$

$$p_0(t+h) - p_0(t) \approx -p_0(t)\lambda h, \qquad n = 0$$

$$p_n'(t) = \lim_{h \to 0} \frac{p_n(t+h) - p_n(t)}{h} = -\lambda p_n(t) + \lambda p_{n-1}(t), \qquad n > 0$$

$$p_0'(t) = \lim_{h \to 0} \frac{p_0(t+h) - p_0(t)}{h} = -\lambda p_0(t) \qquad n > 0$$

For the special case $n = 0$, we have

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t)$$

$$\frac{1}{p_0(t)} dp_0(t) = -\lambda \, dt$$

$$\int \frac{1}{p_0(t)} dp_0(t) = \int -\lambda \;\; dt$$

$$\ln[p_0(t)] = -\lambda t$$

$$p_0 = e^{-\lambda t}$$

For the more general case $n > 0$, we start with $n = 1$

$$\frac{dp_n(t)}{dt} = -\lambda p_n(t) + \lambda p_{n-1}(t)$$

$$\frac{dp_1(t)}{dt} + \lambda p_1(t) = \lambda e^{-\lambda t}$$

$$\int \frac{dp_1(t)}{dt} e^{\lambda t} + p_1(t)\, \lambda e^{\lambda t}\, dt = \int \lambda \, dt, via\ integrating\ factor\ of\ e^{\lambda t}$$

$$p_1(t)\, e^{\lambda t} = \lambda t, \qquad via\ product\ rule$$

$$p_1(t) = \lambda t e^{-\lambda t}$$

For $n = 2$

$$\frac{dp_2(t)}{dt} + \lambda p_2(t) = \lambda^2 t e^{-\lambda}$$

$$\frac{dp_2(t)}{dt} e^{\lambda t} + p_2(t)\, \lambda e^{\lambda t} = \lambda^2 t$$

$$\int \frac{dp_2(t)}{dt} e^{\lambda t} + p_2(t)\, \lambda e^{\lambda t}\, dt = \int \lambda^2 t\, dt$$

$$p_2(t)\, e^{\lambda t} = \frac{(\lambda t)^2}{2}$$

$$p_2(t) = \frac{(\lambda t)^2}{2} e^{-\lambda t}$$

Following this, we conjecture that $p_n(t)$ might follow a Poisson Distribution, which has the general term

$$p_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, n \in [0, \infty)$$

We can prove that this is true via induction. The base case of $p_0(t)$ has already been proven above. Now, we assume that for some arbitrary integer $k \geq 1,\ p_k(t)$ is true. We need to prove $p_{k+1}(t)$ is true, this referring to

$$p_{k+1}(t) = \frac{(\lambda t)^{k+1}}{(k+1)!} e^{-\lambda t}$$

We make use of the previous recurrence relation and substitute $p_k(t)$ into it

$$\frac{dp_{k+1}(t)}{dt} = -\lambda p_{k+1}(t) + \lambda p_k(t)$$

$$\frac{dp_{k+1}(t)}{dt} + \lambda p_{k+1}(t) = \lambda \frac{(\lambda t)^k}{k!} e^{-\lambda}$$

Multiplying both sides by the integration factor of $e^{-\lambda}$ and solving the first-order linear differential equation, we get

$$\int \frac{dp_{k+1}(t)}{dt} e^{\lambda t} + p_{k+1}(t)\lambda e^{\lambda t}\, dt = \lambda^{k+1} \int \frac{t^k}{k!} dt$$

$$p_{k+1}(t)e^{\lambda t} = \lambda^{k+1} \frac{t^{k+1}}{k! * k + 1} = \frac{(\lambda t)^{k+1}}{(k+1)!}$$

$$p_{k+1}(t) = \frac{(\lambda t)^{k+1}}{(k+1)!} e^{-\lambda t}$$

This matches $p_{k+1}(t)$. Therefore, if $p_k(t)$ is true, $p_{k+1}(t)$ will be true. Since $p_0(t)$ is true, then by mathematical induction, it can be said $p_n(t)$ is true $\forall\, n \geq 0$. This thus proves that the general term for the the probability of $n$ arrivals during time $t, p_n(t)$ follows a Poisson distribution.

We can take this a step further by proving the inverse. That is to say, if the number of arrivals during a specific time period $t$ follows a Poisson distribution, then we can determine the distribution of the interarrival time as being an exponential distribution.

More specifically, the number of arrivals during a specific time period $t$ is a Poisson distribution with the mean $\lambda$. We define $N(t)$ as the number of arrivals. The probability mass function of the Poisson distribution is

$$P(N(t) = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \qquad n \in [0, \infty)$$

We aim to find the probability distribution of the interarrival time $T$, which is the time between two consecutive arrivals. When an arrival takes longer than time $t$ $(T > t)$, then it means that within the time period $t$, there were 0 arrivals. Writing this out,

$$p_0(t) = P(N(t) = 0) = \frac{(\lambda t)^0}{0!} e^{-\lambda} \; = e^{-\lambda}$$

Referring back to the first line stating that the assumption of interarrival time being exponential and how the probability of interarrival time being greater than time $t$ is defined, we can see that this probability found here matches up with that.

We then attempt to find the cumulative distribution function

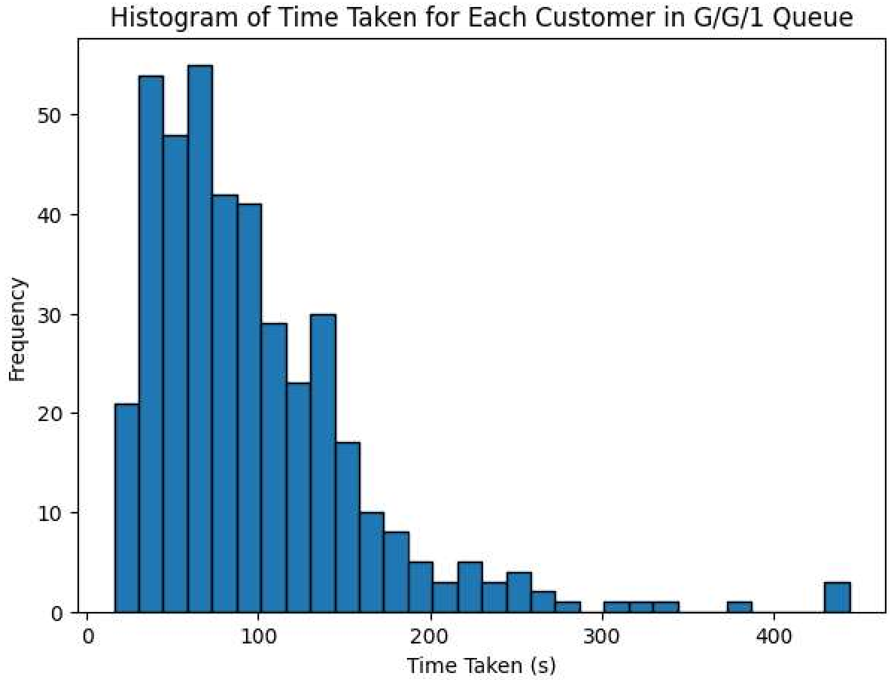
$$\begin{aligned} P\{interarrival\ time \leq t\} &= 1 - P\{interarrival\ time \geq t\} \\ &= 1 - p_0(t) \\ &= 1 - e^{-\lambda t} \end{aligned}$$

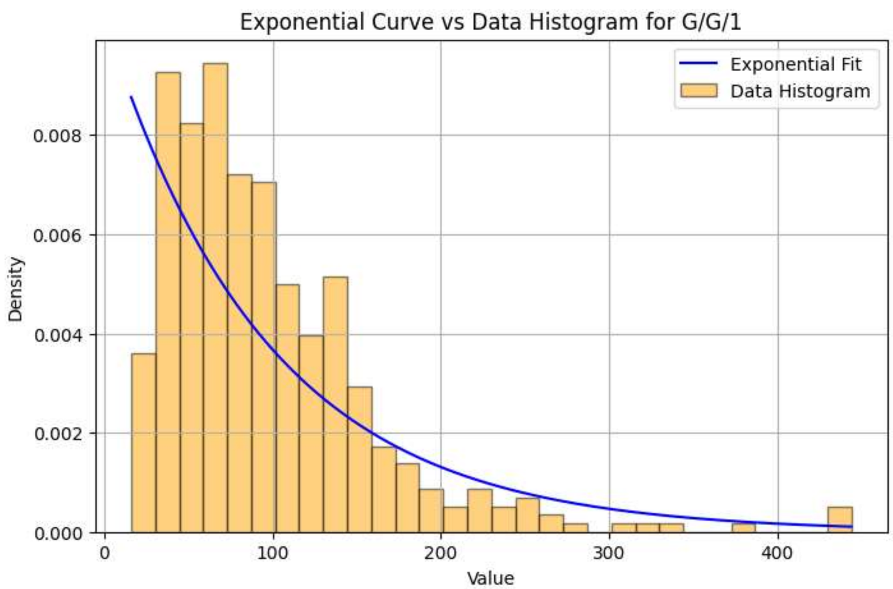
and probability density function of the interarrival time

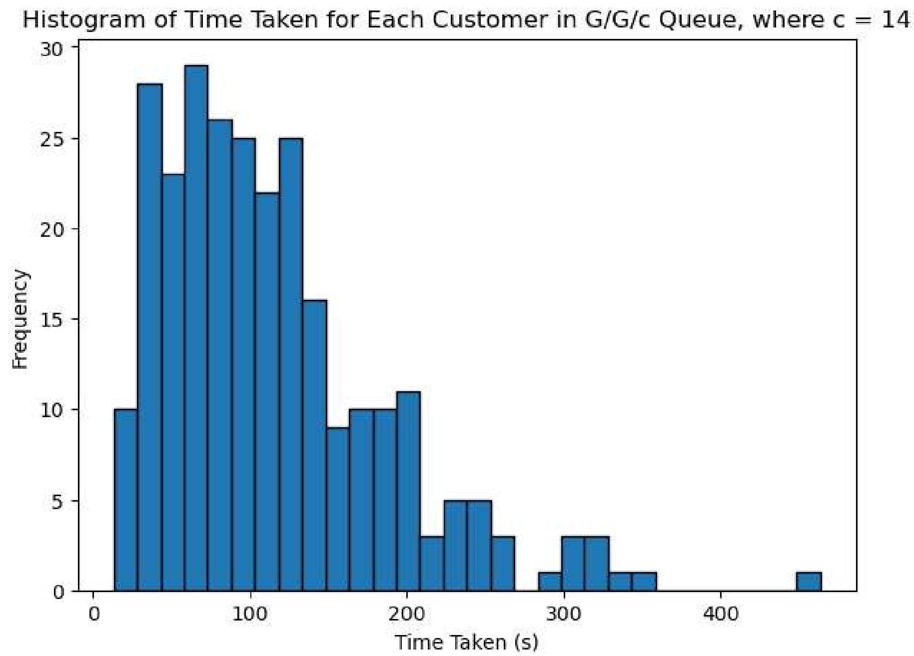
$$f(t) = \frac{dF(t)}{dt} = \frac{d}{dt}\left(1 - e^{-\lambda t}\right) = \lambda e^{-\lambda t}$$

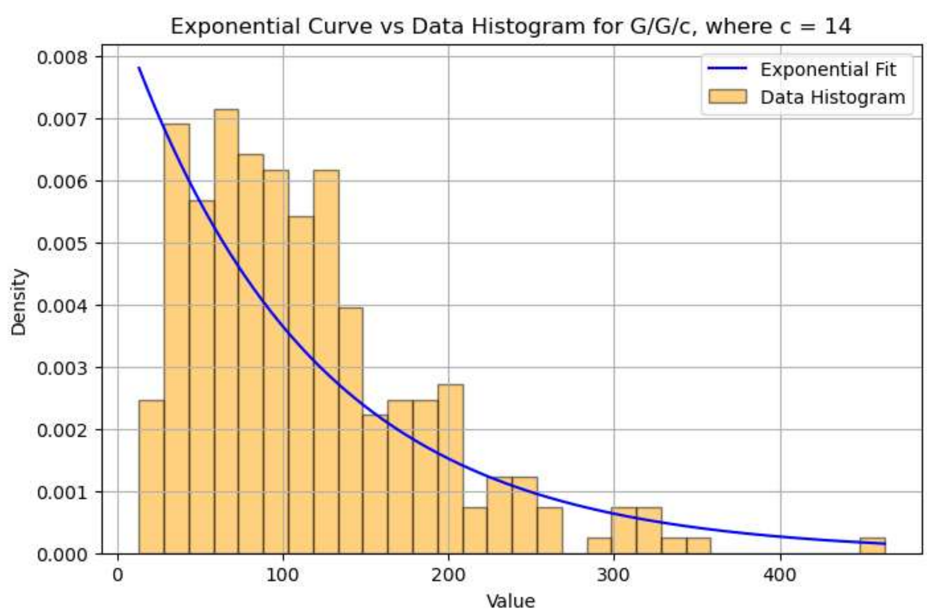
We recognize that both the cumulative distribution function and probability density function match that of an exponential distribution. Thus, this completes the proof that if the interarrival time is exponential, the number of arrivals $n$ during a specific time period $t$ will be Poisson and vice versa.

## ***Diagrams of Collected Data***




Histogram of Time Taken for Each Customer in G/G/c Queue, where c = 14
Frequency
Time Taken (s)
0
5
10
15
20
25
30
100
200
300
400


Exponential Curve vs Data Histogram for G/G/c, where c = 14
Exponential Fit
Data Histogram
Density
Value
0.000
0.001
0.002
0.003
0.004
0.005
0.006
0.007
0.008
0
100
200
300
400

**Formulae for M/M/1 queue:**

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

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

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

$$= (1-\rho)\rho \frac{d}{d\rho}\left(\frac{1}{1-\rho}\right)$$

$$= \boldsymbol{\frac{\rho}{1-\rho}}$$

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

$$\boldsymbol{W_q} = W_s - \frac{1}{\mu} = \frac{\rho}{\mu(1-\rho)}$$

$$\boldsymbol{L_q} = \lambda W_q = \frac{\rho^2}{1-\rho}$$

$$\bar{\boldsymbol{c}} = L_s - L_q = \rho$$

**Formulae for M/M/c queue:**

$$\rho = \frac{\lambda}{c\mu}, where\ \mu\ is\ the\ service\ rate\ for\ each\ counter$$

$$probabiltiy\ of\ 0\ customers\ in\ system, p_0 = \left[\sum_{n=0}^{c-1} \frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!\,(1-\rho)}\right]^{-1}$$

$$probability\ customer\ needs\ to\ wait\ in\ queue, P_w = \frac{(c\rho)^c}{c!} p_0$$

$$\boldsymbol{L_s} = L_q + c\rho$$

$$\boldsymbol{L_q} = \frac{\rho P_w}{\boldsymbol{1} - \rho}$$

$$\boldsymbol{W_s} = \frac{L_q}{\lambda} + \frac{1}{\mu}$$

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

$$\bar{\boldsymbol{c}} = L_s - L_q = \rho$$

**(Exponential) Chi-Square Test for Cashier-operated checkout:**

$\lambda_{\text{est}} = 0.01131$

Note: For the interval column, each row represents the left edge of each bin. The final bin covers the range 329.866 to ∞. (or in this case, 329.866 to 444.000 as the longest recorded time was 444s.)

We sum the values in the $\frac{(\boldsymbol{E_i}-\boldsymbol{O_i})^2}{\boldsymbol{E_i}}$ column to obtain $\chi^2$.

| Interval | $\boldsymbol{O_i}$ | $\boldsymbol{E_i}$ | $\frac{(\boldsymbol{E_i}-\boldsymbol{O_i})^2}{\boldsymbol{E_i}}$ |
|---|---|---|---|
| 0.000 | 0 | 0.000 | 0.000 |
| 16.000 | 21 | 47.416 | 14.717 |
| 30.267 | 54 | 40.914 | 4.185 |
| 44.533 | 48 | 35.304 | 4.565 |
| 58.800 | 55 | 30.463 | 19.763 |
| 73.067 | 42 | 26.286 | 9.393 |
| 87.333 | 41 | 22.682 | 14.794 |
| 101.600 | 29 | 19.572 | 4.542 |
| 115.867 | 23 | 16.888 | 2.212 |
| 130.133 | 30 | 14.573 | 16.333 |
| 144.400 | 17 | 12.574 | 1.558 |
| 158.667 | 10 | 10.850 | 0.067 |
| 172.933 | 8 | 9.362 | 0.198 |
| 187.200 | 5 | 8.079 | 1.173 |
| 201.467 | 8 | 12.986 | 1.914 |
| 230.000 | 7 | 9.669 | 0.737 |
| 258.333 | 5 | 14.702 | 6.402 |
| 329.866 | 5 | 9.339 | 2.016 |

**(Poisson) Chi-Square Test for Self-checkout checkout:**

$\lambda_{\text{est}} = 664.622$

We sum the values in the $\frac{(E_i - O_i)^2}{E_i}$ column to obtain $\chi^2$.

| **Interval** | $O_i$ | $E_i$ | $\frac{(E_i - O_i)^2}{E_i}$ |
|---|---|---|---|
| <620 | 164 | 11.426 | 2037.417 |
| 620-630 | 3 | 13.411 | 8.082 |
| 630-640 | 3 | 22.391 | 16.793 |
| 640-650 | 1 | 32.012 | 30.043 |
| 650-660 | 4 | 39.283 | 31.690 |
| 660-670 | 4 | 41.468 | 33.854 |
| 670-680 | 1 | 37.742 | 35.768 |
| 680-690 | 0 | 29.679 | 29.679 |
| 690-700 | 2 | 20.207 | 16.405 |
| >700 | 88 | 22.382 | 192.372 |

**<u>Graphs for simulation of different combinations of counters:</u>**