CS166: First Project

Modeling Grocery Store Queue

Minerva Schools at KGI

Introduction

We want to build a basic model of checkout queues in a given grocery store. Our final objective is to suggest the optimal number of cashiers needed for the grocery store. In this paper, I built a simple simulation and estimated the expected values of some important parameters (i.e. maximum queue length, average customer waiting time etc.) for different numbers of cashiers. Then analyzing the results, I suggested the optimal number of cashiers that they should employ in the store.

Assumptions and constraints

We are already given some assumptions and constraints in the prompt. We are assuming that the interval between two consecutive customers' arrival in the checkout area is distributed exponentially with an average rate (λ) of 1 customer per minute and after arriving at the checkout store, each customer chooses the shortest queue. Each queue is served by a single cashier whose service time is distributed normally with a mean (μ_1) and standard deviation (σ_1) of 3 minute and 1 minute per customer. 5% of the customers need additional support from the manager, whose service time is also distributed normally with a mean (μ_2) and standard deviation (σ_2) of 4 minute and 1 minute per customer.

Some given constraints of the model are that there is only one manager who has to help everyone among those 5% customers. The grocery store opens at 9 AM and closes at 6 PM and no customer can enter the queue after 6 PM, but the customers already in the queue have to be served before closing the store. Also the maximum number of cashiers they can employ is 10 due to the resource and size of the store.

Some other assumptions that I applied for simplification are that each customer and each cashier are identical and their average arrival rate and average service time stay constant over the day (i.e. I am ignoring the reality that the customers may rush before closing time or the cashier may slow down after lunch). The customers' arrival to the queue is independent of other people's arrival and the grocery store uses a first-come-first-serve method for serving them and their queues can be infinitely long if necessary(infinite buffer). Once the customer chooses a line, they will not change or leave the line before getting the service no matter how long they have to wait.

Variables and their update rules

Some important variables I used for the models include the number of cashiers, time after the opening of the store, cashier service time, manager service time, customer waiting time, length of the queue at a given time and overtime, maximum queue length in a day etc.

The rules to update the variables are mostly followed by our assumptions. We randomly draw a sample from the given exponential and normal distributions to get the next arrival time of a customer and the service time by a cashier or the manager. Customers join the queue with the shortest length and wait until all the people ahead of them in the queue are served. Once a cashier is done serving a customer, s/he takes the next one from his/her queue. Maximum queue length updates by comparing its previous value and the queue length after each customer joins a queue. A random sample from an uniform distribution in range 0 to 1 is drawn for each customer and if it is less than or equal 0.05, s/he needs extended support from the manager.

Empirical Model using Simulation

Using above mentioned assumptions and rules, I built a simulation using python 3 to estimate the important observables for different numbers of cashiers. Two important observables that are given are the estimated average customer waiting time, which is the time needed for a customer from starting in the queue to complete getting the service and estimated maximum queue length in a given day. In order to increase customer satisfaction (and consequently to increase the store branding and profit), we want the customer to wait as low as possible. Also if we can make sure the maximum queue length is smaller then we can use the space more efficiently.

As the third important observable, I chose the extra time that the cashiers (and manager, if necessary) have to stay after 6PM to complete serving the already queued customers. If they need to stay longer then the close time, it will increase their resource-usage and cost and depending on their rules, they either have to extra money to the cashiers or the employee satisfaction will decrease. So we want this extra time to be as small as possible. We calculated this by subtracting the expected closing time (9 hour or 540 minute after opening the store) from the simulation time when they finished serving everyone.

In the simulation, I added one queue per each cashier. When a customer arrives s/he chooses the queue with the smallest number of people in that queue and that cashier. Once s/he is in the queue we start counting his/her waiting time until s/he gets to the server and adds this time with the serving time. Then s/he leaves the server and if s/he needs the manager's support, he enters in the manager queue, where we again count his/her waiting and serving time. When the scheduled time is greater than 540 minute, no additional customer can join the queues and when there is none left in the schedule priority queue, we end the simulation.

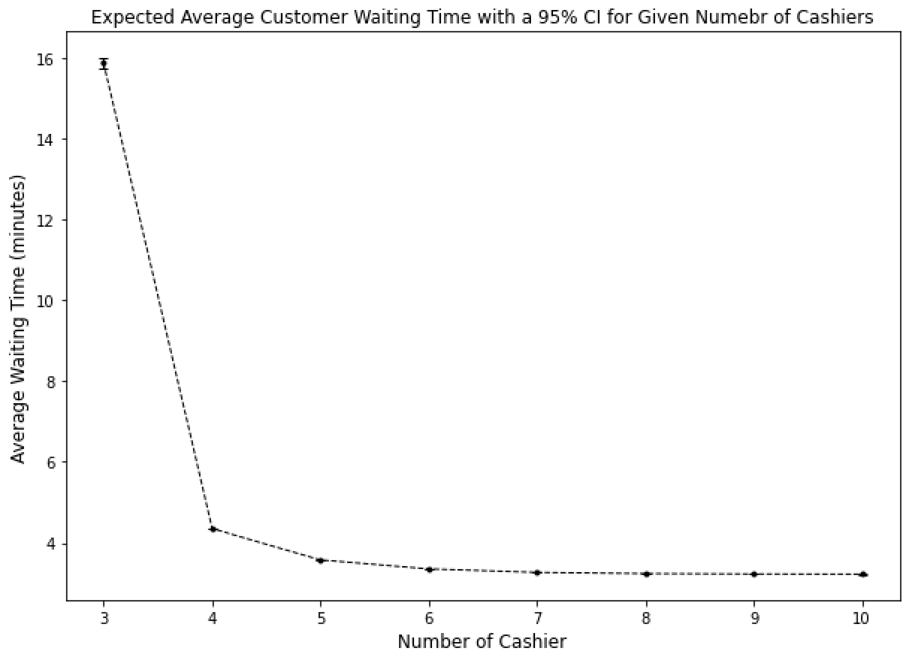

Figure 1: The expected average waiting time for a given number of cashiers are shown from the empirical simulated model. It decreases with an increasing number of cashiers and becomes less than 5 minutes for 4 or more cashiers.

I run the simulation 1000 times with different numbers of cashiers upto 10 to get more observations to calculate the confidence interval. If we have more observations, the standard error ($SD/\sqrt{n}$) becomes smaller, thus our confidence interval becomes narrower and we can be less uncertain (thus more confident) about our estimated mean.

We found that all three observables decreased with increasing numbers of cashiers. The expected average waiting time with 4 cashiers is 4.34 minutes with a 95% confidence interval of 4.33 to 4.35 minutes and it decreases even more for more cashiers (Fig 1). But a waiting time of less than 5 minutes is quite reasonable for a grocery store. Similarly, the estimated maximum queue length decreases from around 10 customers with 3 cashiers to less than 3 customers with 4 or more cashiers (Fig 2). The extra time is also less than 5 minutes for grocery store with 4 and more customers (Fig 3).

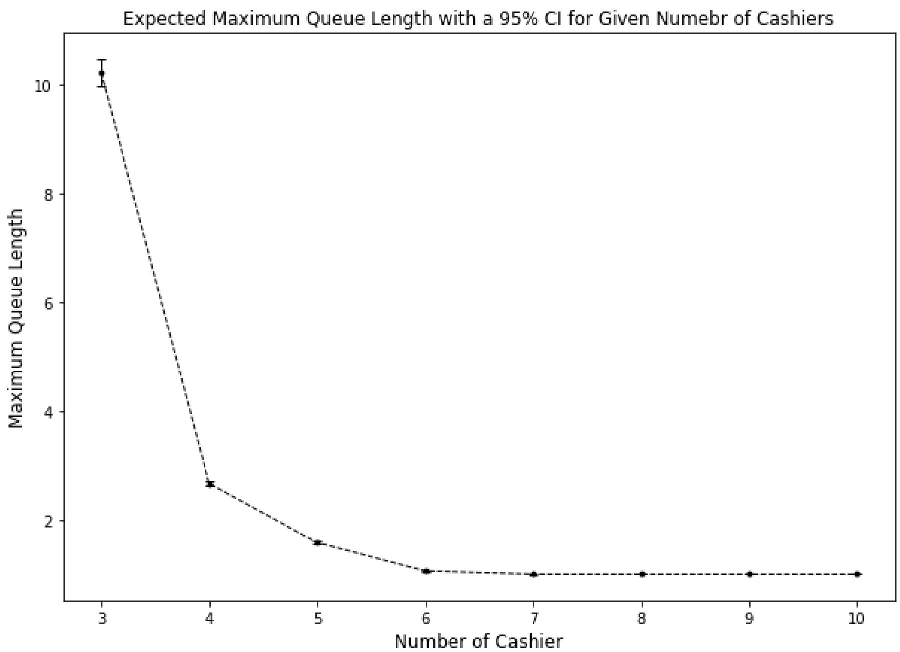

Figure 2: The expected maximum queue length for a given number of cashiers are shown from the empirical simulated model. The maximum queue length decreases with an increasing number of cashiers. The max queue length is around 3 with 4 cashiers and it stabilizes at 1 with 6 or more cashiers.

From the figures, we can see that there is a very strong improvement when we increase the cashier number to 4, but after that, although we have small improvement, that small improvement will come with a large cost as employing a new cashier and making space for a new checkout queue is costly. So based on the simulation, I suggested that the grocery should employ 4 cashiers in total.

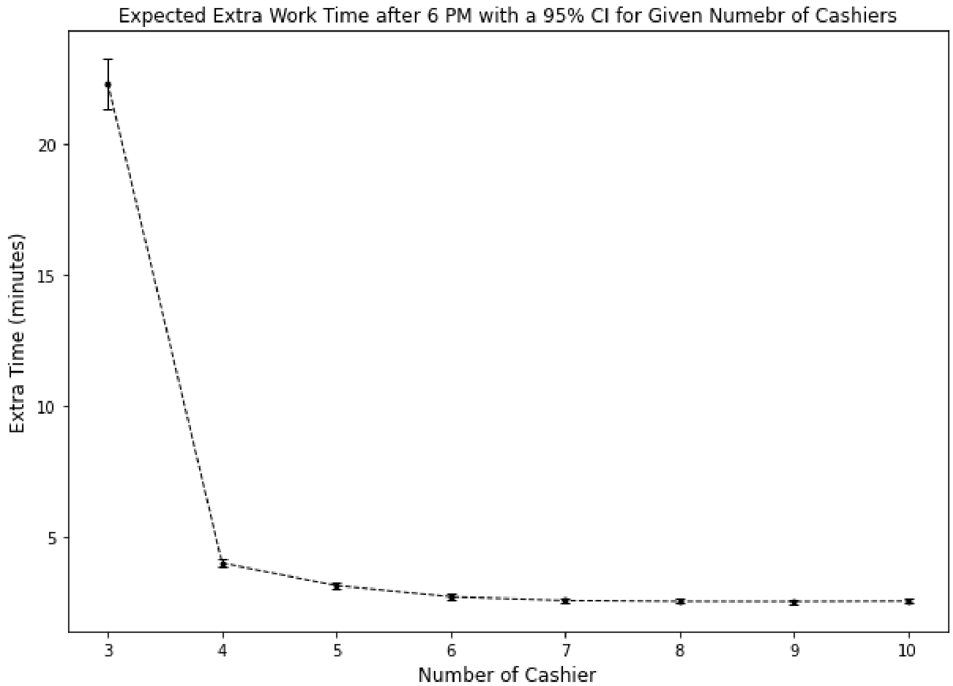

Figure 3: The expected extra times for a given number of cashiers are shown from the empirical simulated model. The extra time falls rapidly and it has a value of fewer than 5 minutes for a store with 4 or more cashiers.

Theoretical Model

Then I again estimated the value of our important observable metrics using a theoretical model to compare the results from both models. Here the arrival time is exponentially distributed which is related to a Markov model, the service distribution is a normal distribution (General model) and for each queue, only one cashier is assigned, so this is a M/G/1 model. But if we have n number of cashiers, then there will be n number of queues. Now if we assume that customers choose the queue completely in random (which contradicts our assumption that people choose the shortest queue), then each queue will also have an exponential arrival distribution with an average rate of $\lambda' = \lambda/n$, where λ is the overall arrival rate. So we can estimate the store as a $n \times M/G/1$ model. We can use the theoretical formula of M/G/1 only with the arrival rate of

λ' for each queue. To estimate the total mean service time, we need to consider that 5% of the time they need support from the manager. Now,

Average service time, $\tau = (3 \times 95\%) + (3 + 4) \times 5\% = 3.2$ *minutes/customer*

Variance of service time, $\sigma^2 = 1$ *minute/customer*

Average arrival rate, $\lambda' = \lambda/n = 1/n$ *customer/minute*

Utilization, $\rho = \lambda' \tau = 3.2/n$

Therefore, in equilibrium, the average waiting time in queue,

$$= \frac{\rho\tau}{2(1-\rho)}\left(1 + \frac{\sigma^2}{\tau^2}\right) = \frac{3.2/n \times 3.2}{2(1-3.2/n)}\left(1 + \frac{1}{3.2^2}\right) = \frac{5.62}{n-3.2}$$

The average total waiting time,

$$= \textit{waiting in queue} + \textit{service time}$$

$$= \frac{\rho\tau}{2(1-\rho)}\left(1 + \frac{\sigma^2}{\tau^2}\right) + \tau$$

$$= \frac{5.62}{n-3.2} + 3.2$$

The average queue length,

$$= \textit{arrival rate} \times \textit{waiting in queue}$$

$$= \lambda' \times \frac{\rho\tau}{2(1-\rho)}\left(1 + \frac{\sigma^2}{\tau^2}\right)$$

$$= \frac{1}{n} \times \frac{5.62}{n-3.2} = \frac{5.62}{n(n-3.2)}$$

Average total customer in the system,

$$= \textit{arrival rate} \times \textit{total waiting time}$$

$$= \lambda' \times \left(\frac{\rho\tau}{2(1-\rho)}\left(1 + \frac{\sigma^2}{\tau^2}\right) + \tau\right)$$

$$= \frac{1}{n} \times \left(\frac{\rho\tau}{2(1-\rho)}\left(1 + \frac{\sigma^2}{\tau^2}\right) + \tau\right)$$

$$= \frac{1}{n}\left(\frac{5.62}{n-3.2} + 3.2\right)$$

I couldn't estimate the maximum queue length over a day theoretically, so I compared the average theoretical queue length with the maximum empirical queue length. For estimating the extra time, I assumed that at closing time (6 PM), the system will still be in equilibrium and have a total customer equals to the average total customer calculated above. Then an average serving time is required per each customer. Thus extra time,

$$= avg\ service\ time\ \times avg\ total\ customer\ in\ system$$

$$= \tau \times \lambda' \times \left(\frac{\rho\tau}{2(1-\rho)}\left(1 + \frac{\sigma^2}{\tau^2}\right) + \tau\right)$$

$$= \frac{3.2}{n}\left(\frac{5.62}{n-3.2} + 3.2\right)$$

Comparison between Theoretical and Empirical model

When we compare the theoretical and empirical expected waiting time, the theoretical waiting time is always higher and the difference decreases with more cashiers (Fig 4). A possible cause for this is the assumptions in the theoretical model that customers choose the queue randomly, thus each queue has an equal probability to be picked irrespective of their queue length. So it is possible that some customers may end up choosing a queue where there are already 2 more people waiting, which will increase the waiting time of that customer, but that is not happening in real life or in the simulation. Then when the number of cashiers (and thus number of queue) increases, the probability of choosing a given queue decreases, thus the probability for choosing a queue where 2 people are already queueing decreases. This is why the difference is decreasing for more cashiers.

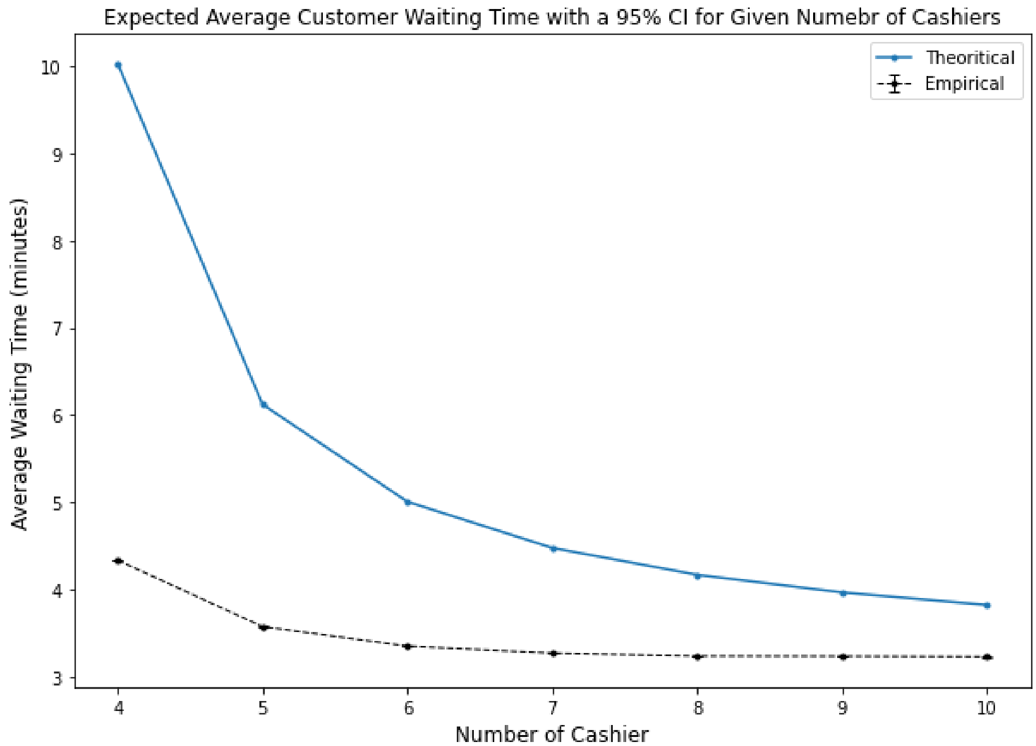

Figure 4: The expected average waiting time for a given number of cashiers are shown for both theoretical and empirical models. The theoretical model shows a higher average waiting time but the difference is decreasing with more numbers of cashiers.

It's hard to compare the average and maximum queue length as the variables that we are measuring are different for each model. But at least we can see that both of them follow a similar trend, where the maximum queue length is constantly higher than the average one (Fig 5). The empirical maximum queue length stables at 1 with 6 or more cashier**s**. When a customer arrives in a line, the queue length becomes 1, so the maximum queue length cannot be less than 1. For the expected extra time, we can see that the theoretical model shows a higher value for fewer cashiers, and later for more cashiers, it is lower than the empirical values. As discussed for waiting times, random choice of the queue might be the cause for a higher value in extra time as some customers may end up in larger queues. But when the cashiers and queue is more, the average number of customers in each queue becomes less than 1 and thus we get an extra time significantly less than the service time by multiplying service time with avg total customer. But in reality, a customer who arrives close to 6PM will choose one of the empty queues and it will take roughly the average service time to serve the customer and close the store. This is why we can see that in simulation, the extra time stabilizes around the avg service time (3 minutes).

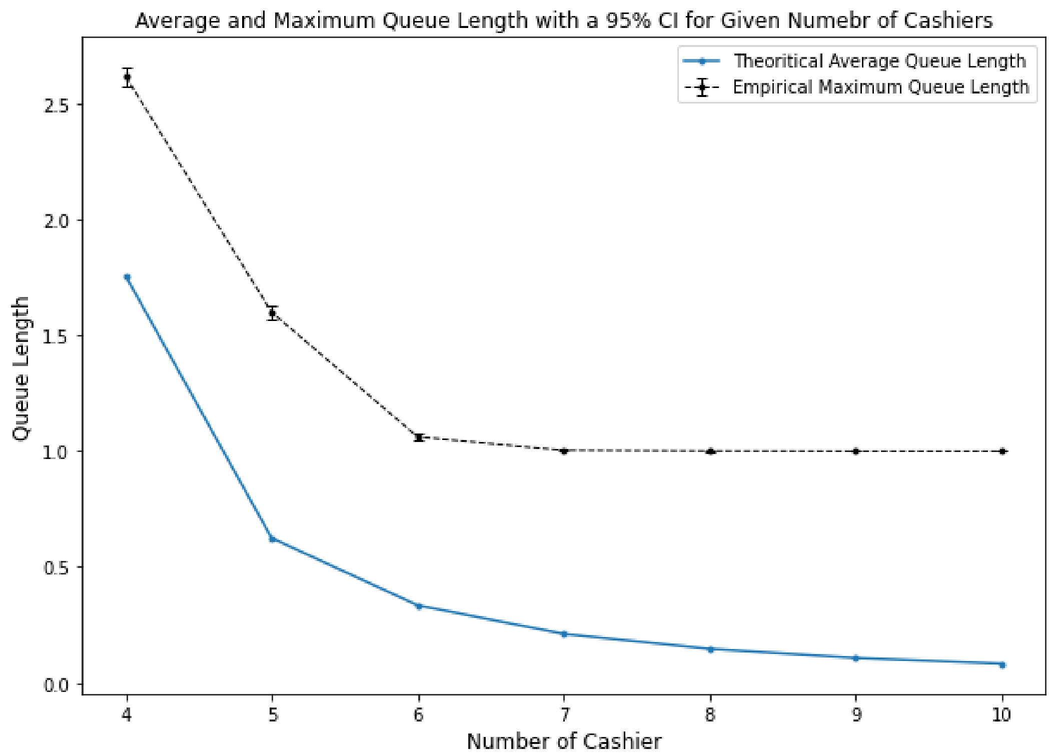

Figure 5: The empirical expected maximum queue length is compared with the theoretical average queue length. Both show a similar trend but the max queue length stabilized at 1.

As the theoretical model did not indicate any major error or unexpected behavior for the simulated empirical model, we can accept the results from the simulation and our final suggestion remains the same that the grocery store should employ 4 cashiers in total.

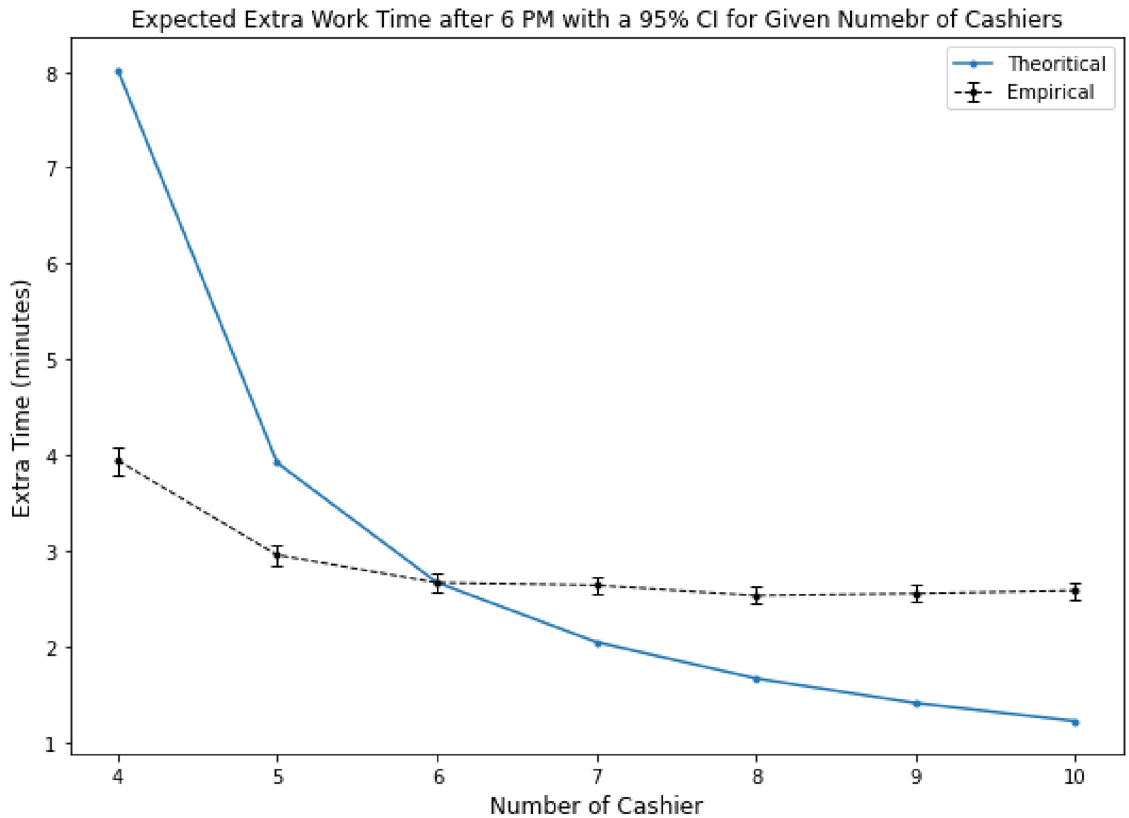

Figure 6: The expected extra times for a given number of cashiers are shown for both theoretical and empirical models. The empirical model stabilizes around 3, where the theoretical model is initially higher but later with more than 6 cashiers, it becomes lower than the empirical one.

Reflection

When deciding the number of cashiers we choose three metrics. The rationale behind choosing these metrics is mainly two: increase its profit and its brand values. As a business, the grocery store needs to analyze and implement a situation that will increase or at least not decrease their profit. But also according to their values and guiding principles, they also want to maintain certain policies for customer satisfaction. Here it might seem that employing more cashiers is a cost, but if having a limited cashier increases the waiting time or queue length, then the customer will not be satisfied with the service and they will switch to other grocery stores if they got better service there. We are assuming that customers will be driven by a set of preference and they may not prefer to stand in a long line even if the products of the store is better and cost-efficient. And we also assumed that it goes against the value and ethical consideration of the store to force its employee (the cashiers) to work more than they are assigned.