Team #4342

The problem we are presented with involves determining minimal changes to the teller system of a bank to achieve the goals set by the manager, namely to minimize the length of the waiting line and the time customers spend waiting. Although the context of our problem is narrow, the models we developed can be applied to any situation involving queues and improving operational efficiency.

In this project, we wanted to make the smallest change to the bank's teller system while satisfying management goals of having the average queue lengths be 2 people or fewer and the average wait time in queue be 2 minutes or less. To achieve these goals, we constructed a system of probabilistic differential equations that related the related the average arrival and departure rates to find the likelihood that our bank will be serving a given number of people at any time. From these probabilities, we determined the expected number of customers we will see waiting in line and the average time they will spend waiting. Using our model we determined that hiring an extra teller would decrease average queue lengths from 2.79 to 0.28 people and wait time from 6.84 to 0.69 minutes. We also accounted for factors such as bank peak hours where we would see an increased arrival rate by adding piecewise components with respect to time to our system of differential equations. Again, hiring an extra teller to serve during those peak hours decreased average queue length from 2.32 to 0.98 people and wait time from 5.68 to 2.4 minutes.

We also created a computational model of the system using stochastic processes. This computational model allows for the analysis of more complex scenarios and results in more accurate numerical data than the mathematical model. The computational model agreed with the mathematical model in that the current system does not achieve the manager's goals for queue length and wait time, and that improvements are necessary. Multiple improvements were modeled and tested, including adding a full-time and a part-time additional teller and installing an ATM. All of these changes brought the queue length and wait time within the manager's specifications, so we performed a cost/benefit analysis to choose the optimal change. We found that, although hiring a part-time teller to work peak hours only costs $8,300/year, installing an ATM will be a cheaper and better solution in the long run, with a payback period of only 3 years. Therefore, we state this solution achieves the objectives while having minimal costs.

HiMCM 2013: Problem B Solution Paper

Introduction

Background

Corporations devote billions of dollars each year to addressing issues like customer satisfaction and operational efficiency. In particular, one major issue businesses face is reducing the time their customers spend waiting in line; research has shown a direct proportion between wait time and customer frustration (Neogi). Thus, companies must find cost-effective ways to alleviate wait times to prevent dissatisfied customers from taking their business elsewhere. In this paper, we investigate several ways a bank can improve its customer service by creating and analyzing both mathematical and computational models of customer arrival rates and teller service times. The models developed in this paper are relevant in a general sense and can be applied to many different situations involving minimizing time spent waiting in queues.

Problem Restatement

At a certain bank, the bank manager has set customer experience goals that 1) the average customer waits less than two minutes for service and 2) on average, the length of the queue line is two people or fewer. Given probability distribution data on the customer interarrival times and customer service times at the bank, as well as information on the bank's daily number of patrons, we first seek to determine whether or not the current service system meets the manager's goals (Pt.1). If the current service is unsatisfactory, we seek to investigate several different measures that the bank could employ in order to assure that management goals are met (Pt.2). We also wanted to account for changes in traffic that banks can expect to see during times such as standard peak business hours.

Assumptions and Justifications

Assumption 1: Customers entering the bank will approach an open teller immediately to be serviced and if no tellers are available, they will wait in the queue.
Justification: Our model will assume that it is in the best interest of bank customers to be served in the quickest manner possible. This assumption is realistic and gives us a specific, logical behavior for customers to follow within our model.

Assumption 2: There is one queue that customers wait in to be serviced by one or more teller.
Justification: Our model follows the format for which most banks arrange their teller system. This format promotes fairness, making sure that the first-come first-serve standard is met. The idea of fairness alone may be considered better customer service.

Assumption 3: Bank managers aim to minimize the cost of operation while maintaining the desired level of service.
Justification: Banks, like every other business, aim to minimize costs to maximize profits. We take this into account when deciding the 'minimal' changes for the bank to make; since it costs money to increase the level of service (more employees, ATM's, etc), we consider and define the "best" changes to the system to be ones that minimize costs while achieving the desired service level.

Assumption 4: Once a customer has been serviced, he or she will depart**.**
Justification: We consider it in the customer's best interest to immediately depart after being serviced to attend to other matters. As a result, in our model served customers are removed from the queue once served and the departure rate is equivalent in value to the service rate.

Assumption 5: There is currently one teller at this bank**.**
Justification: We were not given the number of current tellers. Having one teller simplifies our basic model and provides a point of comparison for situations in which we vary the number of tellers employed.

Assumption 6: The bank opens at 9:00am, and remains open continuously for 6.625 hours or 397.5 minutes before closing.
Justification: 9:00 a.m. is a standard start time for a bank. Given that the bank serves approximately 150 people/day, and that the mean time between arrivals is 2.65 min (math shown in *Mathematical Model Pt. 1*), we find that the bank must be open for about 6.625 hours per day to process the given number of customers. If it were open longer, more than 150 people would be served per day based on the given probability rate.

Assumption 7: The bank is open six days a week for the same number of hours each day.
Justification: For simplicity's sake (given our time restraint), we don't consider days with shortened hours; each day is open the same number of hours (6.625). Additionally, we assume the bank is closed on Sundays, as is protocol for most banks.

Assumption 8: Customers only arrive and are processed at minute intervals.
Justification: The data is given in minute intervals. We do not have the information to make conclusive judgements about how the bank operates on shorter time intervals; therefore, while also acknowledging our time constraint, our model must run on minute intervals.

Part 1: Models

In Part One, we analyze the current system to see if it already satisfies the manager's goals. We use both a mathematical model and a computational model.

Mathematical Model

In order to evaluate the current efficacy of our banking establishment, we first determined the mean interarrival time (***X***) and mean service time (***Y***). Since the data sets for both interarrival time and service time followed a discrete probability distribution, to find ***X*** and ***Y***, we just found the expected value for each corresponding data set, where expected value ***E(Z)*** is:

$$E(Z) = \sum_{z}^{\infty} P(z) \cdot z$$

where ***E(Z)*** is the expected value for some event ***Z, z*** is an individual outcome of ***Z***, and ***P(z)*** is the probability of ***z*** occurring. Our results for ***X*** and ***Y*** are as follows:

$$X\text{: } 2.65 \pm 1.39 \text{ minutes}$$
$$Y\text{: } 2.45 \pm 1.02 \text{ minutes}$$

From ***X*** and ***Y***, we then calculated the average arrivals per minute ( ) and the average customers served per minute ( ):

$$\lambda\text{: } 0.377 \text{ arrivals per minute}$$
$$\mu\text{: } 0.408 \text{ customers per minute}$$

We then constructed a system of linear differential equations where P(S_n) is the probability that the system will be in state S_n, with *n* denoting the total number of people (sum of the people waiting in the queue plus those being served by a teller). Additionally, P_n(t) denotes the probability of being in state S_n at time *t*.

We now look at several different cases for the probabilistic model:

Case 1: S_0

Consider $P_n(t + \Delta t)$, the case of being in S_n at time $t + \Delta t$, where Δt is small enough to guarantee that there will not be multiple arrivals or departures in single Δt.

There are three different ways to reach $\boldsymbol{S_0}$(the state of having zero people) after Δt, along with their corresponding probabilities of occurring:

1) Starting at S_0, one person will arrive and one person will leave:
2) Starting at S_0, zero people will arrive and zero people will leave:
3) Starting at S_1 zero people will arrive and one person will leave:

Where is the mean rate of customer arrival and is the mean rate of service. Therefore, the total probability of being in state 0 at time $t + \Delta t$ is:

Simplifying the expression yields:

From this probabilistic model, we can then solve for the differential:

$$\lim_{\Delta t \to 0}(\frac{P_0(t+\Delta t) - P_0(t)}{\Delta t}) = \lim_{\Delta t \to 0}(-P_0(t)\lambda + P_1(t)\mu - P_1(t)\lambda\mu\Delta t)$$

$$\frac{dP_0}{dt} = -\lambda P_0(t) + \mu P_1(t)$$

Case 2: S_n, $\mathbf{n \geq 1}$

There are four different ways to reach the state of having n people after *Δt* (along with their corresponding probabilities):

1) Starting at S_n, one person will arrive and one person will leave: $P_n(t)(\lambda\Delta t)(\mu\Delta t)$
2) Starting at S_n, zero people will arrive and zero people will leave: $P_n(t)(1-\lambda\Delta t)(1-\mu\Delta t)$
3) Starting at S_{n+1} zero people will arrive and one person will leave: $P_{n+1}(t)(1-\lambda\Delta t)(\mu\Delta t)$
4) Starting at S_{n-1} one person will arrive and zero people will leave: $P_{n-1}(t)(\lambda\Delta t)(1-\mu\Delta t)$

Therefore the total probability of being in state n at time *t* + *Δt* is:

$$P_n(t+\Delta t) = P_n(t)(\lambda\Delta t)(\mu\Delta t) + P_n(t)(1-\lambda\Delta t)(1-\mu\Delta t) + P_{n+1}(t)(1-\lambda\Delta t)(\mu\Delta t) + P_{n-1}(t)(\lambda\Delta t)(1-\mu\Delta t)$$

Simplifying the expression and finding the differential yields:

$$P_n(t+\Delta t) = 2P_n(t)\lambda\mu\Delta t^2 + P_n(t) - P_n(t)\lambda\Delta t - P_n(t)\mu\Delta t + P_{n+1}(t)\mu\Delta t - P_{n+1}(t)\lambda\mu\Delta t^2 + P_{n-1}(t)\lambda\Delta t - P_{n-1}(t)\lambda\mu\Delta t^2$$

$$P_n(t+\Delta t) - P_n(t) = 2P_n(t)\lambda\mu\Delta t^2 - P_n(t)\lambda\Delta t - P_n(t)\mu\Delta t + P_{n+1}(t)\mu\Delta t - P_{n+1}(t)\lambda\mu\Delta t^2 + P_{n-1}(t)\lambda\Delta t - P_{n-1}(t)\lambda\mu\Delta t^2$$

$$\frac{P_n(t+\Delta t) - P_n(t)}{\Delta t} = 2P_n(t)\lambda\mu\Delta t - P_n(t)\lambda - P_n(t)\mu + P_{n+1}(t)\mu - P_{n+1}(t)\lambda\mu\Delta t + P_{n-1}(t)\lambda - P_{n-1}(t)\lambda\mu\Delta t$$

$$\lim_{\Delta t \to 0}\frac{P_n(t+\Delta t) - P_n(t)}{\Delta t} = \lim_{\Delta t \to 0} 2P_n(t)\lambda\mu\Delta t - P_n(t)\lambda - P_n(t)\mu + P_{n+1}(t)\mu - P_{n+1}(t)\lambda\mu\Delta t + P_{n-1}(t)\lambda - P_{n-1}(t)\lambda\mu\Delta t$$

$$\frac{dP_n}{dt} = -P_{n-1}(t)\lambda - P_n(t)(\lambda+\mu) + P_{n+1}(t)\mu$$

Considering the first-order system of differential equations with constant coefficients constructed by the differential equations from Cases 1 and 2, a matrix can be used to express the system.

$$\frac{dP}{dt} = Ax \qquad (1)$$
$$x(0) = C$$

Here $x(t)$ is a column vector that describes the probabilities of being at state S_n, $x: R \to R^n$, and *a* is an $n \times n$matrix. Using an exponential of the matrix, a solution can be obtained for the system of linear first-order ordinary differential equations using a power series.

$$e^{t \cdot A} = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

We can implement a finite approximation of the exponential of the matrix numerically using MathCad, a mathematical computing software. Let *A* be a $m \times m$ matrix. Then the exponential, e^{tA}, can be approximated after counting *K* terms in the series, where *K* represents a finite limit to the queue length. This approximation can be used to approximate the solution to equation (1).

$$e^{t \cdot A} \approx \sum_{n=0}^{K} \frac{t^n}{n!} A^n$$

$$x(t) = C \cdot e^{tA}$$

The eigenvalues of matrix *A* are important in determining exact exponential and understanding the behavior of the probabilistic model. The eigenvalues are the solutions for the following equation:

$$A \cdot v = \Lambda \cdot v$$

$$(A - \Lambda \cdot I) \cdot v = 0$$

Above, Λ is a real eigenvalue for matrix *A*, *v* is a non-zero eigenvector for Λ, and *I* is an $m \times m$ identity matrix. The above equations imply

$$\det(A - \Lambda \cdot I) = 0,$$

which is a polynomial in Λ, or the characteristic polynomial of the matrix *A*. The roots of the polynomial are known as generalized eigenvalues for matrix *A*. In this case, the generalized eigenvalues are real and distinct. Given the above to be true, the probability for being in any given state can calculated given a prior distribution.

To find the average probability that the system will be in a given state over $a \leq t \leq b$,we performed the following function.

$$\boldsymbol{P_{n,avg}} = \left(\int_a^b \boldsymbol{P_n(t)dt}\right) \div (b - a)$$

Mean queue length was computed using the formula for expected values. *n* denotes all possible values for the state of our system and $\boldsymbol{P_{n,avg}}$ is given above

$$\sum_{n=0}^{9} n \times \boldsymbol{P_{n,avg}}$$

After setting up this system of differential equations, we solved for the $P_n(t)$ for $0 \leq n \leq 9$ as $t \to \infty$ in MathCad, a mathematical computing software, and obtained the following average probability distributions, $P_{n,avg}$. Q_n is the number of people waiting in the queue, which is obtained by subtracting the number of people being serviced from S_n , the total number of people in our system:

S_n	**Q_n**	**$P_{n,avg}$**

S_0	0	0.081
S_1	0	0.0139
S_2	1	0.124
S_3	2	0.111
S_4	3	0.099
S_5	4	0.089
S_6	5	0.081
S_7	6	0.073
S_8	7	0.067
S_9	8	0.061

Table 1. Probabilities $P_{n,avg}(t)$ that our one teller system will be in a state S_n with *n* customers at any given time *t.* Q_n denotes number of customers in queue since S_n includes the one customer who is currently being served.

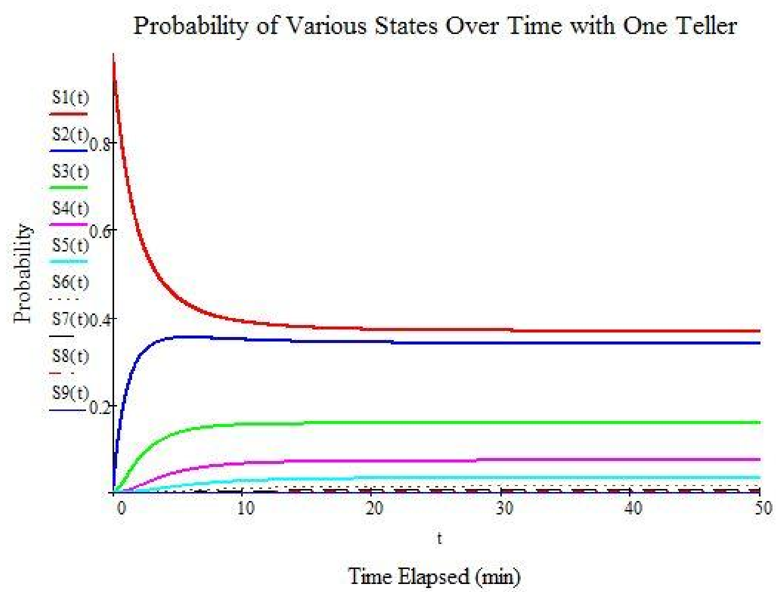


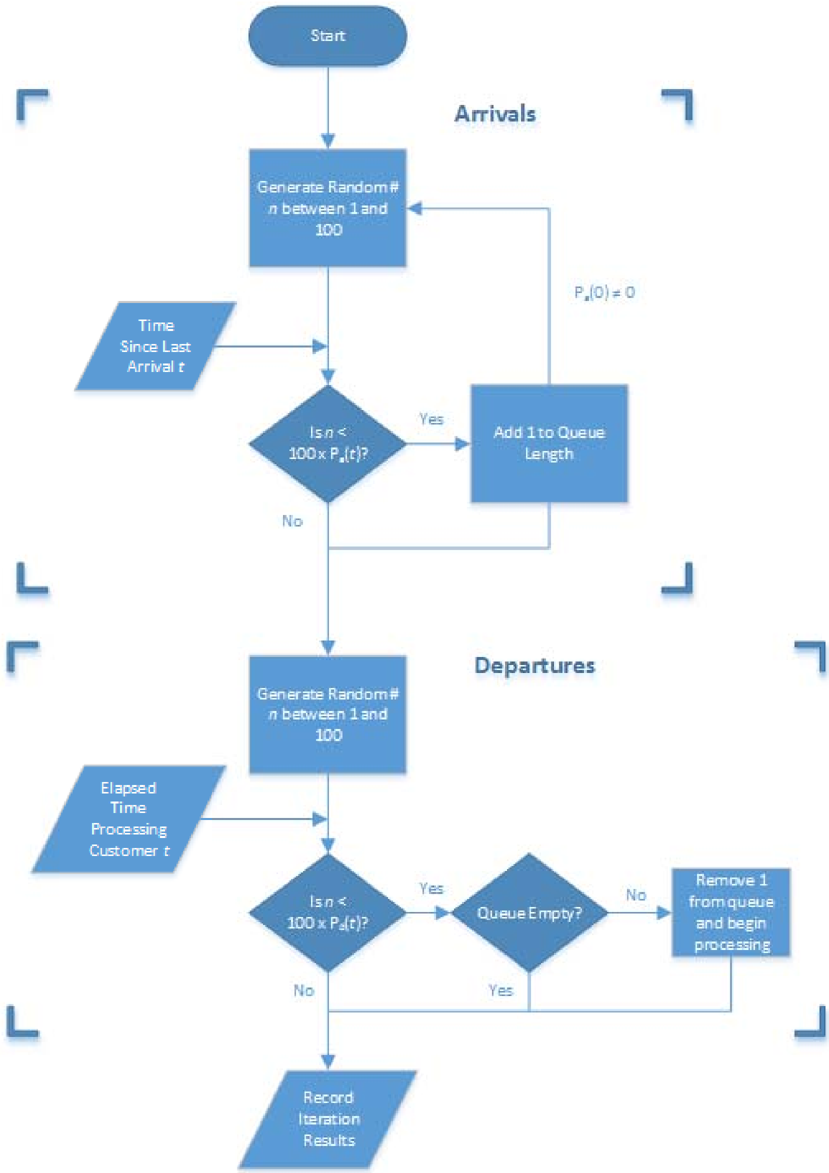
Figure 1. Probability distribution of S_n occurring one teller system. X-axis showing time elapsed is truncated at 50 minutes since probabilities for S_n approach a fixed value.

We bounded *n* at 9 since previous studies have shown that customers will begin to balk and leave once they see that a queue contains 10 or more people (Dhar and Rhaman, 2013). From this data, we calculated the average queue length to be 2.79 people (by finding the expected value), from which we can find the average wait time by multiplying the queue length by the average service rate, μ, which results in an total average wait time of 6.84 minutes. Thus, the current system fails to satisfy the manager's goals, as the average queue length is larger than the desired average of two and the average wait time is greater than the desired two minute wait time. Clearly, teller changes are necessary to reach satisfaction guidelines.

It should be noted that the mathematical model is not entirely numerically accurate. Since the mathematical model does not take into account the given probability functions of interarrival time and service time, but only considers the mean of these probability tables to calculate λ and μ, some inaccuracy results. When constructing the mathematical function, by using only the λ and μ values, we made the assumption that the probability functions were of an exponential distribution. An exponential distribution is the type of distribution that is used to model interarrival times - when collecting data in the field, an exponential distribution accurately models the interarrival times for random events such as people entering a building or telephone call rate - anything that occurs through the Poisson process. Therefore if we were not given a probability distribution, we would assume it to be an exponential probability distribution. However, when we inspect the probability distribution we find that it is not in fact an exponential distribution, but is in fact closer to a normal distribution in shape. This is strange, because in addition to the fact that we would expect this process to follow an exponential distribution, a Gaussian distribution does not have the property of "memorylessness" like an exponential distribution does, and which is necessary to accurately model a function of the type which we are modeling. Due to these factors, it would be very difficult to model this system accurately using a closed-form process - the model is not unlike a G/G/1 queuing system, which is very difficult to solve accurately in a mathematical model. While the general pattern and shape of the graphs and data shown is extremely useful for analysis of the system, the numerical results may have some slight error. Therefore, to collect accurate numerical data, a different model type is necessary which can take the given probability distribution into account. We find that a computational model is the perfect type of model to numerically solve this system.

Computational Model

Due to the limitations of our mathematical model, namely the inaccuracies present due to the inaccurate modelling of the rate of arrival and service and the difficulty of modelling various complex scenarios with such a closed-form model, we decided that it would be useful to build a complementary computational model. This computational model is stochastic and models the arrival and processing of people randomly based on the given probabilities. Due to these properties, the computational model returns slightly different results each time that it is run.



Flowchart of basic one-teller computational model

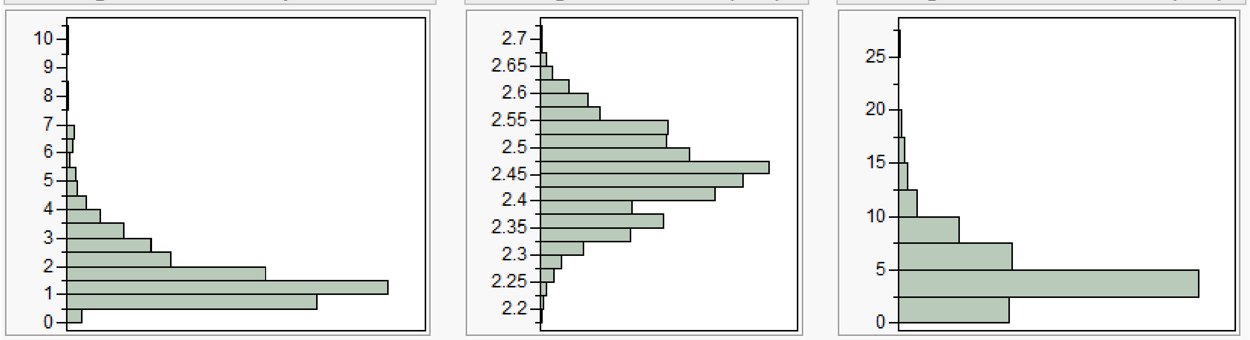
To create a computational model of the system, we must first consider the most simple case before extending the model to more complex situations. The most simple case is that which is presented in the problem statement, namely a constant one-teller system. In this system, people enter the system at a rate λ with a probability distribution of time between arrivals as given in the problem statement. Similarly, people leave the system at a rate μ with the probability distribution as given. To model this system computationally, we will evaluate the number of people entering and exiting the system at discrete time intervals of 1 minute. We will then iterate over these one minute intervals to determine the behavior of the system.

We will model the entry of people into the system by considering the time since the last arrival *t*. Since we wish to know the probability that a person will arrive given a time since last arrival *t*, we must find the probability P(t) based on the given probabilities of time between arrival times. To do this, we consider the probability based on the remaining probabilities - for example P(3) = .35/(.35+.25+.05) = 54%. It should be noted that at the last value of *t* given (*t* = 5), the result of this calculation is 100% probability, because the probability of a time between arrival at *t* > 5 is 0. To evaluate stochastically whether or not a person will arrive based on the calculated probability P(t), we generate a random number *n* between 1 and 100 and test whether n < P(t) x 100. If the test returns positive, the queue length is iterated by one. Additionally, since P(0) > 0 (the probability of two people walking in simultaneously), we must generate a new random number and test if *n* < P(0) x 100, repeatedly iterating the queue until the test fails.

The model for the processing of people is very similar to that of their entry. We first create a new table of P(t) for each *t* based on the given probabilities between arrival, and then, based on the elapsed time *t* that a person has been processed for finds the probability that the person will finish being processed in the current iteration of the program. The same random number test is used to determine stochastically whether or not the person leaves the system. Since we have stated the assumption that when a person finishes being processed they will immediately leave the system, the rate of processing is equal to the rate of exit from the system.

To evaluate the current system with our stochastic computational model, we programmed the model to run one minute at a time - until the number of people that have been processed that day equals 150. Since the model is stochastic, it will return a slightly different result each time it is run (each day that we simulate). Therefore, we run the simulation multiple times, and analyze the distributions of data generated. For our simulations, we ran 1000 simulations for each scenario (1000 discrete days simulated). We recorded several parameters from each run, namely the average number of people in the queue, the average service time, and the average wait time in the queue. We then plotted the distributions and compared the means for each of these parameters in order to compare the different scenarios and to evaluate if the manager's goals are met.

For a bank with only one teller, we saw that there was an average of 1.83 people in the queue, a average service time of 2.45 min, and an average wait time of 4.91 min in the queue. The results for wait time and service time are close to the mathematical calculations we obtained previously. Note, however, that the average queue length derived from our system of differential equations slightly exceeds the value given by the computational model. This is likely due to how we were unable to account for the probability distributions in our mathematical model while the computational simulation reflected those variances by using the given probabilities to generate queue lengths.



Average Number of People in Queue

Quantiles

Summary Statistics

Mean	1.8326975
Std Dev	1.2471949
Std Err Mean	0.0394398
Upper 95% Mean	1.9100918
Lower 95% Mean	1.7553032
N	1000

Average Service Time (min)

Quantiles

Summary Statistics

Mean	2.4496333
Std Dev	0.0808538
Std Err Mean	0.0025568
Upper 95% Mean	2.4546507
Lower 95% Mean	2.444616
N	1000

Average Wait Time in Queue (min)

Quantiles

Summary Statistics

Mean	4.90528
Std Dev	3.2316515
Std Err Mean	0.1021938
Upper 95% Mean	5.1058191
Lower 95% Mean	4.7047409
N	1000

Figure 3. Simulation results for banking system with one teller with a constant arrival rate. Charts show frequency distribution of average number of people in system, average service time, and average wait time in queue for 1000 runs with each run simulating an average day.

Peak Hours

In making a realistic model, we also wanted to consider peak hours: hours at which the bank are busiest during the day. We thus produced alternate variable-rate models for both the mathematical and computational models that take into consideration a bank's peak hours. Normal peak times at banks typically occur for a two-hour period from 11:00am-1:00pm (120 - 240 minutes) (White, 2012).

For the mathematical model, although most likely the rate of entry would be some smooth curve with a peak from 11:00am-1:00pm, for the purposes of the model we considered a discrete rate change from off-peak to peak times for simplicity. The increase in rate of arrival during peak hours was chosen to be 50% compared to normal business hours. This was reflected in the increased . We defined a peak time to have a 50% high entry rate λ than off-peak times. In incorporating this change into our model, we first found the mean and standard deviation of the entry rate, and then calculated the mean of the new rates based on the equation below, where *T* represents time and λ the entry rate.

Since we defined , we can solve for the new λ's based on the original λ. We accounted for this change by modifying our system of differential equations for one teller to be piecewise with different values of (peak or off-peak) based on whether or not *t* is within peak hours. To reflect the use of multiple tellers (see Part Two of paper) during peak hours, we changed the piecewise component describing $P_n(t)$ to multiple teller model described in case 3 along with selecting the peak value when *t* is within peak hours. was also calculated with the peak hour time periods as the upper and lower integral bounds.

For our initial or prior probability distribution, $x(0) = C$, in the peak hour component of our model we used the posterior probability distributions obtained from the one teller and λ=$\lambda_{off-peak}$model. For the post peak hour piece, we used the posterior probability distributions obtained from having $m$tellers during peak hours and λ=λ_{peak}component.

Below is a table showing data collected from running the alternate mathematical model. From the following table, we calculate the average queue length to be 2.32 people with an average wait time of 5.68 minutes. Again, this exceeds the limits set by management's satisfaction guidelines and changes are necessary.

S_n	Q_n	$P_{n,avg}$, 0 ≤ t < 120	$P_{n,avg}$; 120 ≤ t ≤ 240	$P_{n,avg}$; 240 < t ≤ 397.5
S_0	0	0.045	0.084	0.057
S_1	0	0.201	0.05	0.129
S_2	1	0.149	0.057	0.107
S_3	2	0.111	0.063	0.09
S_4	3	0.082	0.069	0.077
S_5	4	0.062	0.075	0.066
S_6	5	0.045	0.084	0.057
S_7	6	0.034	0.095	0.05
S_8	7	0.026	0.109	0.043
S_9	8	0.02	0.129	0.036

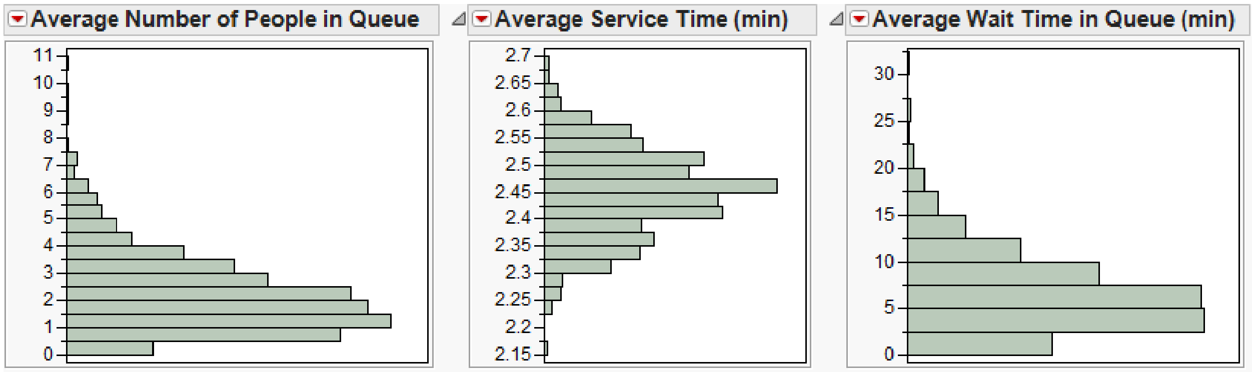
Table 2. Probabilities $P_{n,avg}$ that our one teller system with peak hours be in a state S_n with *n* customers at any given time *t*. Q_n denotes number of customers in queue since S_n includes the one customers currently being served. Peak time interval is from 120-240 minutes. $P_{n,avg}$ is piecewise with respect to peak hour time intervals.

The computational model that accounted for peak hours used several of the same guidelines as the mathematical model; the peak hours are from 11:00 a.m. -1:00 p.m., a peak time has a 50% higher entry rate λ than off-peak times, and for simplicity reasons, there is a discrete rate change from off-peak to peak times. We also used the same equation:

$$\lambda_{original} \times T = T_{peak} \times \lambda_{peak} + T_{off-peak} \times \lambda_{off-peak}$$

Since we defined $\lambda_{peak} \equiv 1.5 \times \lambda_{off-peak}$, we can solve for the new λ's based on the original λ. We then create new probability functions for each new λ (peak and off-peak), and add a function in the code to implement the proper probability function at the proper time (peak times being 11:00am-1:00pm).

Below are the simulation results. The average queue length is seen to be 2.30 people and the average wait time is 6.87 minutes. These results align with the conclusion drawn from our mathematical model that one teller is not enough to provide adequate service.



Quantiles

Summary Statistics

	Average Number of People in Queue	Average Service Time (min)	Average Wait Time in Queue (min)
Mean	2.2894202	2.4473933	6.8723267
Std Dev	1.4690164	0.0826448	4.2071506
Std Err Mean	0.0464544	0.0026135	0.1330418
Upper 95% Mean	2.3805795	2.4525218	7.1334001
Lower 95% Mean	2.1982608	2.4422648	6.6112533
N	1000	1000	1000

Figure 5. Simulation results for banking system with one teller with peak hours from 11 am - 1 am. Charts show frequency distribution of average number of people in queue, average service time, and average wait time in queue for 1000 runs with each run simulating an average day.

Part 2: Extensions

In the second part of our solution, we look at ways to improve the current system so that that the manager's goals are satisfied.

In general, there are two ways we can increase productivity. We can add units of service (in this case, additional bank tellers or ATM's), or we can increase labor efficiency.

Additional Units of Service

Multiple Tellers

To modify our mathematical model to account for multiple tellers, at one time, we modified our system of probabilistic differential equations such that when the system is at S_1, there will only be one teller working (rate of service would be), and when the system is at S_2, there will be two tellers working (rate of service would be), so that the rate of service will be for . The rate of service is for all S_n where n so that all tellers are assisting a customer.

Probability distributions are shown in the following table and graph. Our average queue length is then 0.28 people which gives us an total average wait time of 0.69 minutes thus meeting our satisfaction guidelines.

S_n	Q_n	$P_{n,avg}$
S_0	0	0.00697
S_1	0	0.34
S_2	0	0.156
S_3	1	0.072
S_4	2	0.033
S_5	3	0.015
S_6	4	0.00697
S_7	5	0.003205
S_8	6	0.001475
S_9	7	0.006795

Table 2. Probabilities $P_{n,avg}$ that our two teller system will be in a state S_n with *n* customers at any given time *t.* Q_n denotes number of customers in queue since S_n includes the two customers currently being served.

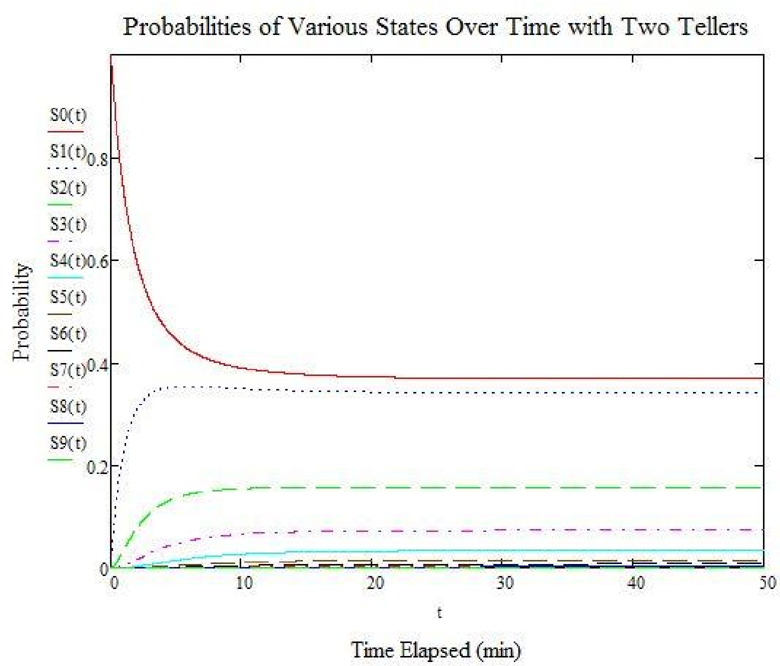


Figure 2. Probability distribution of S_n occurring with two tellers. X-axis showing time elapsed is truncated at 50 minutes since probabilities for S_n approach a fixed value.

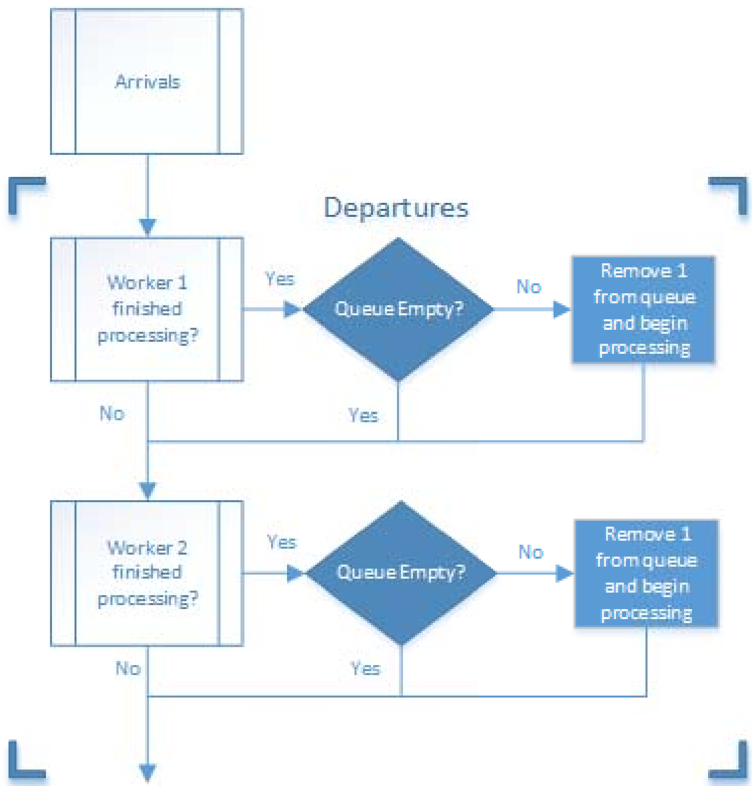
Another simple solution to satisfying the manager's goals would be to hire an extra teller *only* during peak hours to expedite service. This would result in shorter queues with an expected length of just 0.98 people and an average waiting time of 2.40 minutes. Although the waiting time is slightly longer than desired, the queue length is twice as short as management goals.

S_n	Q_n	$P_{n,avg}$, 0 ≤ t < 120	$P_{n,avg}$; 240 < t ≤ 397.5	Q_n ; 120 ≤ t ≤ 240	$P_{n,avg}$; 120 ≤ t ≤ 240
S_0	0	0.045	0.024	0	0.021
S_1	0	0.201	0.071	0	0.0228
S_2	1	0.149	0.059	0	0.14
S_3	2	0.111	0.048	1	0.086
S_4	3	0.082	0.038	2	0.053
S_5	4	0.062	0.031	3	0.033
S_6	5	0.045	0.024	4	0.021
S_7	6	0.034	0.019	5	0.013
S_8	7	0.026	0.015	6	0.008324
S_9	8	0.02	0.012	7	0.005194

Table 3. Probabilities $P_{n,avg}$that our two teller system with peak hours be in a state S_n with *n* customers at any given time *t.* $P_{n,avg}$ is piecewise with respect to peak hour time intervals. Q_n is also piecewise with respect to peak time intervals since the number of tellers and consequently number of customers being serviced varies depending on the time interval.Peak time interval is from 120-240 minutes.

To add an additional teller to our computational model, we simply created a duplicate worker that processes the queue in parallel with the first worker. This second worker has the same properties as the first worker, and follows the same probability function. Therefore, the probability distribution of service time remains unchanged with the addition of a second teller.

If we consider the utilization factors (λ/μ) in the scenario with one worker at peak and off-peak times we note that, although the utilization factor at peak time is significantly greater than 1 (~1.2), the utilization at off-peak times is less than 1 (.75). Therefore, while the single worker can handle the off-peak condition, the queue expands rapidly in the peak time when the worker can't keep up with arrivals. Therefore, it is logical to consider adding a second worker during peak times and having only one teller at off-peak times. We will incorporate the scenario in the computational model by starting the day with one worker, then enabling the second worker during peak times, and then disabling the worker once peak times are over. By having two tellers at peak times, $\mu_{combined}$= 2μ, which halves the utilization factor (to about ~.6), allowing the two tellers to easily keep up with the increased rate of arrival.



Two-worker flowchart, highlighting departure block

Figure 4 displays results from adding an additional *full-time* teller. The change decreases the average number of people in queue to 0.04 and the average wait time to 0.07 minutes.

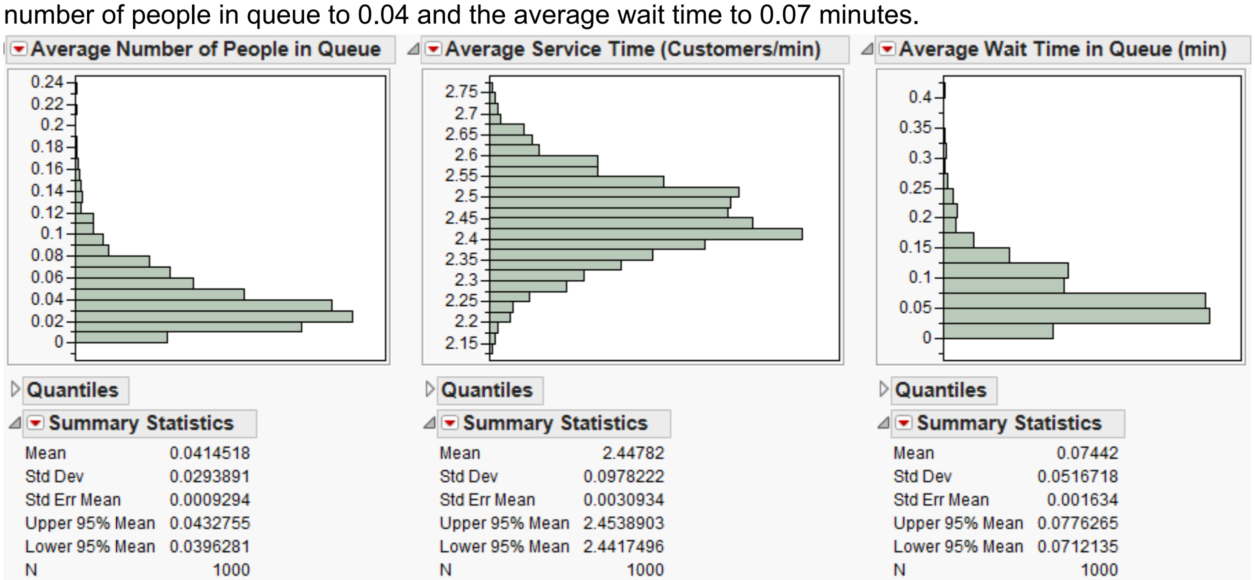


Figure 4. Simulation results for banking system with two full-time tellers with a constant arrival rate. Charts show frequency distribution of average number of people in queue, average service time, and average wait time in queue for 1000 runs with each run simulating an average day. Note changes in mean from one teller system.

As in our mathematical model, we also modeled hiring an additional *part-time* teller to serve during peak hours only (11 a.m. to 1 p.m.) This resulted in average queue lengths of 0.10 people and average wait times of 0.27 minutes thus solving our problem and satisfying our customers.

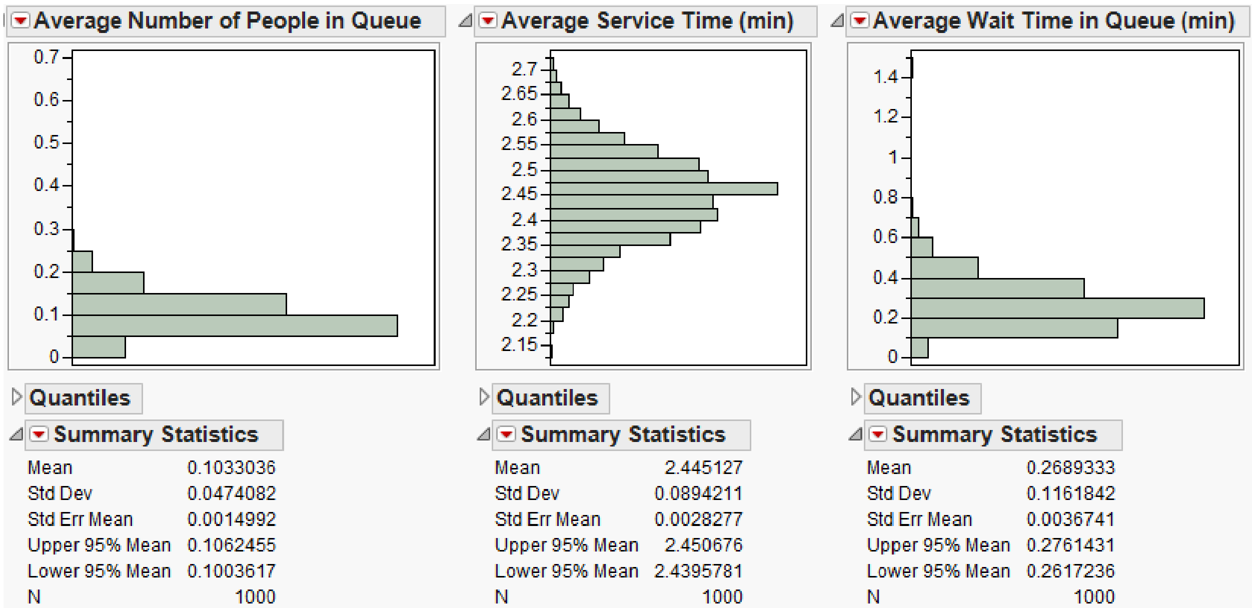


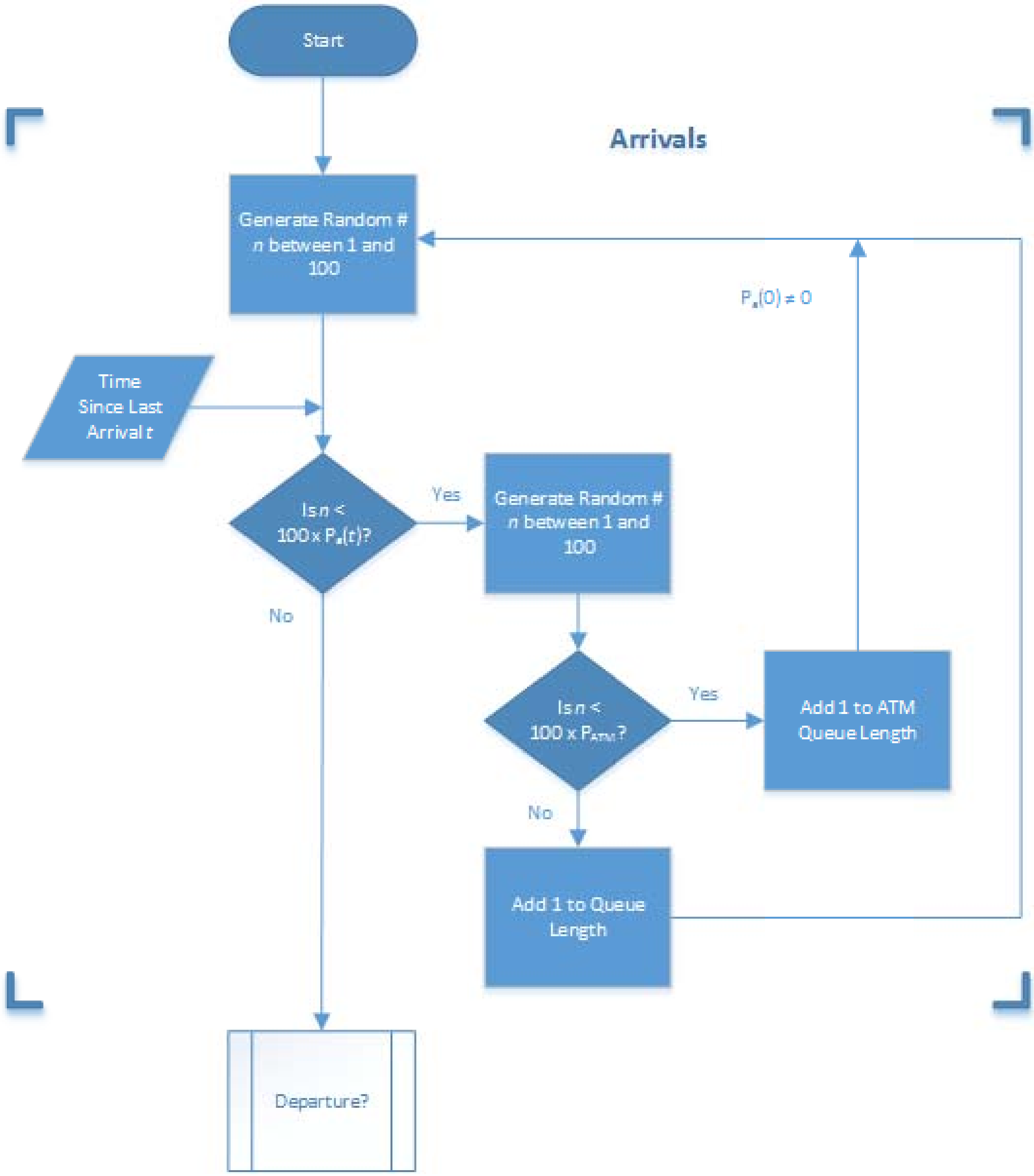
Figure 6. Simulation results for banking system with one extra part time teller (two total) at peak hours from 11 am - 1 pm. Charts show frequency distribution of average number of people in queue, average service time, and average wait time in queue for 1000 runs with each run simulating an average day.

ATM Machine

Another way that a bank can reduce the load on the tellers is to add an ATM. An ATM can be placed outside a bank and allow some of the people who would otherwise need to talk to the teller to instead be able to perform their transaction at the ATM. Since this causes some people to not go into the teller queue, the load on the teller is reduced and the queue length and wait time are also reduced.

To model this scenario we can consider that some fixed percent (ex. 25%) of people arriving to the bank are able to perform their transaction at the ATM and are willing to do so. We define a behavior that states that, if the ATM queue is shorter than the teller queue, the 25% of people able to use the ATM will chose to do so. Otherwise, they will talk to the teller. The ATM queue is processed at a fixed rate of 1 customer per 3 minutes for simplicity.

This system was only modeled computationally since a mathematical system of differential equations is difficult to create for the closed form expression used to represent ATM usage.



Flowchart of system with ATM, highlighting arrival block

Running this model with peak hours (we chose to only run it on peak hours since we felt peak hours represents a more realistic scenario) gave us an average queue length of 0.34 people and a mean wait time of 1.37 minutes.

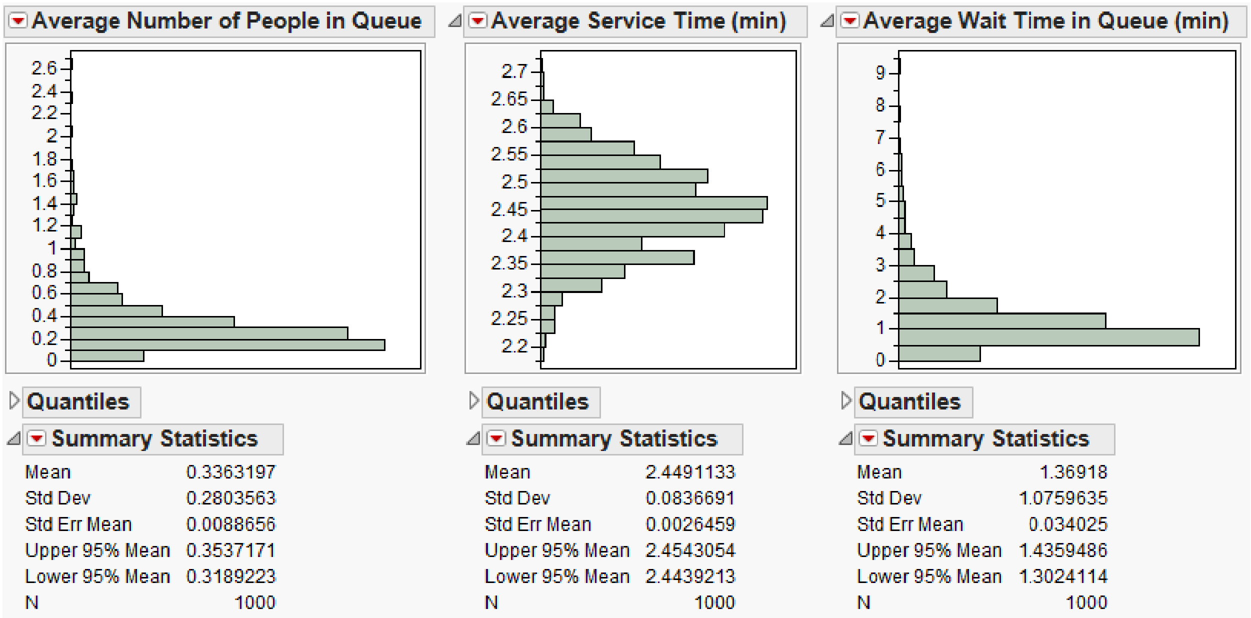


Figure 7. Simulation results for banking system with one ATM and one teller with peak hours from 11 am - 1 pm. Charts show frequency distribution of average number of people in queue, average service time, and average wait time in queue for 1000 runs with each run simulating an average day.

Increase Labor Efficiency

In order to increase the efficiency of the tellers, we considered modeling several options such as improving access to reference materials, improving the computer skills of tellers, giving the tellers access to more advanced technology, improving the mental math skills of tellers, and improving the customer communication skills of tellers.

However, because we lack information on the current efficiency state of the tellers, it was hard to perform a conclusive analysis. In addition, different work habits and learning styles mean every bank teller has to find his or her own way of becoming the most efficient at his or her job.

For these reasons, our study on the impact of improving labor efficiency on queue lengths and wait times is inconclusive. In the future, with access to animate resources (to find more and better options to improve teller efficiency), as well as more data on the current efficiency state of the tellers at the bank in question (for example, has the bank already done things to improve efficiency? How advanced is the software the tellers use? What are the education levels of the tellers?), we could analyze this further.

Data Analysis

The customer service goals as defined by the bank manager were to have the average queue length be shorter than two people and the average wait time for service (time in queue) to be less than two minutes. When we evaluate the bank as it is currently, we consider our two models for this condition - with and without peak hours. Without considering peak hours, we find an average of 1.8 people in the queue with a wait time of nearly 5 minutes. When we consider peak hours, this value increases to 2.3 people in the queue and a wait time of nearly 7 minutes. Although these two scenarios differ slightly (as expected), neither falls within the manager's goals, and therefore some change is necessary.

Our proposed improvements involve hiring more tellers or adding an ATM. By hiring a second teller for the entire day, we mathematically found that we can decrease the average number of people in the queue to 0.28 and the wait time to 0.69 minutes. Although this is a huge improvement, it may be seen as overkill that comes at an unnecessarily high cost. A similar, but much less expensive solution would be to hire a part-time teller to only work during peak hours. In this simulation, we find that the average queue length is 0.98 person and the mean time in queue is 2.4 minutes. This solution is adequate for achieving the goals of the bank manager.

A different approach to solving the problem involves adding an ATM instead of additional tellers. In this solution, the ATM reduces the load on the teller, resulting in a mean queue length of 0.34 people and a wait time of 1.37 minutes when peak hours are considered. This solution also addresses the goals of the manager.

Cost/Benefit Analysis

Since we have generated multiple solutions, each of which achieve the goals of the manager, we must consider which would be the best solution for the bank to employ. We will define "best" to be the solution that minimizes cost to the business while still achieving the goals of the manager.

First we must evaluate the cost of hiring a new bank teller. A bank teller makes on average $12.40/hour (Occupational Employment and Wages 2013). To hire a full-time teller for this bank (based on the 6.625 hour average time open per day stated in Assumption 6 and a 6-day work-week) would cost $25,630/year. Additionally, hiring a full-time employee typically requires some compensation in the form of insurance or benefits, possibly increasing this value. However, this is relatively close to the average mean annual wage for a bank teller of $25,790, so we will ignore insurance and benefits for now. If, instead of hiring an additional full-time teller, we hired a part-time teller for just the peak hours (2 hours/day and slightly more on the weekend) would cost approximately $8,300 - a significant cost reduction. ATM's, on the other hand, have a high initial cost yet have no hourly cost. Therefore, an ATM installation, while an expensive initial investment, would save significant amounts of money in the long run. For example, a typical ATM costs around $25,000 (Welch), and therefore is equivalent to three years of a part-time worker's salary. However, after the three-year payback period, the bank would be saving $8,300 a year. When compared to hiring an additional full-time teller, the ATM cost is about the same as the yearly salary of the teller, and would be saving the bank $25,000 after the first year.

The ATM carries additional benefits that are not considered in our model. For example, an ATM is open 24/7, whereas our bank is only open for about 7 hours a day, 6 days a week. Therefore, installing an ATM quadruples (168 vs 42 hours) the time which our bank can perform transactions for customers. Since some current customers may use the ATM at times when the bank is not open, the load on the teller would be further reduced beyond what we consider in the computational model as there would be less people entering the bank during its hours of operation. Additionally, some people that may have difficulty going to the bank when it is open due to work schedules or other factors could now use our bank, and our bank would gain additional customers and profits.

Sensitivity Analysis

To test the sensitivity of our result to changes in the input parameters, we performed a sensitivity analysis on our mathematical model. For each experiment, we changed one variable (eg. λ) by a small percent while holding the other variables constant. We then measured the change in the output. We did not perform this analysis on the computational model as the simulations take non-negligible time to run, and because the results should very closely mirror the mathematical solution.

Independent Variable % Change	Dependent Variable % Change
+10% λ	+24.7% Queue Length
-10% λ	-24.5% Queue Length
+10% μ	-21.4% Queue Length
-10% μ	+26.4% Queue Length

While we note that a 10% change in an input variable creates a larger (~25%) change in the output, all the changes were relatively consistent and there were no abnormal results. This suggests that, while our model can be affected by errors in input parameters, the error will propagate in an expected and repeatable fashion, allowing it to easily be discovered and removed instead of propagating in a strange and inexplicable fashion which would suggest an error in our model.

Strengths

- **Our model is highly adaptable to the processing speed of the teller and arrival rate of consumers:** The only two inputs necessary to solve the system of probabilistic differential equations that provide our average queue lengths and average wait times are the rates at which tellers can service customers (μ) and the rate at which customers arrive (λ). Thus our model is easily adaptable to situations with different rates.
- **Since our model utilizes a probabilistic stochastic process, it accounts for some randomness:** Since we compute the average queue length and wait time by taking into account the probability with which we expect to see various numbers of customers in the bank, we account for some variance and randomness in our model.
- **Computational and mathematical component of our model is adaptable for variety of different changes to system:** We evaluate the results of various changes to our system both computationally and mathematically to strengthen the conclusions we draw. Also having both a computational and mathematical component of our model allows us to broaden the range of changes we can consider without overcomplicating our model or making unnecessary assumptions. For example, implementing the effects of adding an ATM was best done computationally since modeling ATM usage mathematically would have required significant additional piecewise components.

Weaknesses

- **Our differential equations do not take into account probability distributions of λ and μ:** λ and μ were calculated using mean interarrival and service times which have associated deviations. However, in our system of probabilistic differential equations, we could not generate ranges of values for probability distributions or conduct statistical analysis that reflected the uncertainties inherent in these values.
- **Our waiting times and queue lengths were estimated in the mathematical model:** Since waiting times and queue lengths were derived from expected values, there are possible outcomes in which those times and lengths could be larger than what was set in satisfaction guidelines. However, we were able to check this with the computational model and verify the values generated, leading to much stronger conclusions.
- **Our mathematical model only calculates the probability distributions for a maximum of 10 people in the system:** Determining $P_n(t)$ for various values of *n* required solving a system of differential equations in an *n x n* matrix. Since our processor power was limited we could not solve a matrix of infinite size (or at least up to 150). As *n* became larger than 10, $P_n(t)$ returned probability values that were small enough to be unlikely to occur in our system that only serves 150 customers/day so we bound n from 0 to 9. Again, however, our computational model took this into account and solved instead for an infinite queue.

Conclusion

Using both our computational and mathematical models, we observed that the initial rates for customer arrival intervals and service rates maintained an average queue length of 2.79 people and an average wait time of 6.84 minutes. These did not fulfill the management's desires to ensure that the queue length be two or fewer people and for the average customer to wait less than two minutes. We came up with two different possible solutions that could be implemented to lower both the queue length and wait time to fit within the management's request. Our first solution was to hire an additional teller to help with the rate of service. With the addition of our a second teller, our mathematical model approximated that there would be an average queue length of 0.28 people and an average wait time of only 0.98 minutes. Our computational model (considered to be more numerically accurate) calculated that there would be an average queue length of only 0.04 people and an average wait time of 0.07 minutes.

We extended our study to include the second teller only during the peak hours. Our computational model approximated that there would be an average queue length of 0.10 people and an average wait time of 0.28 minutes. We also investigated adding a peak-time teller, which would reduce the average queue length to 0.1 people and the average wait time to 0.27 minutes. Though the addition of a second teller would fall within the management's standard for customer service, we decided to also examine the addition of an ATM in our bank. Using our computational model, there would be an average queue length of 0.34 people and an average wait time of 1.37 minutes during the peak hours. Since the ATM was a one time purchase, met the standards of management, and is cheaper than hiring a second teller in the long run, we decided that an ATM would be a better solution.

In the future, we would try to improve our current models by using supported, rather than predicted, values for inputs such as percentage of people at banks who use ATM's (we arbitrarily predicted 25% in our solution) and additional traffic flow during peak hours (which we arbitrarily assigned as 50%). We would also like to perform more in-depth sensitivity analysis, for example on the computational model in addition to the mathematical model.

Additionally, we would like to investigate other ways the bank could improve its customer service to meet the goals of the manager. For example, we might look at the effects of lengthening the time the bank is open; it follows that increasing the hours the bank is open will cause an overall decrease in the arrival rate throughout the day. Another interesting approach that we would like to investigate would be to give the tellers additional training to increase their speed of processing the customers. Finally, we may consider investigating mobile banking, both in the short-term and long-term (in the long-term, how much of banking can be done online?) to see how it may affect the usage of brick-and-mortar banks in general.

References

Dhar, S.K., and T. Rahman. May 2013. Case Study for Bank ATM Queuing Model. IOSR Journal of Mathematics (7): 1-5. Retrieved from: http://www.iosrjournals.org/iosr-jm/papers/Vol7-issue1/A0710105.pdf

Neogi, S. (n.d.). Queuing Theory Approach to Enhance Decision Making in the Service Sector, 1–18. Retrieved from http://www.apicsterragrande.org/CKNelson 2013 Winners/Queuing Theory Approach to Enhance Decision Making in the Service Sector - A case study.pdf

Occupational Employment and Wages: Tellers. (2013, March 29). *Bureau of Labor Statistics*. Retrieved November 16, 2012, from http://www.bls.gov/oes/current/oes433071.htm

The Real Cost of Running Your Own ATM. (n.d.). Welch ATM. Retrieved November 17, 2013, from http://www.welchatm.com/real-cost-of-running-atms-infographic.html

White, R. 2012. What Is a Peak Time Teller?. *Chron*. Retrieved November 17, 2013, from http://work.chron.com/peak-time-teller-26558.html

Appendix

Bank Hours

Bank of America

- Mon-Thurs: 9-5; Fri: 9-6; Sat: 9-1; Sun: closed
- http://locators.bankofamerica.com/locator/locator/DURHAM_NC/branch_and_atm_locations/locations.html

Wells Fargo:

- Mon-Thurs: 9-5; Fri: 9-6; Sat: 9-1; Sun: closed
- https://www.wellsfargo.com/locator/search/Durham_NC/

Coastal:

- Mon-Fri: 8:30-5:30; Sat-Sun: closed
- https://www.coastal24.com/locations/

Chase:

- Mon-Fri: 9-7; Sat: 9-4; Sun: closed
- https://www.chase.com/

Bank Teller Salary

Employment	Employment RSE	Mean hourly wage	Mean annual wage	Wage RSE
541,770	0.6 %	$12.40	$25,790	0.2 %

Percentile	10%	25%	50% (Median)	75%	90%
Hourly Wage	$9.44	$10.41	$11.99	$14.07	$16.50
Annual Wage	$19,630	$21,660	$24,940	$29,260	$34,320

http://www.bls.gov/oes/current/oes433071.htm

Code Sample

Relevant functions are shown below, namely the main processing function and the functions which model the addition and removal of people from the queue. Variable definitions, UI controls, data logging, and other functions have been removed for clarity.

```
void process() //Main Loop, runs simulations in iterations of 1 minute
        {
            if (peakTimes) //Sets peak and off-peak probabilities and # of workers if enabled
            {
                t++;
                if (t == 0)
                {
                    timeBetweenArrival = new double[] { 0, 3, 15, 45, 35, 80, 95, 100 };
                    twoWorkers = false;
                }
```

```
            if (t == 120)
            {
                timeBetweenArrival = new double[] { 10, 20, 45, 80, 95, 100, 100, 100 };
                //twoWorkers = true;
            }
            if (t == 240)
            {
                timeBetweenArrival = new double[] { 0, 3, 15, 45, 35, 80, 95, 100 };
                twoWorkers = false;
            }
        }

        if (atmQueue != 0) //Processes ATM queue every 3 iterations
        {
            atmTimer++;
            if (atmTimer >= 3)
            {
                atmQueue--;
                atmTimer = 0;
            }
        }

        peopleInQueue += numberOfPeopleArriving();

        //Data Recording
        int numPeopleLeaving = worker1() + worker2();
        peopleInQueueData.Add(peopleInQueue);
        totalQueueWaitTime += peopleInQueue;
        totalPeopleProcessed += numPeopleLeaving;

        updateUI(); //Updates UI elements

    }

    int numberOfPeopleArriving() //Number of People Arriving at the bank each iteration
    {
        timeSinceLastArrival++;

        if (random.Next(0, 1000) < 10 * timeBetweenArrival[timeSinceLastArrival])
        {
            if (ATMActive)
            {
                if (random.Next(0, 100) < 25)
                {
                    if (atmQueue < (peopleInQueue + 1))
                    {
                        timeBetweenArrivalsData.Add(timeSinceLastArrival);
                        timeSinceLastArrival = 0;
                        atmQueue++;
                        return 0;
                    }
                }
            }

            timeBetweenArrivalsData.Add(timeSinceLastArrival);
            timeSinceLastArrival = 0;
            if (random.Next(0, 100) < timeBetweenArrival[0])
            {
                timeBetweenArrivalsData.Add(0);

                if (random.Next(0, 100) < timeBetweenArrival[0])
                {
                    timeBetweenArrivalsData.Add(0); //Chance of >3 people per iteration: .01%
(negligible)
                    return 3;
                }
                return 2;
            }
```

```
            return 1;
        }

        else
        {
            return 0;
        }
    }

    int worker1() //Teller: returns 0 if still working, 1 if finished
    {
        returnValue = 0;
        if (processing)
        {
            timeBeingServed++;

            if (random.Next(0, 1000) < (10 * serviceTime[timeBeingServed]))
            {
                timeBeingServedData.Add(timeBeingServed);
                processing = false;
                returnValue = 1;
            }
        }

        if (!processing && peopleInQueue > 0)  //possible issue? minutes
        {
            processing = true;
            peopleInQueue--;
            timeBeingServed = 0;
        }
        return returnValue;
    }

    int worker2() //Teller 2: Enabled if twoWorkers is set to true
    {
        if (twoWorkers)
        {
            returnValue2 = 0;
            if (processing2)
            {
                timeBeingServed2++;

                if (random.Next(0, 1000) < (10 * serviceTime[timeBeingServed2]))
                {
                    timeBeingServedData.Add(timeBeingServed2);
                    processing2 = false;
                    returnValue2 = 1;
                }
            }

            if (!processing2 && peopleInQueue > 0)  //possible issue? minutes
            {
                processing2 = true;
                peopleInQueue--;
                timeBeingServed2 = 0;
            }
            return returnValue;
        }
        return 0;
    }
```

Team #4342

Dear Manager,

We are writing to report the results of our study on the company's current customer service state. Last week, you gave us the task of analyzing whether our branch currently satisfies the customer service standards you set forth, and if not, determining the most minimal change required to restore the level of customer service to the level defined in these standards.

First, as employees, we would like to let you know we fully support the intentions of this study. Through our interactions with customers each day, we are fully aware that quality customer service -- whether it be the positive attitudes of the employees or the services the bank offers as a whole -- is integral to keeping our customers happy and coming back and improving the efficiency of our bank.

As a result, we wholeheartedly agree with the standard of customer service you set forth stating that, on average, customers should experience a wait time no longer than two minutes, and that no queue should have more than two people in it on average.

For this study, we first set up a mathematical model of our current system with one teller using data on existing arrival and service rates. From this we saw that customers can expect to spend on average 6.84 minutes waiting in lines with an average line length of 2.79 people. These values are clearly larger than those outlined in our satisfaction guidelines.

To complement our mathematical model, we also created a computer simulation of our current system. Although less elegant than a mathematical model, the computer simulation allows us to model more complex situations like the peak hours that our bank experiences daily where lines get especially long. During these peak hours, our computer simulation shows that we can have lines over 10 people long, corresponding to wait times of over 20 minutes. These wait times are aggravating our loyal customers, and we are pleased to see that you have taken an interest in these issues and are searching for a solution to this problem.

After thoroughly evaluating our current banking service system using our mathematical and computer models, we would like to present several possibilities for improving the customer service at our bank. Of course, since we are also interested in maximizing the profit of our bank, we will also consider the cost of implementing these changes. The most basic solution, which I'm sure you've considered, is that of hiring a new teller. There's plenty of space behind the desk to house an additional teller, and adding a second teller would greatly ease the load on the overworked teller that is handling all of our current customers. Currently our single teller is serving a customer 92% of the time, so he/she has very little down time. Based on our computer simulations, hiring an additional teller would reduce our average line lengths to .04 people and average wait time to a mere 4 seconds. However, the annual salary of a bank teller is about $25,000 so this would not be an insignificant cost to our bank.

Additionally, having two full-time tellers is unnecessary. Therefore, we've also considered the possibility of hiring an additional part-time teller who would work during peak time thus helping to reduce the load on our current teller in these busy yet important parts of the day. We find that in this scenario, with a part-time teller working only 2 hours/day, we can reduce the queue length from its current state to about 0.1 people long on average and the wait time to about 16 seconds. This solution satisfies your customer service criteria, and also saves our bank a lot of money when compared to hiring a full-time teller. A teller is paid about $12.40/hour, and at 13 hours a week (2/weekday and 3 on Saturday), this would only cost our bank $8,300 a year!

Although hiring an additional part-time teller seems like an attractive solution, we still considered additional options, like adding an ATM to our facility. By placing a drive-up ATM outside our bank, some people could use it instead of talking to the teller. Therefore, the load on the teller is reduced. Given the assumption that 25% of people would be willing to use an ATM instead of waiting to talk to the teller, we find that, without hiring any tellers, we can reduce the queue length of both queues (ATM and teller) to less than 0.33 people long on average, and maintain an average wait time of 1.4 minutes, even when considering peak hours of operation. This achieves your goals for customer service and also is very cost-effective. A typical ATM installation costs less than $25,000. This equates to three years of paying a part-time employee, or just a year of salary for a full-time employee. Therefore, you will recoup your investment very quickly, and, when the payback period is over in 1-3 years, be saving thousands of dollars per year while maintaining excellent customer service!

Installing a drive-through ATM at our bank also has several other benefits. First, an ATM is open 24/7, whereas our bank is only open on average about 7 hours a day, 6 days a week. This corresponds to a fourfold increase in the amount of time in which our customers can perform transactions at our bank - from 42 hours currently to 168 hours. Our customers will appreciate these extra hours, as they will never again drive to our bank only to find it closed! Additionally, due to the greatly increased hours of operation at our bank due to the ATM, potential customers that had conflicts with our hours of operation due to work or other factors may now be able to use our bank, leading to increased revenue.

Although we considered modeling numerous other options for improving customer service beyond what is listed here, such as extending the hours of operation of the bank or improving the processing speed of tellers (through job training, hiring more experienced tellers, improving computing systems and workflow, etc), we believe that either hiring a peak-time teller or installing an ATM are the best solutions for achieving your standards for customer service. If you are looking for a quick, low-investment fix, hire the peak-time teller. However, if you want a long-term money saver, the ATM is your best option - after a few year payback period, it will save our bank thousands of dollars a year, with the added bonus of improving hours of operation and customer satisfaction.

Sincerely,

Your Devoted Employees