Comparing Single Line, Multi-Register and Multi-Line Single Register Queuing Models

Purpose:

The purpose of this research is to determine how the mean and variance of waiting times compare between two queuing structures. The two structures being analyzed are the single line, multi-register model and the multi-line single register model.

In the single line, multi-register set up, all customers are in a long, single line that is serviced by multiple registers. As any of the customers at one of the multiple registers finishes, the next customer in line is serviced. In this model, there is no strategy that the customer can use to attempt to reduce wait time.

In the multi-line, single register model, customers can choose which line they will use. Each line only has one register; so this model is more vulnerable to an increase in wait times due to slower customers. With this model, the customer can use the strategy of choosing the shortest line in order to attempt to lower wait times. Since the amount of time each customer takes to checkout is a random variable, there is not guarantee that the shortest line will result in the shortest wait time.

The Models:

In order to compare the two line types, I created two models using Python. Here I will describe the methodology and the assumptions behind both models.

Single Line, Multi-Register:

To model the single line, multi-register queue model, I first created a list of n random numbers between 0 and 1. These numbers represent a unit-less time measurement that is assigned to each customer to serve as a time-at-register metric. I then remove the r first customers (r representing the number of registers) from list and put them into another list representing the customers currently at the register. In the list of customers at the register, the lowest time-at-register is subtracted from all r customers' time-at-register and appended to a list that holds the total passage of time. This brings the lowest customer's wait time to 0. The lowest customer is then removed from the list and the next customer in line is added to the register list and taken away from customers in line list. This process continues until the last customer is serviced. After the last customer is serviced, the list that represents the total passage of time is summed and the summation is the output. The model output represents how long the n^{th} customer has to wait given the number of customers (n) and the number of registers (r). This base model is used to iterate through multiple trials in order to create a sample of times that a customer waits in line given number of customers (n) and number of registers (r).

Assumptions:

The model assumes that customer's time at a register is represented by a uniform distribution.

Multi-Line, Single Register:

To model this queuing structure, I first assign all n customers a random value between 0 and 1. Just like in the other model, this represents how much time each customer will take at the register. I then use a random integer generator to select an integer between 1 and r (the number of registers). The random integer is used to assign each customer, but the last, to a line. After the assignments have been made, the r number of lists populated randomly by customers represent the state space that the n^{th} customer is presented with. The n^{th} customer is assumed to select the shortest line in order to maximize likelihood of a lower wait time. In the model, the n^{th} customer's wait time is represented by adding her/his wait time to the summation of the wait times of the shortest line. This is the singular output of the model. As with the other model, the single trial created in this model is iterated multiple times to get a sample of waiting times given n number of customers and r number of registers.

Assumptions:

(1) Like the single line, multi-register model, it assumed that time at register is distributed as a uniform random variable between 0 and 1. (2) An additional assumption is that the customer always chooses the shortest line. This may not always hold true in application for a couple of reasons. One reason is that the customer may not have information on the length of all lines. Some stores have very large checkout areas or checkout areas that have shopping racks on both sides of each register obscuring the view of how many people are in line. Another reason this may not hold is because customers may attempt to predict the approximate check out time of other customers. For example, if I'm at the grocery store and I see a line with two people, each buying a package of gum and another line with one person with a cart completely full of groceries, I'm going to choose the longer line, which violates the assumption of the model. Despite my concerns with this assumption, I think the model is a good measure of how most customers choose a line most of the time. (3) The third assumption is that the n^{th} customer is presented with a completely random state space, where 3 lines with 0, 0 and 15 customers is equally likely to three lines of 5, 5 and 5 customers. While there does appear to be some amount of randomness in line lengths (anecdotally from my personal observation at the store), it doesn't seem to be completely random.

Initial Analysis:

Below are the tables showing the results of running 200 trials through both models on various numbers of customers and registers. The tables include the sample mean and variance for each model and the 99% confidence interval for the difference in means and ratio of the variances between the two models.

Cashier Number	Customer Number	Multi-Line Mean	Multi-Line Variance	Single Line Mean	Single Line Variance	Mean Difference 99% CI	Variance Ratio 99% CI
2	20	4.51	1.2	4.70	0.42	[-0.07,0.35]	[0.24,0.51]
3	30	4.26	1.23	4.60	0.24	[0.14,0.24]	[0.14,0.28]
4	40	4.00	1.12	4.62	0.19	[0.47,0.85]	[0.12,0.25]
5	50	3.70	1.02	4.61	0.17	[0.73,1.10]	[0.11,0.24]
6	60	3.52	0.95	4.66	0.13	[0.97,1.32]	[0.09,0.20]
7	70	3.42	1.00	4.66	0.12	[1.06,1.41]	[0.08,0.17]
8	80	3.31	1.04	4.66	0.09	[1.17,1.53]	[0.06,0.12]

Cashier Number	Customer Number	Multi-Line Mean	Multi-Line Variance	Single Line Mean	Single Line Variance	Mean Difference 99% CI	Variance Ratio 99% CI
2	10	2.36	0.08	2.09	0.18	[-0.43,-0.11]	[0.16,0.33]
3	15	2.01	0.53	2.08	0.14	[-0.06,0.21]	[0.19,0.40]
4	20	1.82	0.56	2.12	0.10	[0.16,0.43]	[0.11,0.23]
5	25	1.81	0.53	2.11	0.07	[0.17,0.43]	[0.09,0.19]
6	30	1.60	0.49	2.1	0.06	[0.37,0.61]	[0.09,0.19]
7	35	1.60	0.42	2.15	0.05	[0.44,0.66]	[0.08,0.17]
8	40	1.58	0.45	2.15	0.05	[0.45,0.69]	[0.08,0.16]

Cashier Number	Customer Number	Multi-Line Mean	Multi-Line Variance	Single Line Mean	Single Line Variance	Mean Difference 99% CI	Variance Ratio 99% CI
2	150	35.44	8.97	37.24	2.98	[1.23,2.37]	[0.23,0.48]
3	150	22.7	6.14	24.7	1.46	[1.51,2.42]	[0.16,0.34]
4	150	16.21	5.02	18.23	0.68	[1.63,2.42]	[0.09,0.20]
5	150	12.34	3.24	14.65	0.47	[1.99,2.63]	[0.10,0.21]
6	150	10.16	2.92	12.14	0.33	[1.68,2.74]	[0.08,0.17]
7	150	8.17	2.46	10.37	0.21	[1.94,2.48]	[0.06,0.12]
8	150	6.91	2.10	9.01	0.17	[1.85,2.35]	[0.06,0.12]

Conclusions:

Mean Comparison:

The multi-line mean had a statistically shorter wait time in all but the 2 register, 20 customer; 2 registers, 10 customer; and 3 registers, 15 customer trials. The difference in mean times appear to be a function of number of customers and registers. As the number of customers and registers increases, the difference tends to increase. I believe the explanation for the trend has to do with how the lines are set up in the multi-line model. The advantage of the multi-line model is the potential presence of shorter lines that customers can choose to increase the chance of a shorter wait time. I'll demonstrate that as customer number increase, even as cash registers increase proportionately, the probability of the existence of a shorter line increases.

Think of the probability that a customer chooses a line as a single Bernoulli trial. The customer either chooses the line or she doesn't. Using the assumption of an equal likelihood of choosing a line independent of line length, the formula for the probability of a single line having a proportionate number of customers is below:

$$\binom{C}{C/L}\left(\frac{1}{L}\right)^{C/L}\left(1-\frac{1}{L}\right)^{C(1-(1/L))}$$

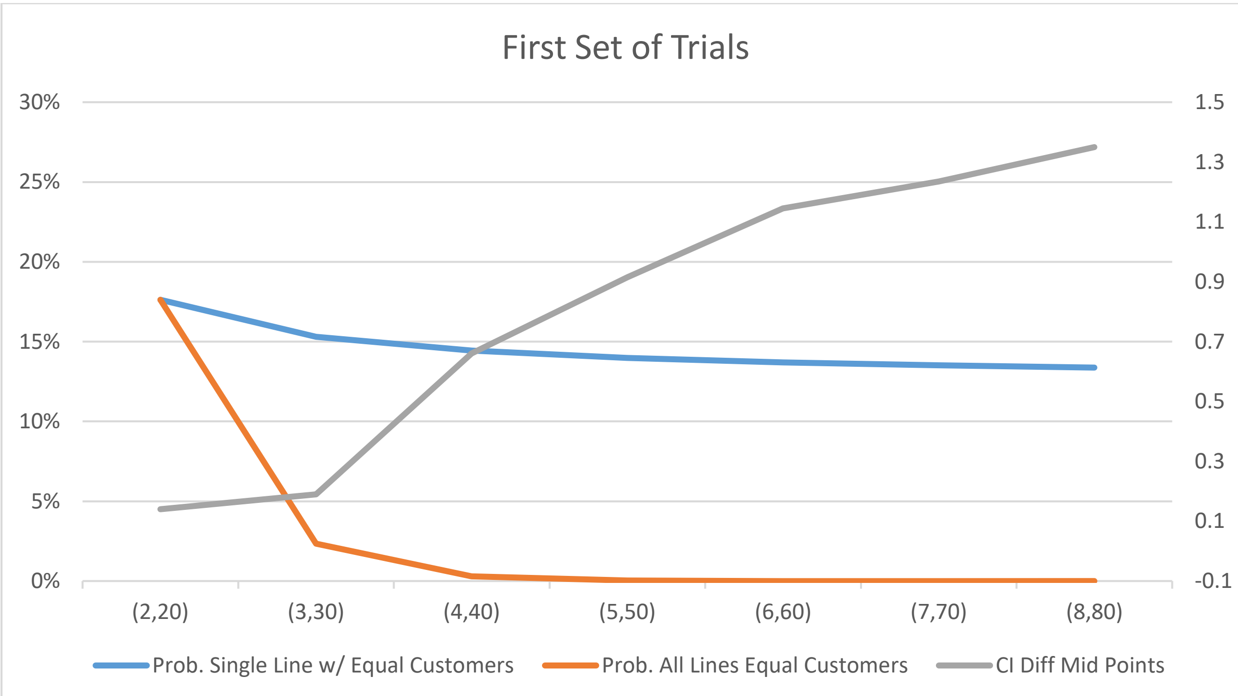
Where C is the number of customers and L is the number of lines.

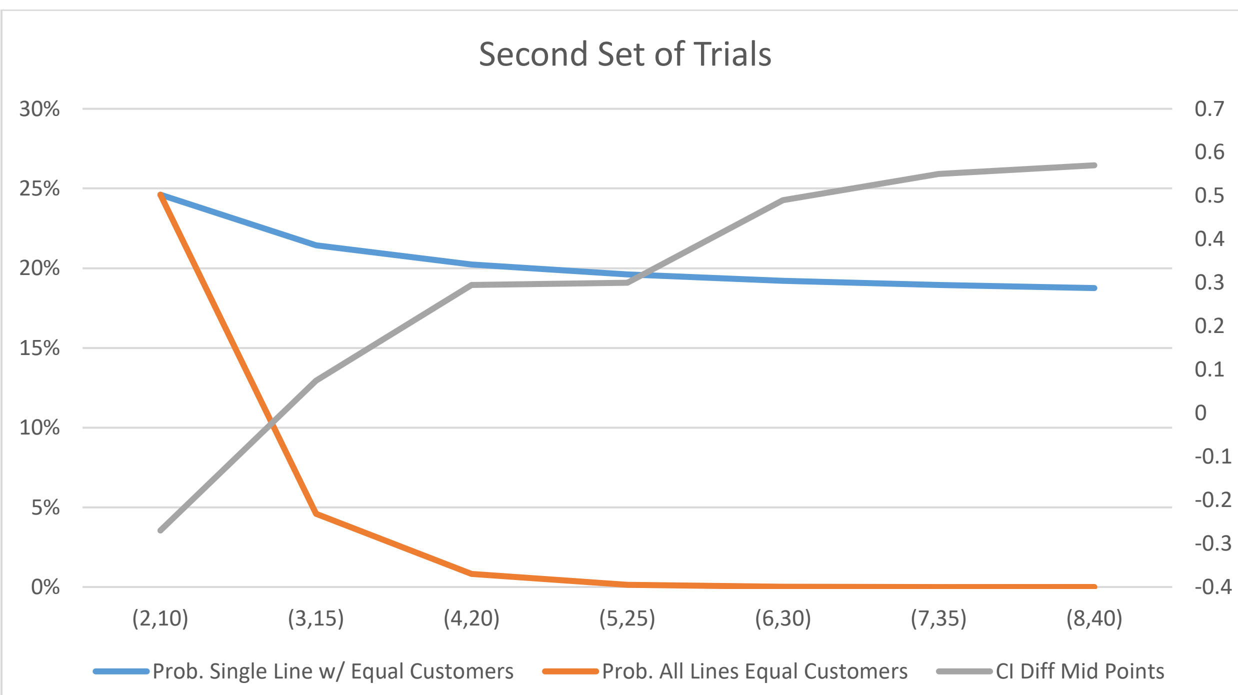
This formula represents the probability that one line has the proportionate number of customers in it (i.e. number of customers divided by the number of lines).

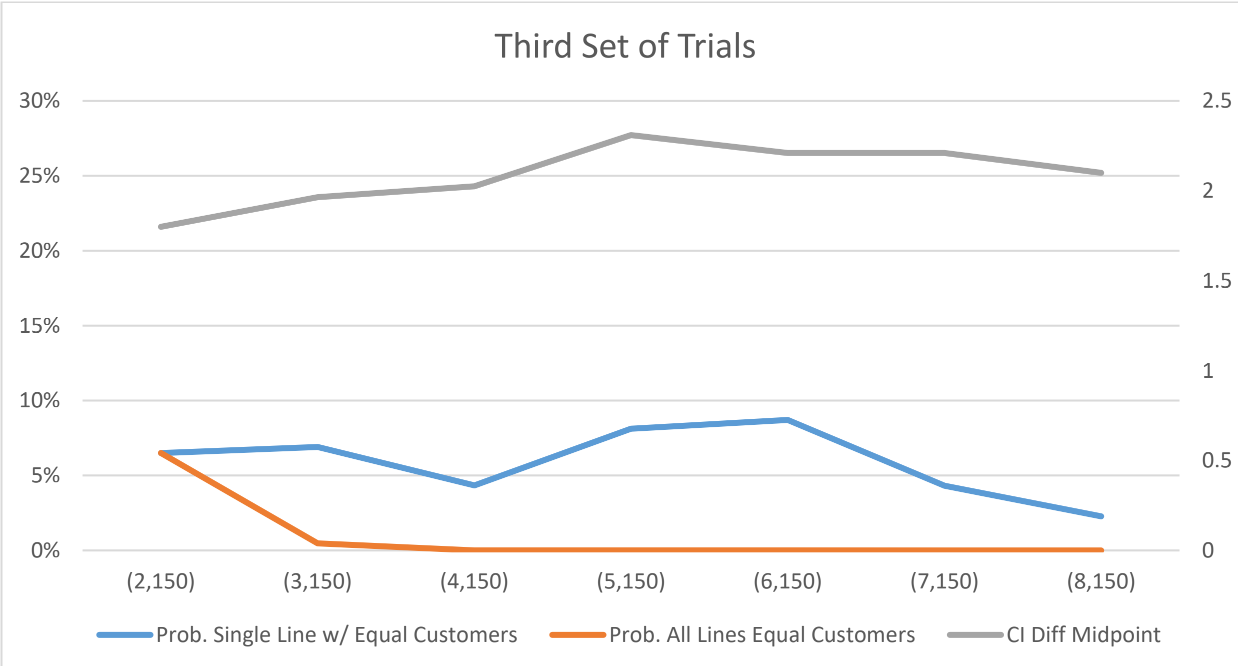
In order to calculate the probability that all lines have the same number of customers, we must raise that probability that one line has a proportionate number of customers to (L-1). We subtract 1 from L because we cannot assume that the last line's customer length is independent of the other lines. If line 1 has 5 customers and line 2 has 5 customers then line 3 must have 5 customers if there are a total of 15 customers. Determining the length of customers in the last line is not probabilistic, therefore we do not include the last line as a trial. The formula for the probability that all of the lines have the same length is shown below:

$$\left(\binom{C}{C/L}\left(\frac{1}{L}\right)^{C/L}\left(1-\frac{1}{L}\right)^{C(1-(1/L))}\right)^{(L-1)}$$

Here are the charts depicting the probability of a single line having a proportionate number of customers in blue, the probability of all the lines having the same number of customers in orange and the midpoint of the mean difference CI in grey (Note: Mean CI Diff is on the right Y-axis).







As the charts show, as the number of lines increase, the probability of all lines having the same number of customers falls close to 0 quickly. The probability of all lines having the same number of customers and the midpoint of the mean difference CI have an inverse relationship.

Since there is quickly a high probability of the existence of a shorter line, it is more likely that the n^{th} customer's wait time will be lower since he or she will be able to find a line shorter than the number of customers divided by the number of lines. I believe this is why we see the multi-line wait time's performance increase as the number of lines increase.

Variance Comparison:

The variance of the single line model is a small fraction of the multi-line model. As the number of registers and customers increase, the variance of the single line model becomes smaller relative to the multi-line model. Based on the data, it looks like with 8 registers, the ratio of the variances seems to converge to about [0.06,0.12], which means that with 99% confidence the variance of the single line model is 6-12% of the variance of the multi-line model. This is a major drawback to the multi-line model.

Analysis with Time at Register as an Exponentially Distributed Variable:

I decided to change the time-at-register from a random number between 0 and 1 to an exponential random variable with lambda of 0.5 (same mean as the random number from a uniform distribution). I think that this is more representative of how lines actually work. I don't think it is equally likely that a customer take the max amount of time as it is that the customer take the mean amount of time, which is what using a random number implies. With the exponential distribution, there are rare events that take a lot more time than the average. This could be when there is some sort of a dispute and a manger is called or when someone decides to pay with a bunch of coins. After making the change to the models, I ran the same tests through them. The results are tabulated below.

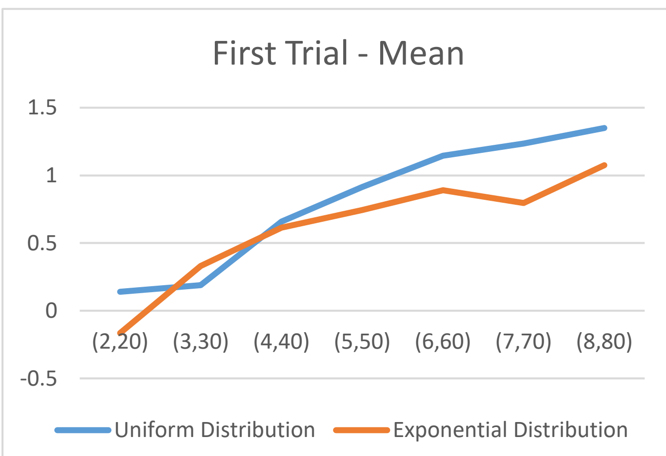
Cashier Number	Customer Number	Multi-Line Mean	Multi-Line Variance	Single Line Mean	Single Line Variance	Mean Difference 99% CI	Variance Ratio 99% CI
2	20	4.68	1.15	4.52	1.23	[-0.42,0.09]	[0.74,1.54]
3	30	4.17	0.97	4.50	0.85	[0.11,0.55]	[0.61,1.27]
4	40	3.97	1.41	4.58	0.52	[0.39,0.84]	[0.25,0.53]
5	50	3.70	1.05	4.45	0.37	[0.55,0.94]	[0.24,0.51]
6	60	3.61	1.17	4.5	0.34	[0.69,1.09]	[0.20,0.42]
7	70	3.65	0.91	4.44	0.33	[0.61,0.98]	[0.25,0.53]
8	80	3.39	1.01	4.47	0.23	[0.89,1.26]	[0.16,0.33]

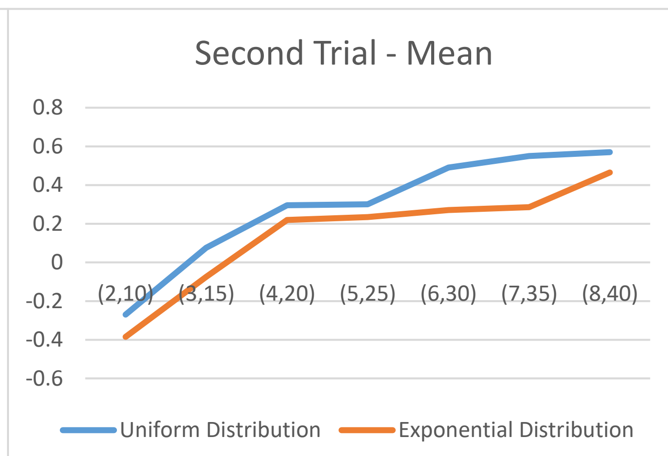
Cashier Number	Customer Number	Multi-Line Mean	Multi-Line Variance	Single Line Mean	Single Line Variance	Mean Difference 99% CI	Variance Ratio 99% CI
2	10	2.45	0.7	2.06	0.4	[-0.56,-0.21]	[0.40,0.83]
3	15	2.07	0.51	2.00	0.29	[-0.22,0.07]	[0.39,0.82]
4	20	1.84	0.58	2.06	0.24	[0.07,0.37]	[0.28,0.59]
5	25	1.74	0.53	1.97	0.16	[0.10,0.37]	[0.21,0.44]
6	30	1.71	0.44	2.00	0.15	[0.14,0.40]	[0.24,0.49]
7	35	1.73	0.51	2.02	0.14	[0.15,0.42]	[0.19,0.40]
8	40	1.60	0.43	2.06	0.1	[0.34,0.59]	[0.17,0.35]

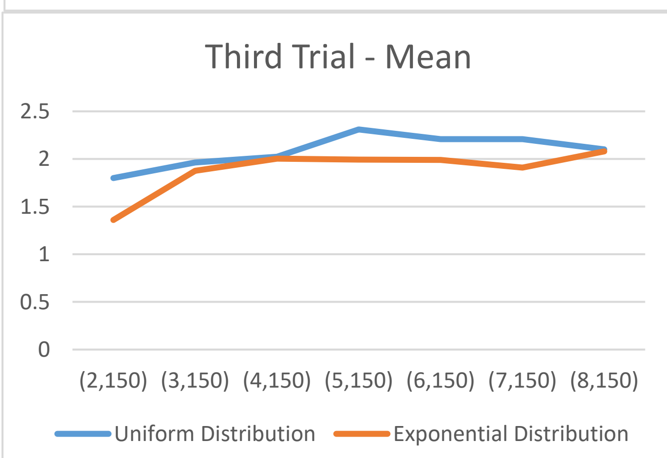
Cashier Number	Customer Number	Multi-Line Mean	Multi-Line Variance	Single Line Mean	Single Line Variance	Mean Difference 99% CI	Variance Ratio 99% CI
2	150	35.56	9.53	36.92	9.02	[0.65,2.07]	[0.65,1.37]
3	150	22.6	6.42	24.47	3.73	[1.35,2.40]	[0.40,0.84]
4	150	16.05	1.78	18.05	1.78	[1.59,2.42]	[0.27,0.57]
5	150	12.53	3.37	14.53	1.73	[1.62,2.37]	[0.35,0.74]
6	150	9.93	3.18	11.92	0.98	[1.65,2.33]	[0.21,0.44]
7	150	8.21	2.34	10.13	0.75	[1.62,2.20]	[0.22,0.46]
8	150	6.80	2.30	8.90	0.54	[1.80,2.36]	[0.16,0.34]

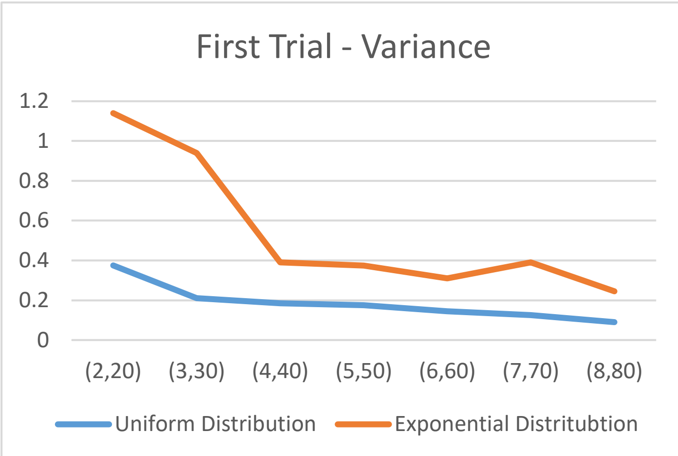
Conclusions:

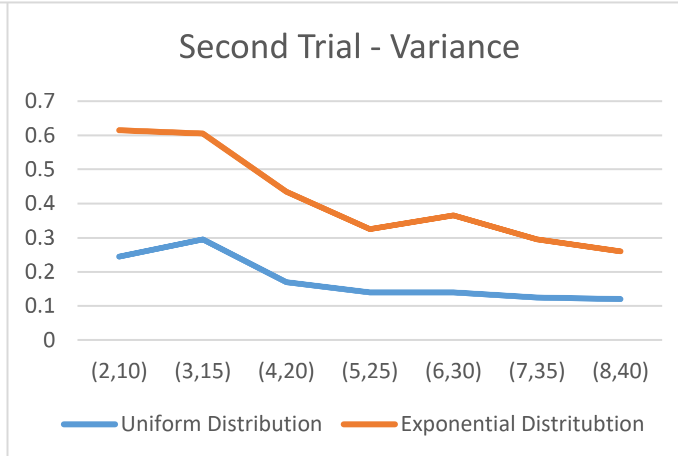
Below are the charts of the mid points of the CI of the mean difference and CI of the variance ratio:

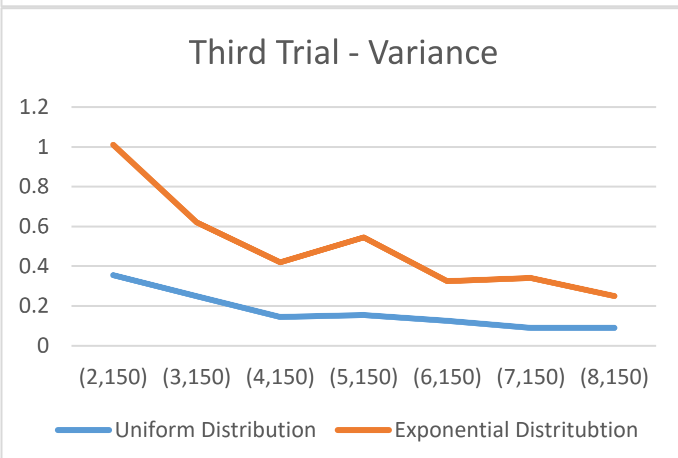












Conclusions:

It appears that the exponential distribution has a lower difference in means but a higher ratio of variance than the uniform random distribution in the model. Although the magnitudes are different, they follow a similar pattern and result in essentially the same statistical conclusions.

I think the exponential distribution makes more sense for this application because it is a distribution that is not truncated to the right. Meaning there is a low probability of customers who take a really long time. In the uniform distribution, all wait times have the same probability. Intuitively it doesn't make sense that a customer that will take the maximum amount of time at a register has the same probability as a customer that will take the average time at the register. For that reason, I think the exponential distribution does a better job at modeling the true behavior of the time-at register random variable.

Overall Conclusions:

Based on the results of the models and statistical tests, I conclude that the multi-line model has a lower average mean (in almost all cases) but a much higher variance. Intuitively this makes sense, the multi-line model has a varying length of line and each line is at a higher risk of getting slowed down by exceptionally slow customers. In the single line model, the length of the line is the same and an exceptionally slow customer is not as influential since there are other registers that continue to move customers through the line while the slow customer is being served.

So, in a situation where the lowest average wait time is the objective, the multi-line model is best. In a situation where the lowest variance is the goal, the single line model is the best. In the more likely situation where an appropriate balance between mean and variance is desired, the decision is not clear and depends on the specifics of the objectives. I should reemphasize that the variance of the multi-line model is much higher, in some cases as much as ten times higher than the single line model. This should be taken into account when the objective is something other than lowest average time or lowest variance of time. Customer satisfaction may be higher in the single line model because fewer will have extremely long wait times due to the lower variance.

Next Steps:

There are two parts of the analysis that could use further exploration. (1) The first is customer behavior in the multi-line model. How do customers really choose what line they wait in? Is it based on length of line alone? Proximity of line to current position? Is it completely random as modeled (except for the last customer who chooses the shortest line)? Does it follow some kind of stochastic process with higher probabilities for shorter lines and lower probabilities for longer lines? With a better understanding of how customers really choose lines I would be able to better model the multi-line scenario and therefore have more accurate conclusions. Depending on how different customer's line choice is from random, the output of the model could change drastically. (2) The second is modeling customers coming into the line as a Poisson random variable. I made my models similar to how it would be if I went to the store. I come to a line or lines that already have people in it, I start my timer and see how long it takes me to complete my purchase. Each iteration of the trial yields just my wait time. If I were to compare my wait time to everybody else's, I would get a biased metric because I'm the last person and no one else is coming behind me. If I was to set the model up as a Poisson random variable and start clocking the times of all customers after the model had run for a little while, I could likely get a more life-like result. I could see random influxes and out fluxes of customers as they randomly go to the check-out lines and see the performance of both queuing set ups in a more application-oriented model.