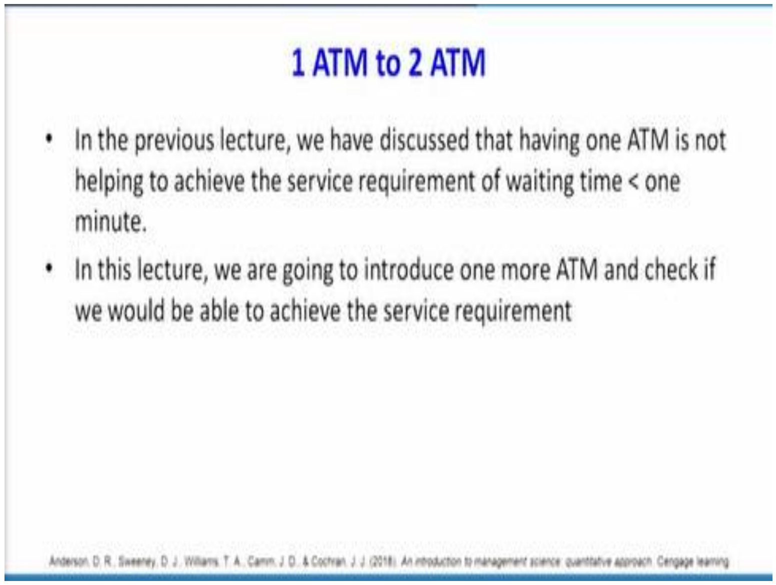
Decision Making with Spreadsheet
Prof. Ramesh Anbanandam
Department of Management Studies
Indian Institute of Technology, Roorkee

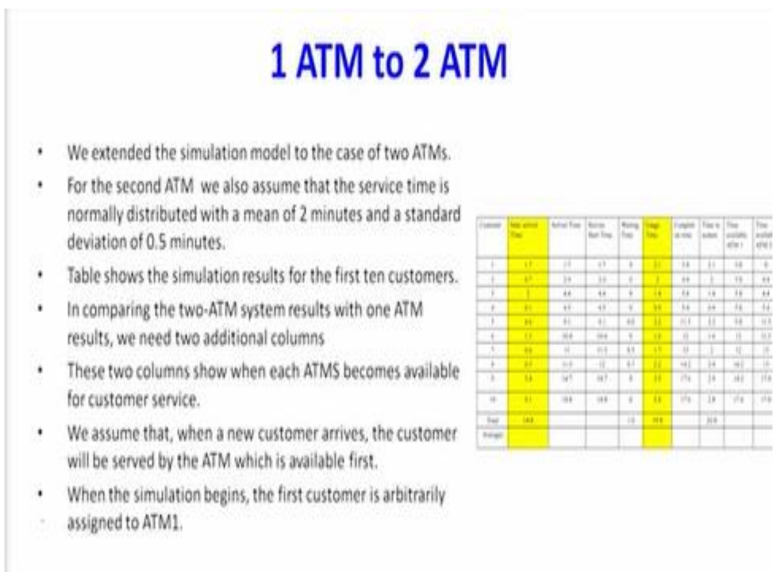
Lecture - 45
Simulation - 4

Dear students, in the previous lecture I discussed how to use simulation by having 1 ATM to meet the service requirement of the average waiting time should be less than 1 minute. But we have seen that we are not able to achieve that service requirement. So, in this class, I am going to discuss introducing new ATMs and whether we are able to meet the service requirement or not.

So, the agenda for this lecture is a waiting line simulation with 2 ATMs after that verification and validation.



In the previous lecture, we discussed that having 1 ATM is not helping to achieve the service requirement of waiting less than one minute. So, in this lecture, we will introduce one more ATM and check if you can meet the service requirement.



So, now 1 ATM to 2 ATM, so, we extended the simulation model to the case of 2 ATMs. For the second ATM, we also assume that the service time is normally distributed with a mean of 2 minutes and a standard deviation of 0.5 minutes. The table shows the simulation result of the first 10 customers. Look at the table on the right-hand side. So, I have given simulation results for the first 10 customers.

In comparing the 2 ATM system results with 1 ATM result, we need two additional columns. Previously, there was only 1 column of time available for ATM. Now, we have introduced one more column. We assume that when a new customer arrives, the customer will be served by the ATM, which is available first. That is the assumption that there may be 2 ATMs. Both are idle, but a new customer is entering into that room.

So, he will go to the ATM, which is waiting for a longer time, which is ideal for a longer time. That is one assumption. And when the simulation begins the first customer is arbitrarily assigned to ATM 1. That is the second assumption.

Two ATM Simulation Model

Customer	Inter arrival Time	Arrival Time	Service Start Time	Waiting Time	Usage Time	Completion time	Time in system	Time available ATM 1	Time available ATM 2
1	1.7	1.7	1.7	0	2.1	3.8	2.1	3.8	0
2	0.7	2.4	2.4	0	2	4.4	2	3.8	4.4
3	2	4.4	4.4	0	1.4	5.8	1.4	5.8	4.4
4	0.1	4.5	4.5	0	0.9	5.4	0.9	5.8	5.4
5	4.6	9.1	9.1	0.0	2.2	11.3	2.2	5.8	11.3
6	1.3	10.4	10.4	0	1.6	12	1.6	12	11.3
7	0.6	11	11.3	0.3	1.7	13	2	12	13
8	0.3	11.3	12	0.7	2.2	14.2	2.9	14.2	13
9	3.4	14.7	14.7	0	2.9	17.6	2.9	14.2	17.6
10	0.1	14.8	14.8	0	2.8	17.6	2.8	17.6	17.6
Total	14.8			1.0	19.8		20.8		
Averages									

Now I will explain the 2 ATM simulation model. So, here I have written the customer inter-arrival time. This was randomly generated as usual; this also follows a uniform distribution. And arrival time, you see 1.7 randomly generated inter-arrival time. So, the arrival time is 1.7. So, immediately you will receive the service. So, the service start time is 1.7; then usage time follows a normal distribution. This also randomly generated then the completion time.

Completion time is service start time plus usage time. So, 1.7 + 2.1 so 3.8. So, time in the system is completion time minus arrival time. So, that is 3.8 - 1.7 = 2.1. And the first customer is going to get the service from the ATM 1. So, ATM 1 is available for service only after 3.8 minutes. So, the second item is free. So, look at the second customer. The second customer inter-arrival time is 0.7 minutes. So, for this arrival time, I have written on a continuous scale, so 0.7 + 1.7.

So, it will be a 2.4 service start time. Now you see that service start time. So, when we are starting the service start time, he is comparing these two. What is happening for 1 ATM? The time available for ATM 2 is 0. So, he will directly enter into the ATM 2. So there will not be any waiting time. So, he will take 2 minutes usage time. So, the completion time is 4.4. How? 2+2.4. The time in the system is 4.4 - 2.4. It is a 2. Now this updation is important.

So, what is the minimum available time for the previous customer? For example, 3.8 and 0. So, 0 is the minimum. So, the completion time will be updated for the second ATM. So, time is available for ATM 1 because it is not the minimum that will be carried to the next customer. So, I will explain the whole process with the help of Excel. Now I go to Excel. I will explain how to use the 2 ATM simulation model.

Now, I will explain the bank with the 2 ATMs with the help of Excel. So, the inter-arrival time follows a uniform distribution. So, the smallest value is 0, the largest value is 5, and the service time follows a normal distribution mean to standard deviation of 0.5. So, what I have done? Here, I have generated 1000 customers. What is our target? So, after doing a simulation for 1000 customers, we have to see what is the mean waiting time. So, for example B16 is inter arrival time.

So, for the first 5 customers, the values are static. So, look at B21. So, B21 I have used this formula. (A + B4 + RAND). So, (B5 - B4A + RAND (B – A)). Like that, I have developed inter-arrival time. I have generated inter-arrival times for 1000 customers. Now you see the arrival time. So, the arrival time is 9.1 + 0. That is (C20 + B21) because this arrival time, I am writing on a continuous scale. The service starts time. Here, you have to see if C21 is less than or equal to a minimum of I20 and J20.

Then the service start time is a minimum of I20 and J20; otherwise, C21. Suppose the sixth customer arrives at 9.1 minutes, but ATMs 1 and 2 will be available. ATM 1 is available after 5.8 minutes, and ATM 2 is available after 11.3 minutes because this 9.1 is greater than 5.8. He will directly enter to get the service. Now the waiting time is column 3 - 2, which is a service start time minus arrival time. So, the service time is the service time that follows a normal distribution.

Mean 2 standard deviation 0.5 that I have randomly generated. Now the completion time is column 3 plus column 4. That is service start time + service time. So, the time in the system is column 5 minus 2, (completion time - arrival time). Now how I am updating the time available. If I20 is for the fifth customer, if I20 is a minimum of I20 and J20, G21 will be updated. Otherwise, I20 will be updated. For example, here, I20, when we compare 5.8 and 11.3, 5.8 is the minimum.

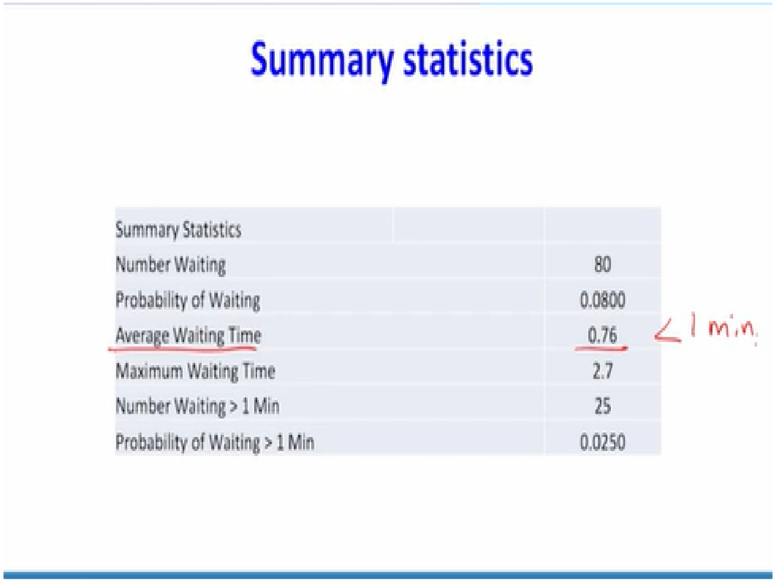
So, for ATM 1 the completion time will be updated from 5.8 to 11.2. The second one, the 11.3 just carried to the next customer. Look at the seventh customer inter-arrival time is 2 minutes. So, the arrival time is 2 + 9.1 = 11.1 minutes. So, when he will be starting to get the service? You see the previous ATMs 1 and 2. ATM 1 is available for 11.2 minutes. ATM 2 is available at 11.3 minutes. But he is coming early. Early means 11.1 minutes.

But he will start to get service only at 11.2 minutes. So, the waiting time is 11.2 minus service time this is randomly generated. Now you see the completion time is 13.6 minutes. Now the ATM 1 completion time for the 6th customer is 11.2. The time available for ATM 2 is 11.3. So, this 13.6 will be updated wherever there is a minimum time available. For example, in ATM 1, the minimum time available is 11.2. So, this completion time will be updated below the 11.2.

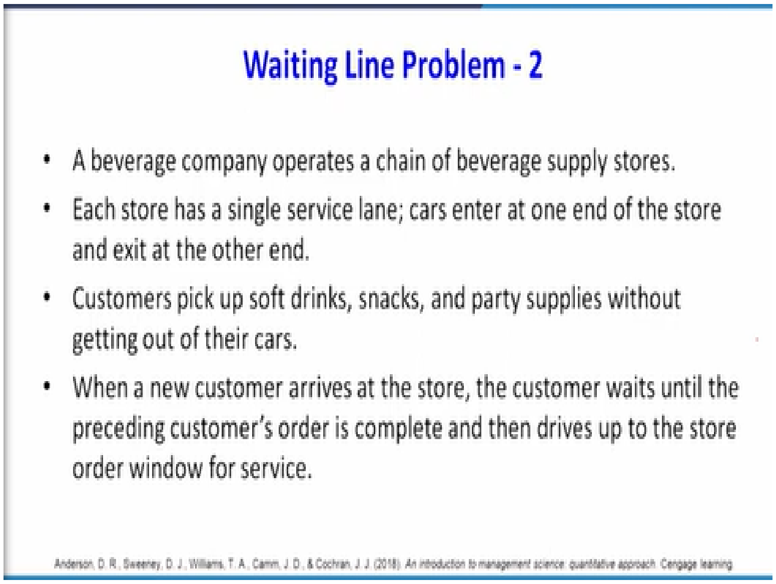
So, what we are saying is that a customer will enter the ATM which is waiting idle for a longer period of time. So, here ATM 1 is waiting for a longer time. So, the seventh customer is entering into ATM 1. So, we are updating the time available for the ATM 1. Like this I have completed for 1000 customers. So, here what are the statistics we are using? The number of people; are waiting. We can use count if.

In that E16 to E1015, how many people are waiting? So, 82. What is the probability of waiting? So, the number of people waiting is divided by the total number of customers, so 1000, which is 0.08. So, the average time waiting time is just I find the average. I have summed the total time divided by the number of people waiting. Number of people waiting is here 82.

The probability of waiting greater than 1 minute, say E1022 upon A1015. Now, the important summary statistic for us is the average waiting time. So, if I keep on pressing F9 so I am going to press F9, you see that when I keep on pressing F9, most of the time, the average waiting time is below 1 minute. So, that is what our service requirement is. So, what are we concluding here? If you have 2 ATMs, you can meet your service requirement of an average waiting time of less than 1 minute.

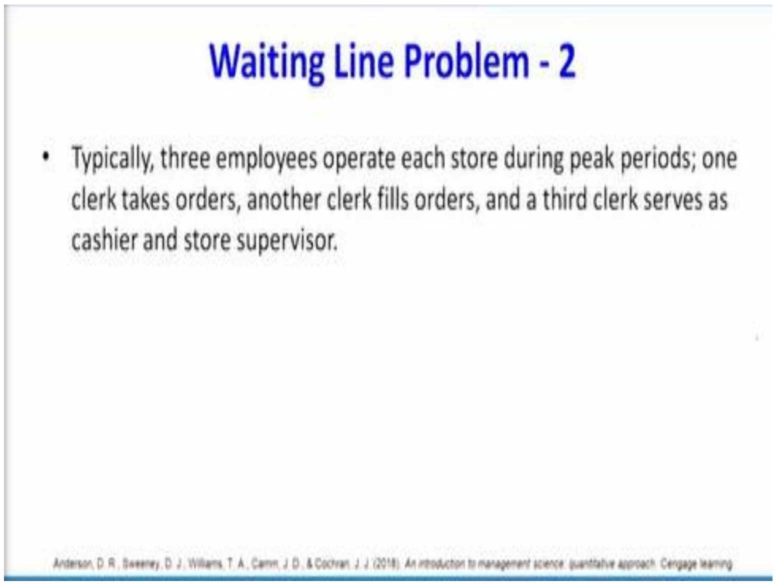


So, I will go back to the presentation. So, the summary statistic says what is more important for us is the average waiting time. So, the average waiting time is 0.76 which is less than 1 minute. So, we are able to achieve the service requirement by having 2 ATMs. That is the conclusion for this simulation problem.



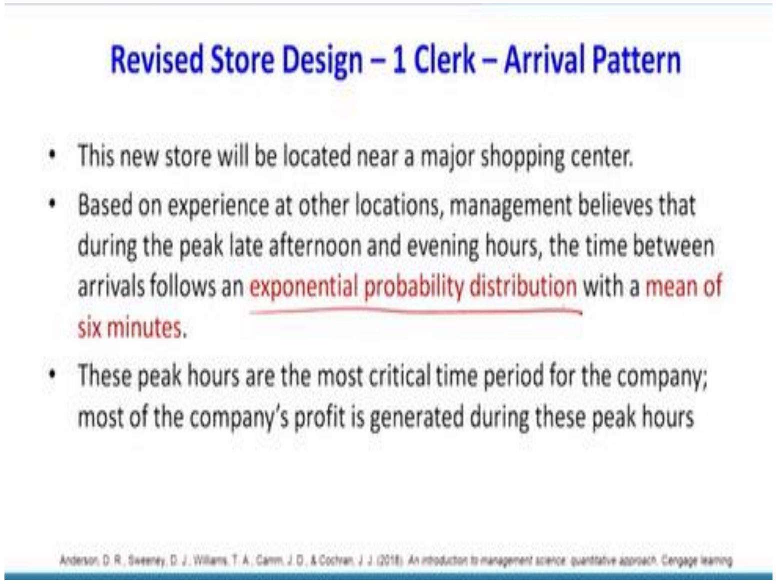
Now I am going to take another problem which is like this ATM problem. So, I am going to solve it just for your understanding. A beverage company operates a chain of beverage supply stores.

Each store has a single service lane. Cars enter at one end of the store and exit at the other end. Customer picks up soft drinks, snacks, and party supplies without getting out of their cars. When a new customer arrives at the store, the customers wait until the preceding customer's order is complete and then drive up to the store order window for service.



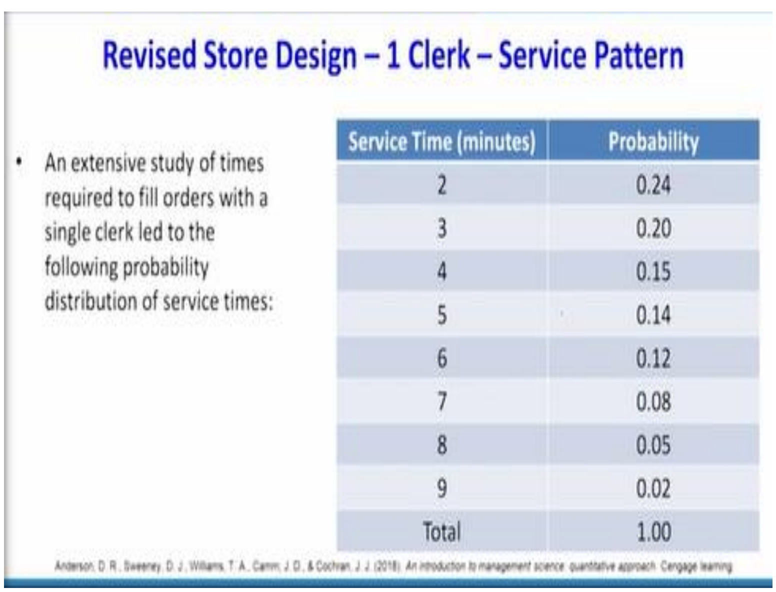
Typically, three employees operate each store during the peak period; one clerk takes orders, another clerk fills the order, and the third clerk serves as a cashier and store supervisor.

The company is considering a revised store design in which computerized order-taking and payment are integrated with specialized warehousing equipment. Management hopes that the new design will permit the operation of each store with one clerk. To determine whether the new design is beneficial, management decided to build a new store using a revised design.



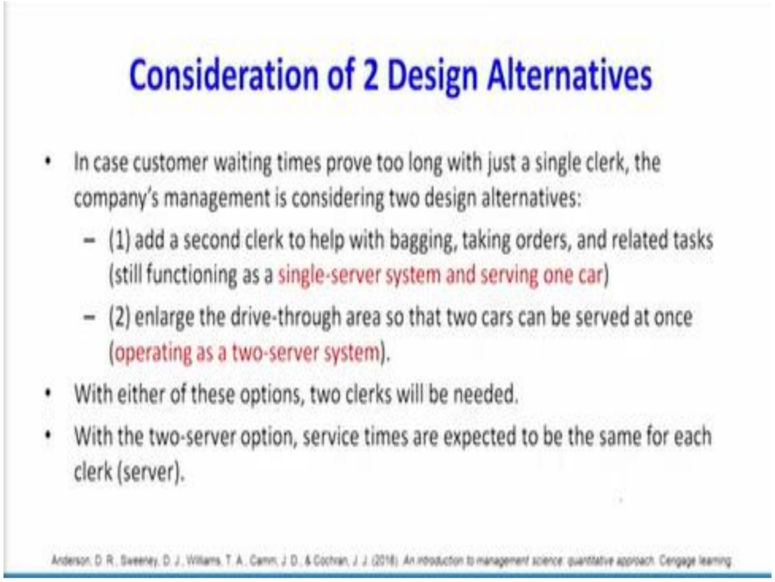
So, what is this revised design? So, revised store design one clerk. This new store will be located near a major shopping center. Based on experience at other locations, management believes that during the peak late afternoon and evening hours, the time between arrivals follows exponential probability distribution. Please remember that in the ATM problem, we have assumed that the arrival follows a uniform distribution.

But here, the arrival follows the pattern of exponential probability distribution with a mean of six minutes. These peak hours are the most critical time period for the company; most of the company's profit is generated during these peak hours.



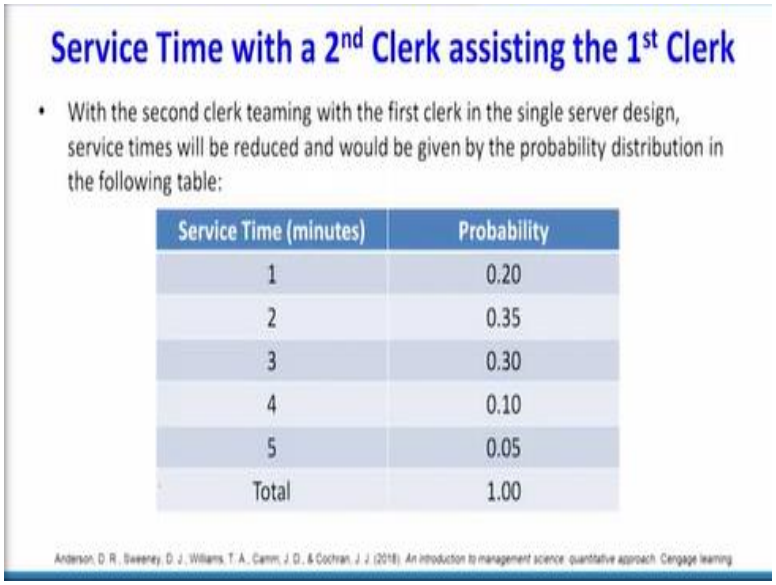
Service Time (minutes)	Probability
2	0.24
3	0.20
4	0.15
5	0.14
6	0.12
7	0.08
8	0.05
9	0.02
Total	1.00

So, we have seen the arrival pattern. Now we will see what is the service pattern. An extensive study of times required to fill orders with a single clerk led to the following probability distribution of service times. So, the service time is 2 minutes. Sometimes, it may take 2 minutes also, 3 minutes, 4 minutes, 5 minutes, 6 minutes. Even there is a possibility to take 9 minutes also, with a 2% probability. So, this is the revised store design when there is only 1 clerk.



So, consider the second design. Consideration of another 2 design alternatives. In case customer waiting time proves too long with just a single clerk, the company's management is considering 2 design alternatives. So, the first one is adding a second clerk to help with bagging and taking orders and related tasks, still functioning as a single server system and serving one car. So, another alternative is to enlarge the drive-through area so the 2 cars can be served at a time.

That is operating as a two-server system. With either of these options, two clerks will be needed. With the two server options, service times are expected to be the same for each clerk, that is, for a single server.



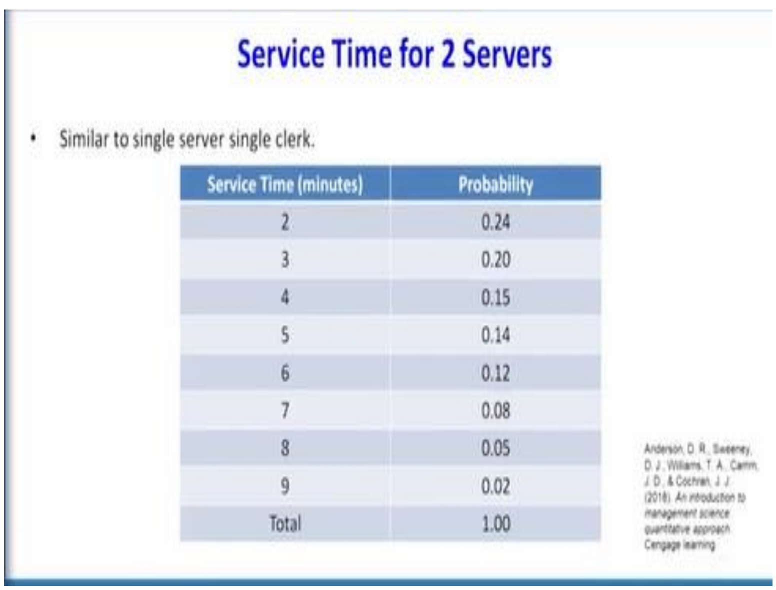
Service Time with a 2nd Clerk assisting the 1st Clerk

- With the second clerk teaming with the first clerk in the single server design, service times will be reduced and would be given by the probability distribution in the following table:

Service Time (minutes)	Probability
1	0.20
2	0.35
3	0.30
4	0.10
5	0.05
Total	1.00

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Cochran, J. J. (2018). An introduction to management science: quantitative approach. Cengage learning

Now service time with the second clerk assisting the first clerk. With the second clerk teaming with the first clerk in the single server design, service time will be reduced and will be given by the probability distribution in the following table. You see that now the service time is decreased when there are 2 clerks assisting a single server. So, the maximum service time is only 5 minutes.



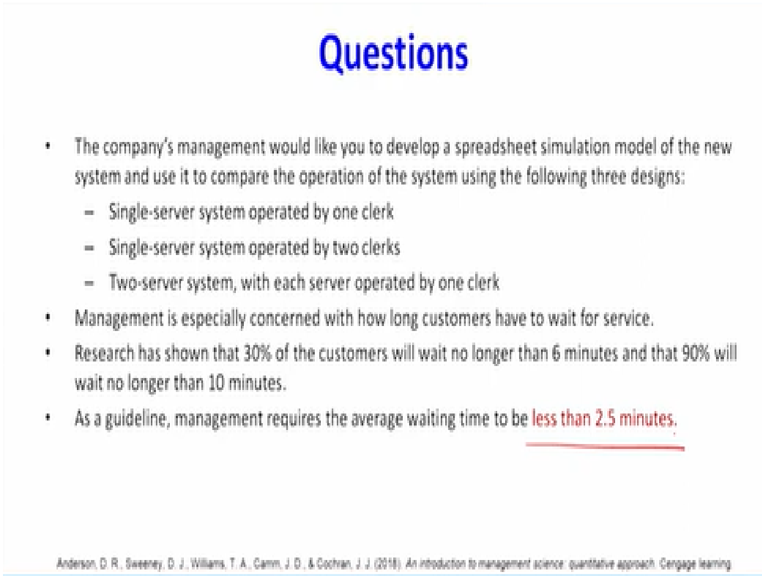
Service Time for 2 Servers

- Similar to single server single clerk.

Service Time (minutes)	Probability
2	0.24
3	0.20
4	0.15
5	0.14
6	0.12
7	0.08
8	0.05
9	0.02
Total	1.00

Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Cochran, J. J. (2018). An introduction to management science quantitative approach. Cengage learning

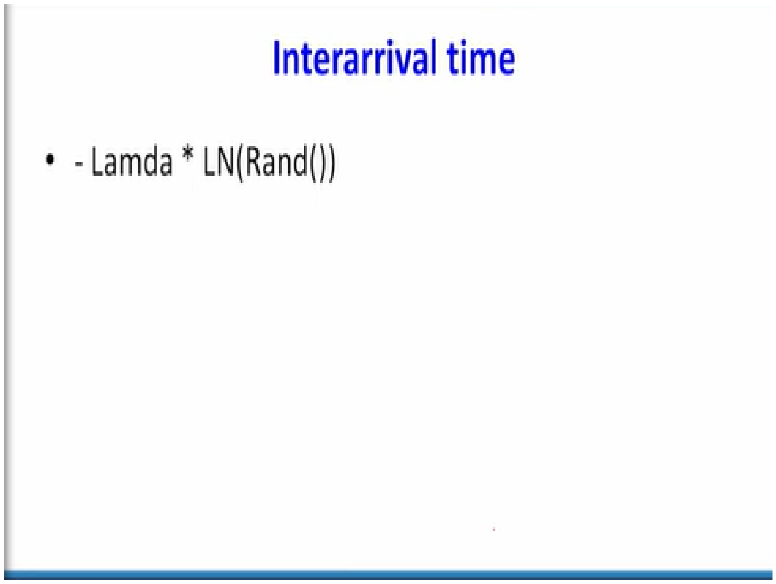
When there are two servers, that means two cars enter at a time, but with a single clerk, the service time is again you see that 2, 3, 4. There is a possibility of 9 minutes also. So, this problem is taken from the book Anderson at all.



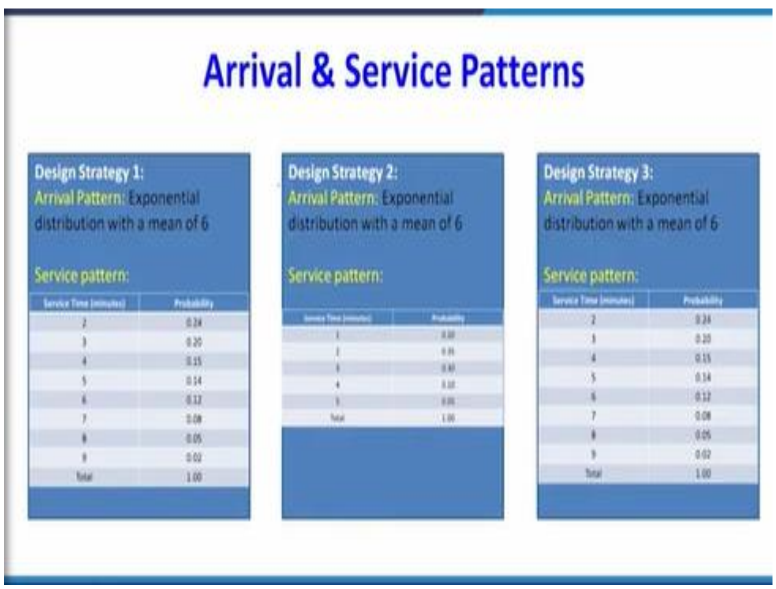
So, what are the questions? The company's management would like to develop a spreadsheet simulation model of the new system and use it to compare the operation of the system using the following 3 designs. What are the 3 designs that are proposed? One is a single server system operated by 1 clerk. The second one is a single server system operated by 2 clerks. Third one two server system with each server operated by one clerk.

So, management is especially concerned with how long customers have to wait for service. Research has shown that 30% of the customers will wait no longer than 6 minutes and 90% will wait no longer than 10 minutes. As a guideline, management requires that the average waiting time be less than 2.5 minutes. So, we are going to simulate all these 3 new designs. And we are going to find out what is the average waiting time.

Whether all these 3 new designs we are able to achieve this target, that is, the average waiting time is less than 2.5 minutes.



Here, to generate a data set that follows exponential distribution in Excel, we have to use - lambda, which is the mean multiplied by the log of Rand. So, this function will be used in Excel for generating random numbers which follow an exponential distribution.



Now, I am going to summarize the arrival and service patterns for the 3 proposed designs. Design 1 arrival pattern exponential distribution with a mean of 6 and service pattern is given. And design 2 is designed for 1 server. There are two clerks. So, here, service patterns and empirical distribution are given. The third design is 1 clerk but 2 servers. It is similar to 2 ATMs. In each server, only 1 clerk is helping. That is why the service time is longer.

So, the problem, is that the service patterns for design 1 and design 3 are the same. So, now we are going to make 3 different types of simulation. For each simulation model, we are going to find out what is the average waiting time. If the average waiting time is less than 2.5 minutes, we are going to recommend that design. Now, I will go back to Excel.
Now, the server is operated by one clerk strategy number 1. So, the inter-arrival time follows an exponential distribution mean of 6. Service patterns follow an empirical distribution. So, the probabilities and the service timers are given. Here, we are going to make the lower random number and the upper random number. So, here is what I have done. You see that the probability is 0.24, and the service time is 2. So, 0.24 I have written 0.0 as the lower random number and the upper random number as 0.24.

Again, it started with 0.24, and the probability for a 3-minute service was 0.2. So, I have added that. So, 0.24 plus 0.2. So, it will be 0.04. The upper random number is 0.44. For the next one, the lower random number is 0.44. You see that the probability is 0.15. So, 0.44 plus 0.15, so you will get 0.59. So, what I have written is a lower random number and an upper random number from the given table. So, I have written on a continuous scale. Now, we will go back to our simulation model.

Look at the customer 1. Now you see the inter-arrival pattern, as I told you. So, the formula for generating random numbers which follow exponential distribution is minus mean B4 log of RAND. So, I have generated 1000 customers. So, the inter-arrival time for the first customer is the same. The service start time for the first customer is the same. Waiting time what is the waiting time? As usual, it is the difference between service start time minus and arrival time.

Then, see that in the service time, I have used the VLOOKUP function. So, VLOOKUP then RAND I selected A9 to C16. What is the A9? A9 is where the lower random number starts. What is the C16? Where the service time for the last value is given here. That is, C16 is 9.0. Then we know that the VLOOKUP standard function is RAND. Then we have to generate the range, 3. Why I have given 3? Because the value will be picked from the third column. That is why it is 3.

So, here, the service time is also generated for 1000 customers, then the completion time. The completion time is service start time + service time. So, the time in the system is completion time - arrival time. Now we will see for the second customer. For the second customer, as usual, I have generated the inter-arrival time. Then how did I get 11.9? So, 11.4 + 0.5 = 11.9. Now you see the logic for service start time. If C24 is greater than G23, what is the G23?

Completion time for the previous customer. The completion time for a previous customer here is 2.5. So, when we compare his 11.9 and 2.5, 11.9 is greater. So, that means he is coming late. So, the service start time is exactly 11.9. So there will not be any waiting time. Then, service time I randomly. Then, the completion time and the time in the system are calculated. Here, I am going to find out the idle time for the server. So, idle time is E24. If E24 = 0, then what is the E24?

If the waiting time of the customer is 0, the idle time for the server is D24 - G23; otherwise, it is 0. So, like this, I have found the idle time for all 1000 iterations. So, finally, if you see the summary statistics and number of people waiting, count in the waiting line column. How many people have waited? Probability of waiting. Most importantly see the average waiting time. The average waiting time is 9.83. But what is our requirement? It should be less than 2.5.

See the maximum waiting time. Sometimes, some customers wait 44.3 minutes. The number of people waiting longer than 6 minutes is 355. The probability of waiting greater than 6 minutes is 0.36. The number of people waiting is more than 10 minutes, 250 people. So, here is the first strategy: keep pressing F9 so that every time I see it, the average waiting time changes. So, most of the time the average waiting time is greater than 4.

But our requirement is that the average waiting time should be less than 2.5. So, the first design is not suitable for us. Now, we are going to explore the second design. Here, what is the second design? There is a single server, but two clerks are helping to provide the service. So, here, the arrival pattern is as usual. It is the same. So, here also the random numbers which I have generated in a continuous scale: 0.00, 0.20, then 0.20, then 0.55 up to 1, then I have entered the service time 1, 2, 3, 4, 5.

Now, what is the difference between the previous case and this case? Previously the service time was longer because there is only one clerk. But now the service time is decreased. So, when the service time decreases, we will see the average waiting time. So, inter arrival time is as usual it is same. Because it follows an exponential distribution, including arrival time, service start time, waiting time, service time, completion time, and time and system.

Similarly, for the second customer. So, arrival time is taken. So, you see the service start time. The second customer is arriving at 2.2, but the completion time is 3. So, he has to wait. How much time he has to wait? So, 3 - 2.2, which is minutes. So, the service start time is randomly generated from the given empirical distribution by using the VLOOKUP function. So, the completion time is service start time + service time. So, the time in the system is completion time minus arrival time.

Now look at the idle time. If the waiting time is 0 for the customer, the idle time for the server is D21 - G20; otherwise, it is 0. Now, I will go back to the summary statistics for the second design. For the second design, the number of people waiting is 409. The probability of waiting is 0.41. See the average waiting time. So, the average waiting time is 2.35 minutes. Suppose I keep on pressing F9. The average waiting time is 1 minute, Rohan. What was our target? 2.95.

So, the average waiting time is exceeding most of the time. Sometimes, it is within 2.5 minutes. But it is very close to our target. But many, time it is exceeding. So, the second design is somewhat more suitable. So, if I press F9, I see that 2.35, 2.56, 2.12, 2.57, and 2.63. So, most of the time, it exceeds 2.5 minutes. But sometimes it is meeting our requirements. So, this is the second design. Here is the conclusion from the second design.

Sometimes, we are meeting; most of the time, we are not able to meet the target, but sometimes, we are meeting the target. So, the second design is very close to our expectations. Now, we are going to see the third design. What is the third design? Two servers are operated by one clerk each. Here, the inter-arrival follows exponential distribution with a mean of 6. And service time follows an empirical distribution. This problem is similar to having 2 ATMs because here, there are two servers is there.

So, two cars can enter the shop to get their service. So, for the customer here, we will also be following the same assumption. What is the same assumption? The customer will go to server 1. After that, if any customer enters, he will enter the server, which is idle for a longer time. So, it is a 4.6. So, the arrival time is also 4.6. Service start timer 4.6. There is no waiting time. Look at the service time. We have taken service time from using the VLOOKUP function.

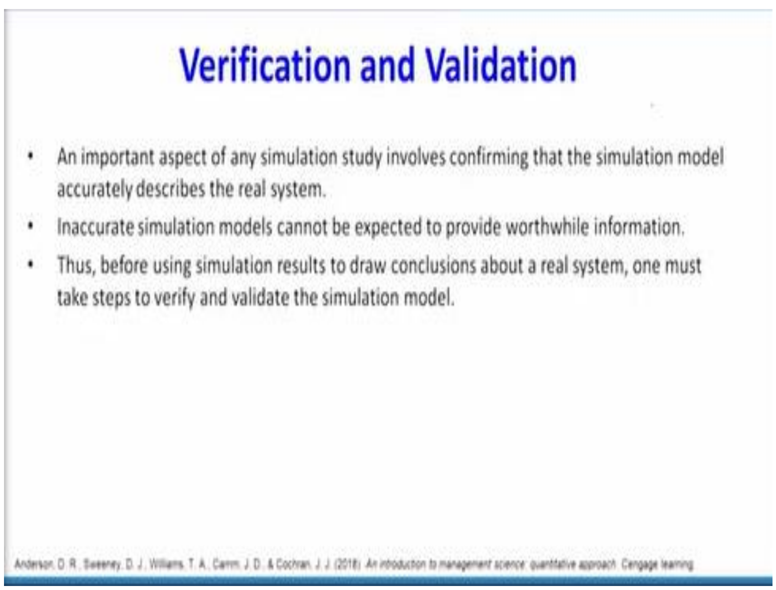
VLOOKUP up RAND A92 C16. The logic of the VLOOKUP function is a random number will be generated. That random number will be traced between A9 to C16. If the random number matches between A9 and C16, that corresponding value in column 3 will be picked as a service time. That is why we got 2. The randomly generated number may be between 0 to 0.24. So, completion time, time in the system, and time available for server 1.

Because we have the assumption that customer 1 will enter server 1, now look at the second customer. The second customer inter-arrival time is 10.7. His arrival time is 15.3 because server 2 is already free. So, he will enter into the system to get the service. So, there is no waiting time. Service time will be picked from the empirical distribution completion time. So, now the completion time, how we are updating the completion time?

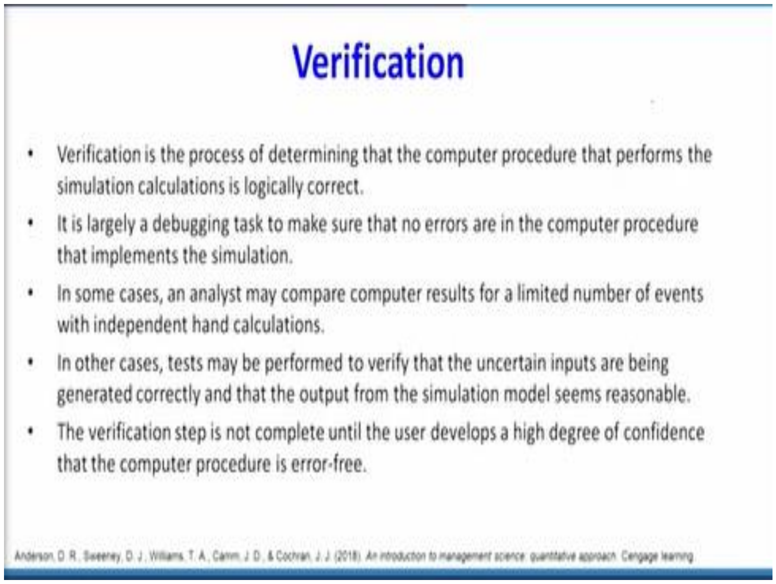
When you look at the previous customer ATM 2 is free now. So, now the completion time will be updated below ATM 2. So, it will become 18.3. The 6.6 will be carried to the next customer. The logic is the same as our 2 ATM problem. So, I have repeated this to 1000 customers. I have found different summary statistics. Number of people waiting, probability of waiting. Look at the average waiting time. The average waiting time is 1.65.

If I press F9, I see that most of the time, the average waiting time is less than 2.5 minutes. So, what are we concluding? The third design is the most suitable design. That means having two servers will help us to achieve our requirement of a waiting time of less than 2.5 minutes. Now, I will go back to the presentation.

So, we have seen what is more important is the average waiting time of 6.27 here, 2.86, here it is 1.7. So, we are recommending the third design. What is the third design? 2 servers operated by one clerk each.

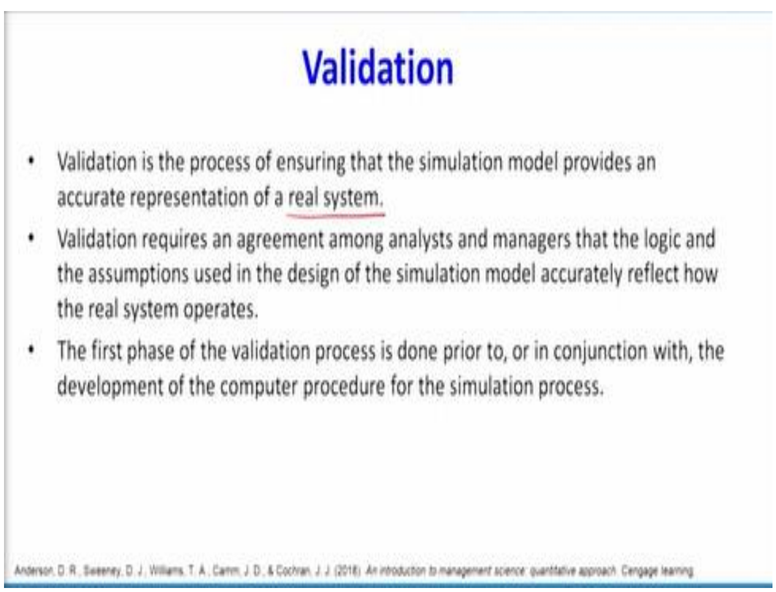


Now, we will go for simulation verification and validation. An essential aspect of any simulation study involves confirming the simulation model accurately describes the real system. Inaccurate simulation models cannot be expected to provide worthwhile information. Thus, before using simulation results to draw conclusions about the actual system, one must take steps to verify and validate the simulation model.



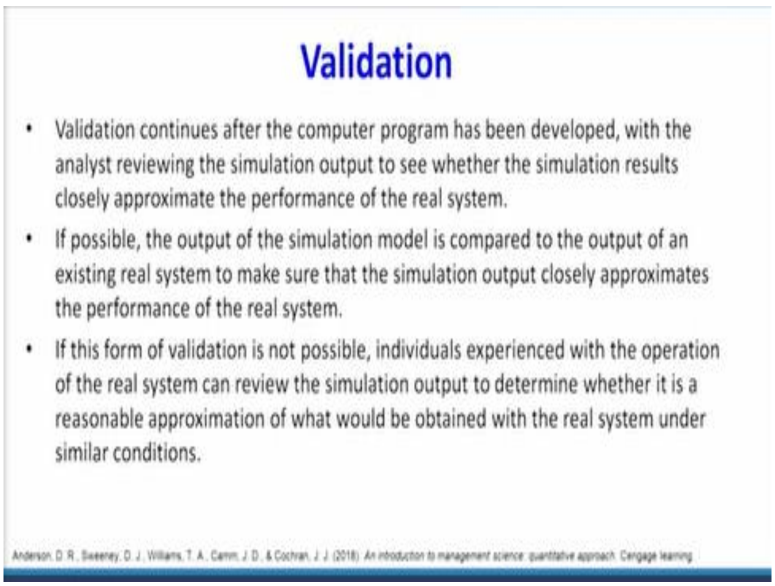
First, we will see what the verification is. Verification is the process of determining that the computer procedure that performs the simulation calculations is logically correct. It is largely a debugging task to make sure that no errors are in the computer procedure that implements the simulation. In some cases, an analyst may compare computer results for a limited number of events with independent hand calculations.

In other cases, the test may be performed to verify that the uncertain inputs are being generated correctly and that the output from the simulation model seems reasonable. The verification step is not complete until the user develops a high degree of confidence that the computer procedure is error-free.



Next, we will see what validation is. Validation is the process of ensuring that the simulation model provides an accurate representation of the real system. The verification is just debugging. Here, the validation is whether the simulation model exactly represents the real system or not. So, validation requires an agreement among analysts and managers that the logic and the assumptions used in the design of the simulation model accurately reflect how the real system operates.

The first phase of the validation process is done prior to or in conjunction with the development of a computer procedure for the simulation process.



Validation continues after the computer program has been developed, with the analyst reviewing the simulation output to see whether the simulation result closely approximates the performance of the real system. If possible, the simulation model's output is compared to the output of an existing real system to ensure that the simulation output closely approximates the performance of the real system.

If this form of validation is not possible, individuals experienced with the operations of the real system can review the simulation output to determine whether it is a reasonable approximation of what would be obtained with the real system under similar conditions. In this lecture, I explained the introduction of 1 more ATM to achieve the service requirement. What is the service requirement in our problem?

The service requirement is that the average waiting time should be less than 1 minute. So, we have found that by introducing 2 ATMs we are able to achieve the requirement. Also, I have taken another problem to explain the concept of a two-server system waiting line problem to meet the service requirement. Finally, I have explained the concept of verification and validation for the simulation model. Thank you.