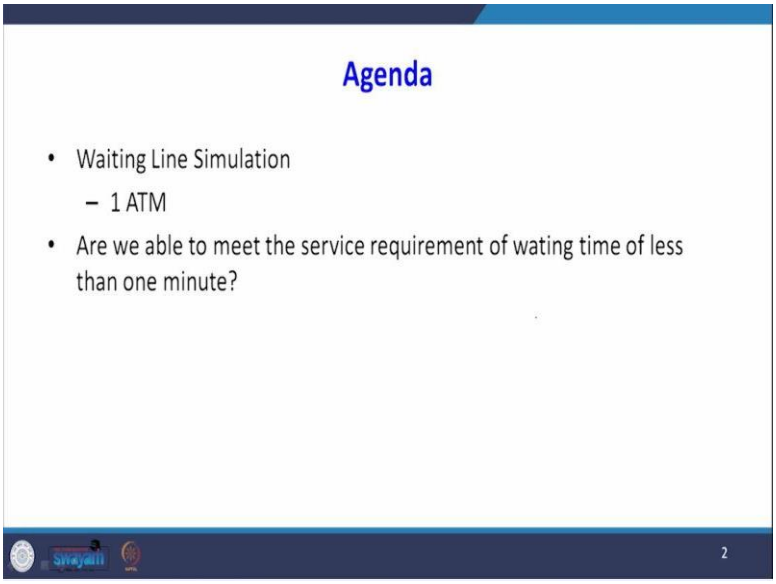
Decision Making with Spreadsheet
Prof. Ramesh Anbanandam
Department of Management Studies
Indian Institute of Technology, Roorkee

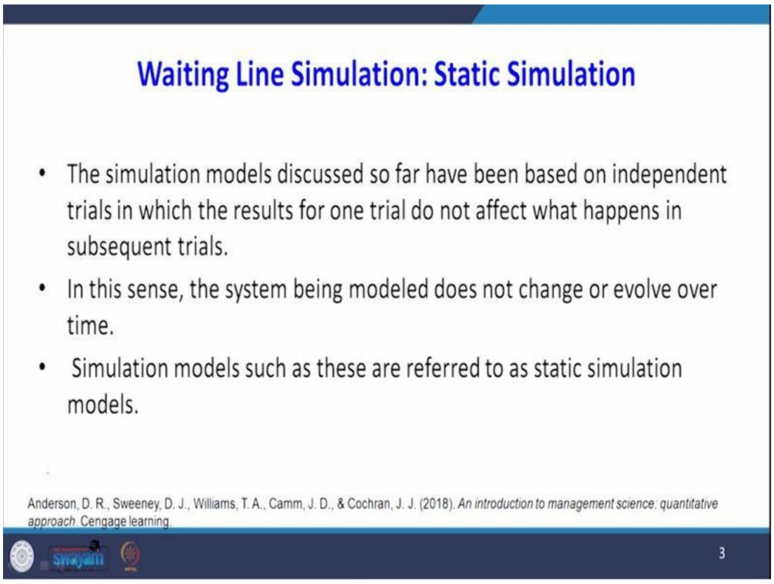
Lecture - 44
Simulation - III

Dear students, in the previous lecture, I have discussed about inventory simulation. In this lecture, I am going to examine another interesting problem, which is called waiting line simulation.



So, the agenda for this lecture is to assume that a bank has only 1 ATM when there is only one ATM, so the bank manager wants to fix certain service requirements. What is the service requirement? Meeting the service requirement of the waiting time for a customer should be less than one minute. What is the service requirement? The bank manager or the vice president of the bag decides that my customer should not wait more than one minute.

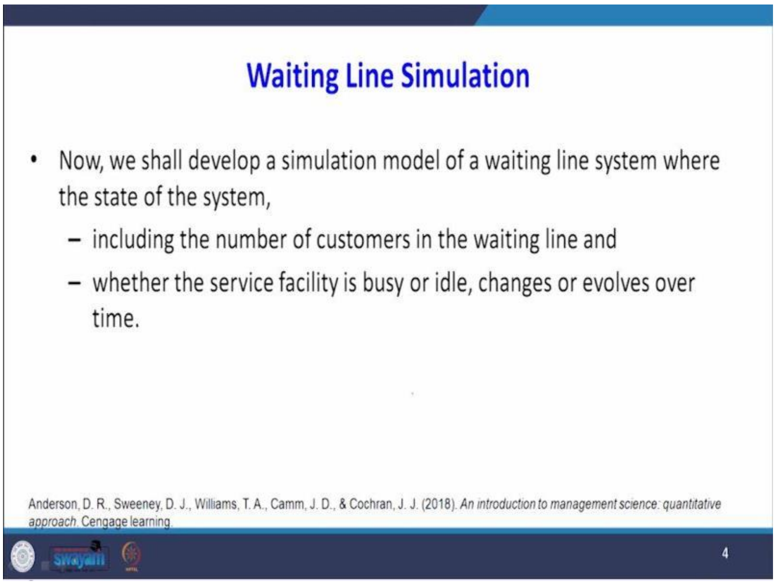
For that purpose, we have only 1 ATM; if I have only 1 ATM, can I meet this service requirement? We are going to provide that answer with the help of a simulation model.



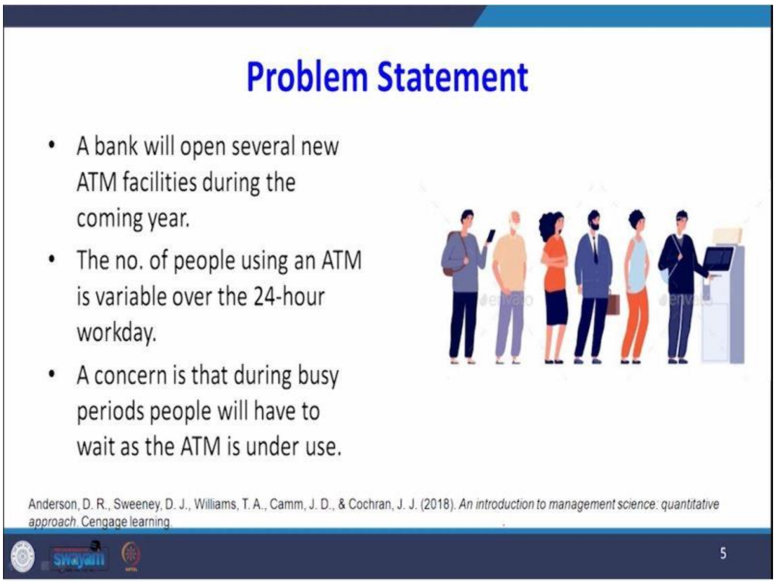
Waiting line simulation, static simulation: The simulation models discussed so far have been based on independent trials in which the result for one trial does not affect what happens in the subsequent trials. For example, in the previous class, we discussed the inventory simulation, so the result of one trial will not affect the subsequent trial. In this sense, the system being modeled does not change or evolve over time.

So, time is not an important factor there; it is independent of time. Simulation models such as these are referred to as static simulation models. What is the meaning of the static simulation model? The outcome of one trial will not affect the outcome of another trial, but in this class, we are going to discuss dynamic simulation. What is the meaning of this dynamic simulation?

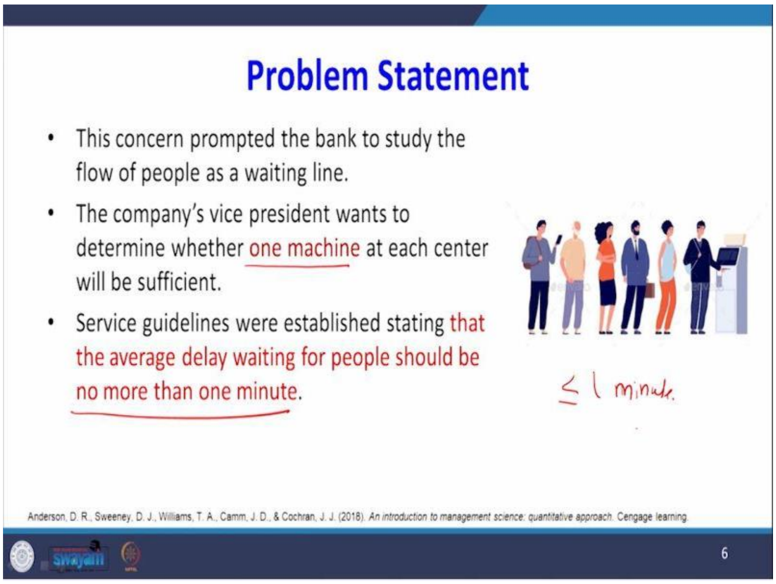
The outcome of one trial will affect the outcome of another trial; the model has been discussed. So far, it is called Monte Carlo simulation in the Monte Carlo simulation. It is independent of time, but in this class, we are going to discuss dynamic simulation.



So, waiting line simulation. Now, we shall develop a simulation model for a waiting line system where the state of the system, including the number of customers in the waiting line, and whether the service facility is busy or idle or changes or evolves over time. So, we are going to discuss the dynamic simulation.



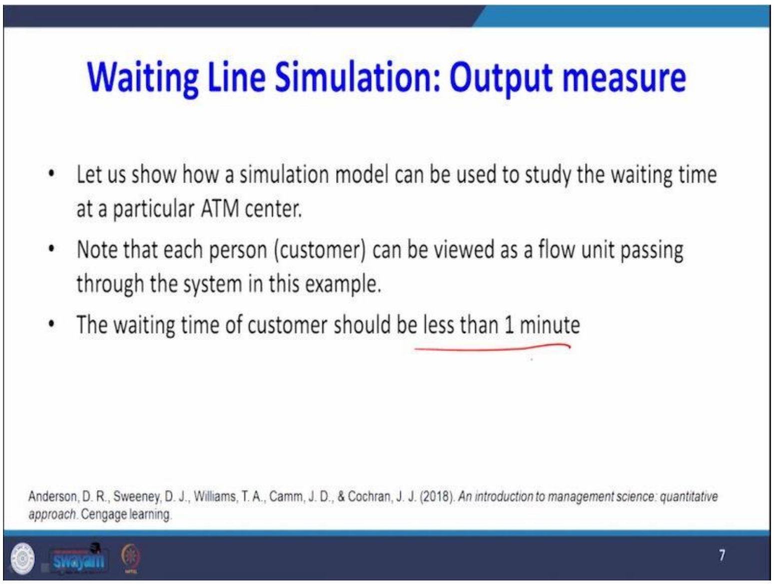
What is the problem? As I discussed in the introduction, a bank will open several new ATM facilities during the coming year; see that there is a bank the people are waiting for. The number of people using an ATM is a variable over 24 hour work day. A concern is that during busy periods, people will have to wait as the ATMs are under use.



This concern prompted the bank to study the flow of people in a waiting line. The company's vice president wants to determine whether one machine at each center will be sufficient. You might have seen in front of the bank, there may be an ATM centre will be there sometimes, there will be one mission or two missions. At present, assume that the bank has only one mission.

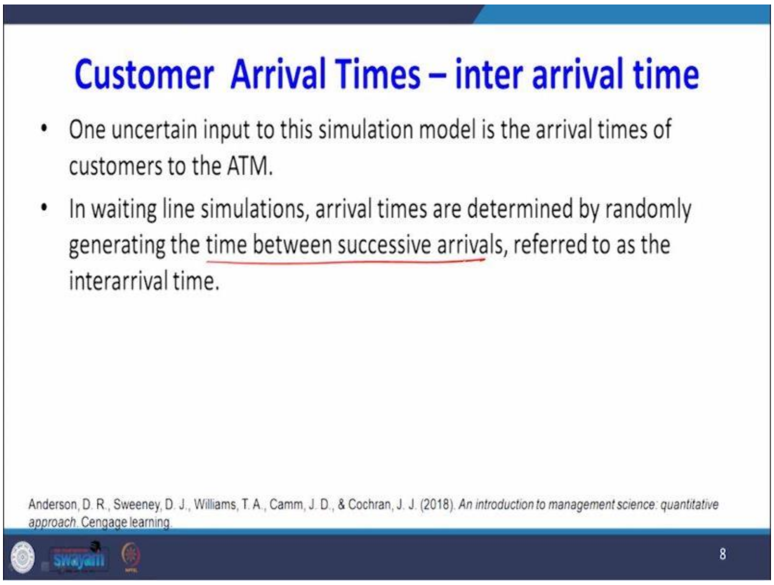
So, the service guidelines were established stating that the average delay waiting for people should be no more than one minute. So, the service requirement is the customers of that bank should not wait more than one minute, which means they can wait a maximum of one minute. But we have only one machine. So, the question is whether we are able to achieve this service requirement.

That is, the waiting line should be less than one minute with the help of this one machine. So, that is what we are going to suggest.

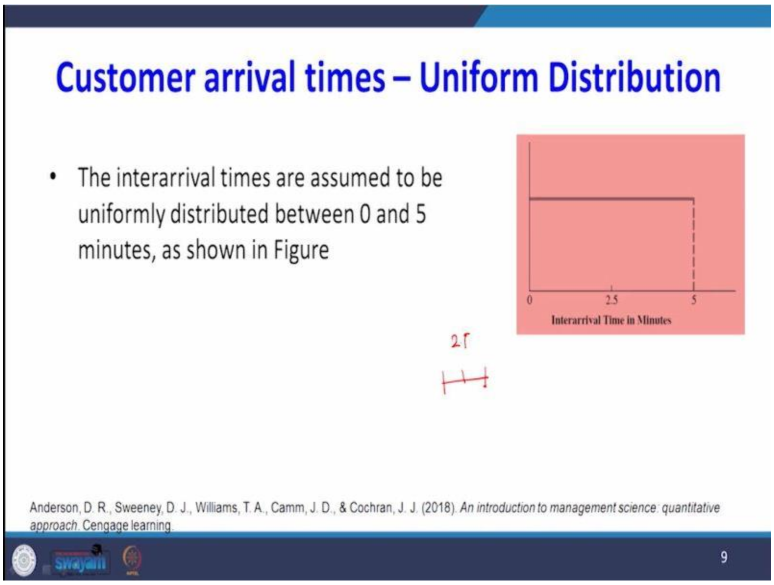


So, let us show how a simulation model can be used to study the waiting time at a particular ATM center. Note that each person, the customer, can be viewed as a flow unit passing through the system in this example. What is the requirement? The waiting time of customers should be less than one minute, but we have only 1 ATM. Now we are going to test whether this service requirement can be achieved or not.

If it is not possible to achieve what we have to suggest we should go for two ATMs in the same room.

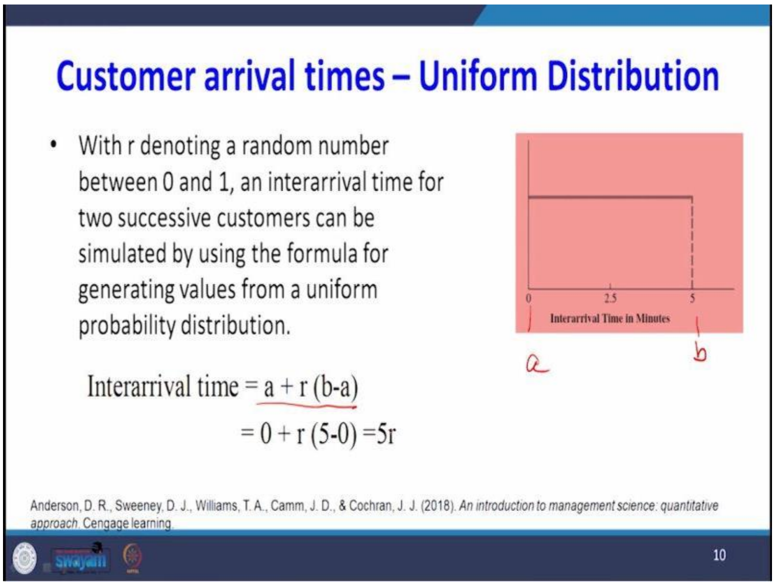


Now, the customer arrival time, which is called inter-arrival time, is one of the inputs for the simulation model. So, one uncertain input to the simulation model is the arrival time of customers to the ATM. In waiting line simulations, arrival times are determined by randomly generating the time between successive arrivals, referred to as the interarrival time. So, we are going to generate interarrival times.



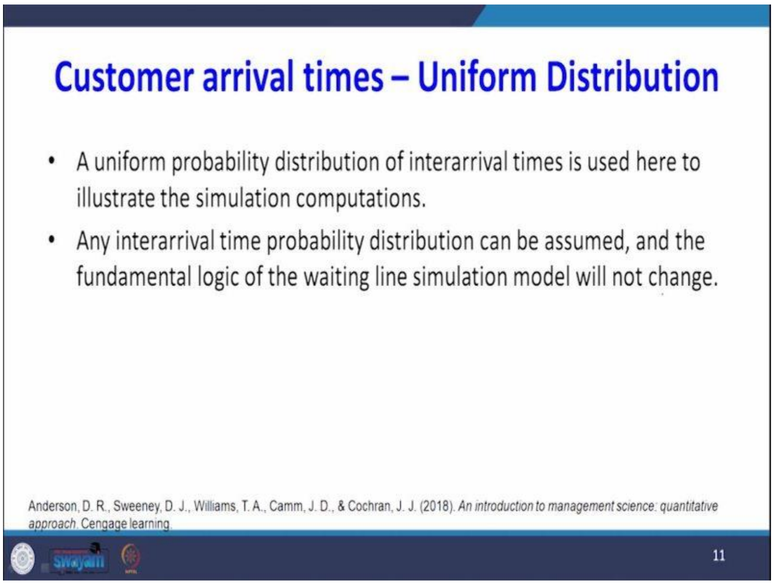
So, we are going to assume that the inter-arrival time follows a uniform distribution. So, the inter-arrival times are assumed to be uniformly distributed between 0 and 5 minutes, as shown in this figure. See arrival pattern follow uniform distribution, say minimum 0 maximum 5, and the average inter-arrival time is 2.5 minutes. That means the arrival between one customer and another customer the average time is 2.5.

Sometimes the customer may come in one minute also in 4 minutes also but the average intra-arrival time is 2.5 minutes.



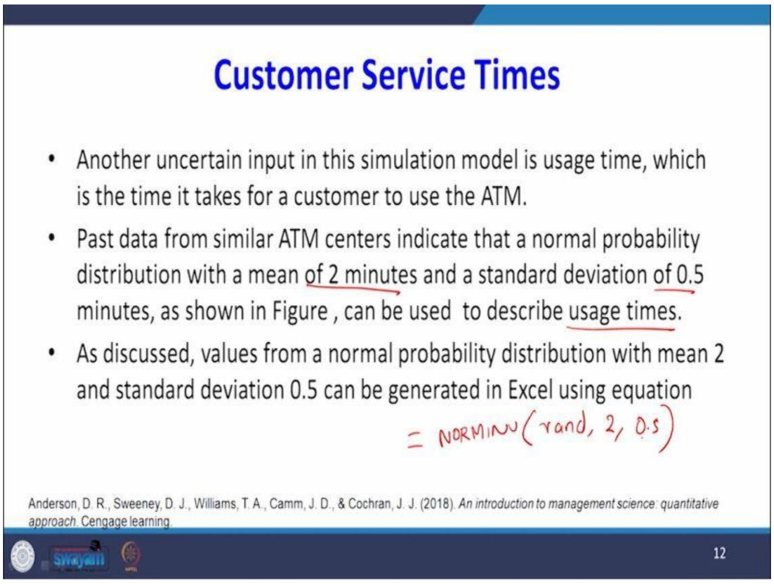
With r denoting a random number between 0 and 1, an interval time for 2 successive customers can be simulated using the formula. This formula we have already shown to you. I will show you what is the formula simulated for generating values from a uniform probability distribution. So, what is that formula? That formula is a + r (b – a). Here, a is this 0, b is this 5, and r is a random number between 0 and 1. So, when you substitute a = 0 .

$$\text{Interarrival time} = a + r\,(b-a)$$
$$= 0 + r\,(5-0) = 5r$$



A uniform probability distribution of inter-arrival time is used here to illustrate the simulation computations. Any inter-arrival time probability distribution can be assumed. So, instead of uniform distribution, you can assume the arrival pattern may follow Poisson distribution also, or it can be exponential. But in this problem, we have taken that inter-arrival time follows a uniform distribution, and the fundamental logic of the waiting line simulation model will not change.

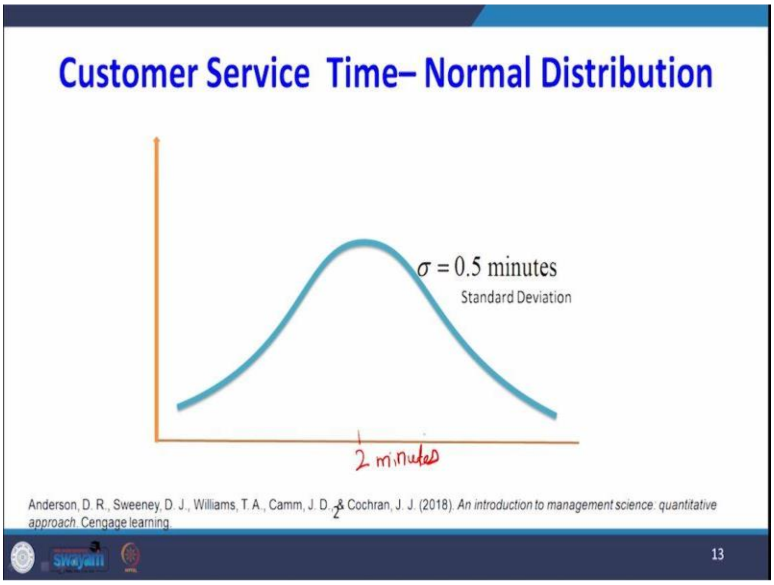
You can assume any type of distribution for the arrival pattern. So, the logic of simulation will not change. But in this problem, we are going to assume that the arrival inter-arrival pattern follows a uniform distribution



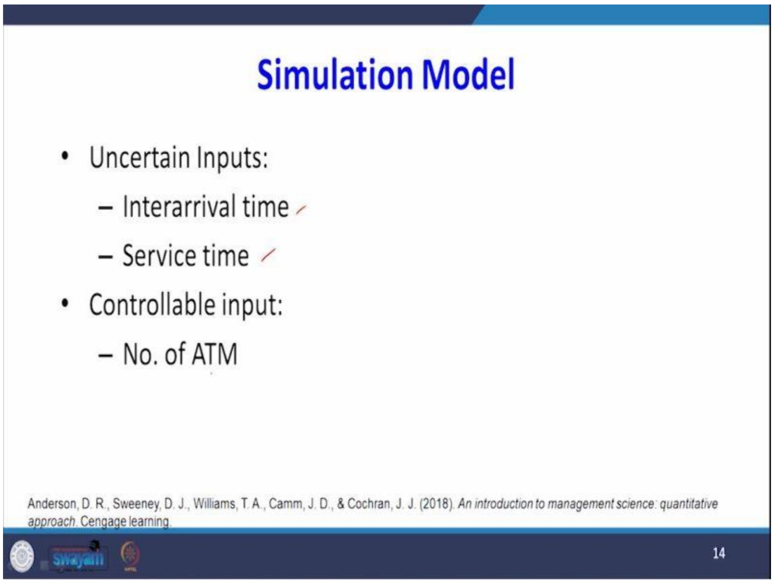
The next input is customer service times; here, the customer service time means that when a customer enters the ATM center, some people may withdraw money. Some people may check their balance; some people may check their change their PIN. Like that, there are different services they can take. So, here, another uncertain input in the simulation model is usage time, which is the time it takes for a customer to use the ATM.

There may be different purposes; some people may use it for withdrawing money, some people for checking the balance, and so on. But that time is going to follow normal distribution. So, past data from similar ATM centers indicate that a normal probability distribution with a mean of 2 minutes and a standard deviation of 0.5 minutes, as shown in the figure that the figure is there in the next slide, can be used to describe the usage time, that is, service time.

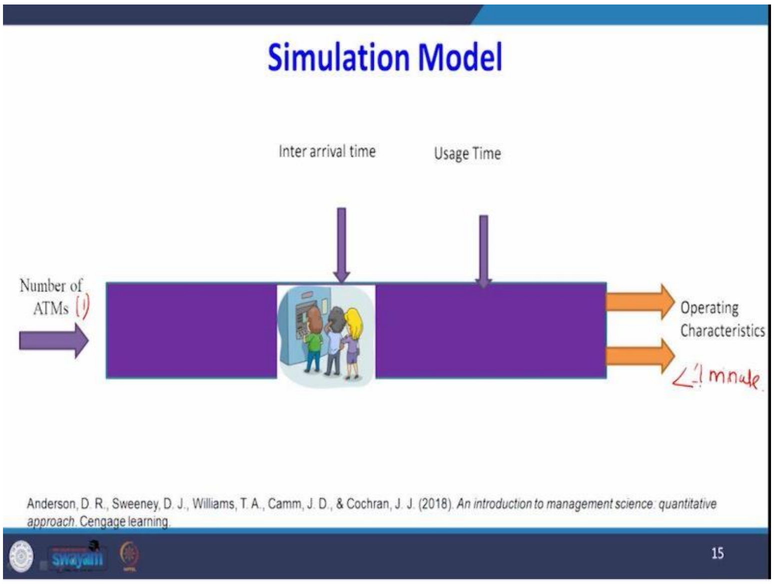
As discussed, the value from a normal probability distribution with a mean of 2 and a standard deviation of 0.5 can be generated in Excel. So, what for the equation we use? Also we have discussed NORM.INV rand, then mean, then standard deviation. So, using this function, we can generate a number that follows the normal distribution.



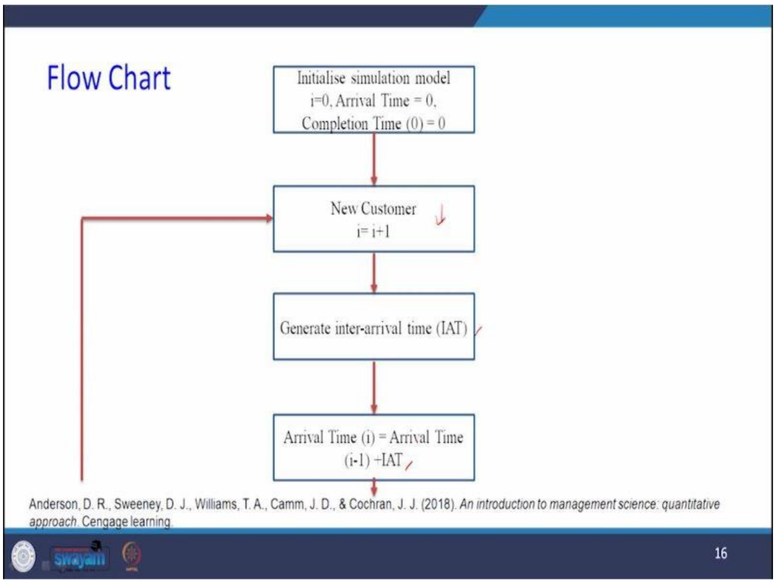
So, here the mean is 2 minutes, and the standard deviation is 0.5 minutes. What is this? The customer service time.



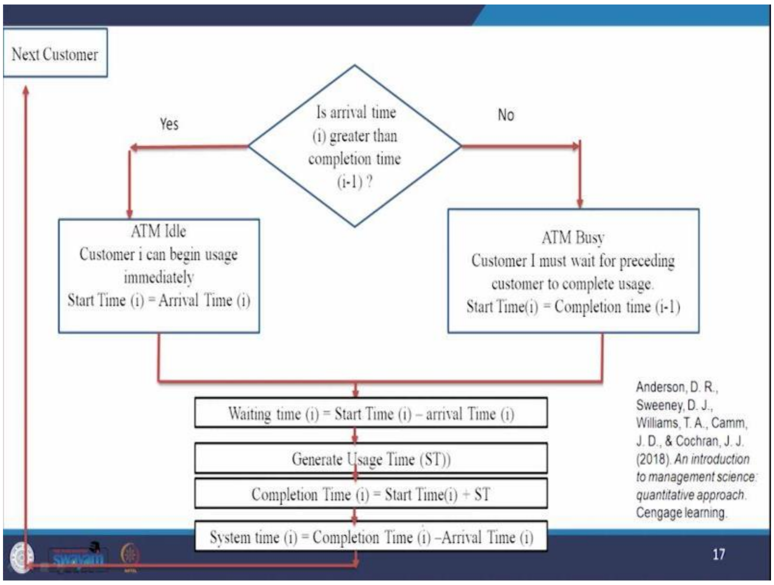
So, in the simulation model, what are the uncertain inputs of inter-arrival time and service time? Inter-arrival time follows a uniform distribution, and service time follows a normal distribution. Here, the controllable input is the number of ATMs. Now, we will work out this model if there is only one ATM, whether the service requirement of waiting time less than one minute can be achieved or not. If it is not achievable, then we suggest 2 ATMs.



Now, the simulation model: What is the simulation? There are a number of ATMs, but at present, we have only one ATM. What are the other inputs? Inter-arrival time is one input, and the usage time is another input. What are the operating characteristics? So, the average waiting time should be less than one minute. That is one of the operating characteristics we are going to test by having one ATM with this operating characteristic. That is, an average waiting time of less than one minute can be achieved or not.

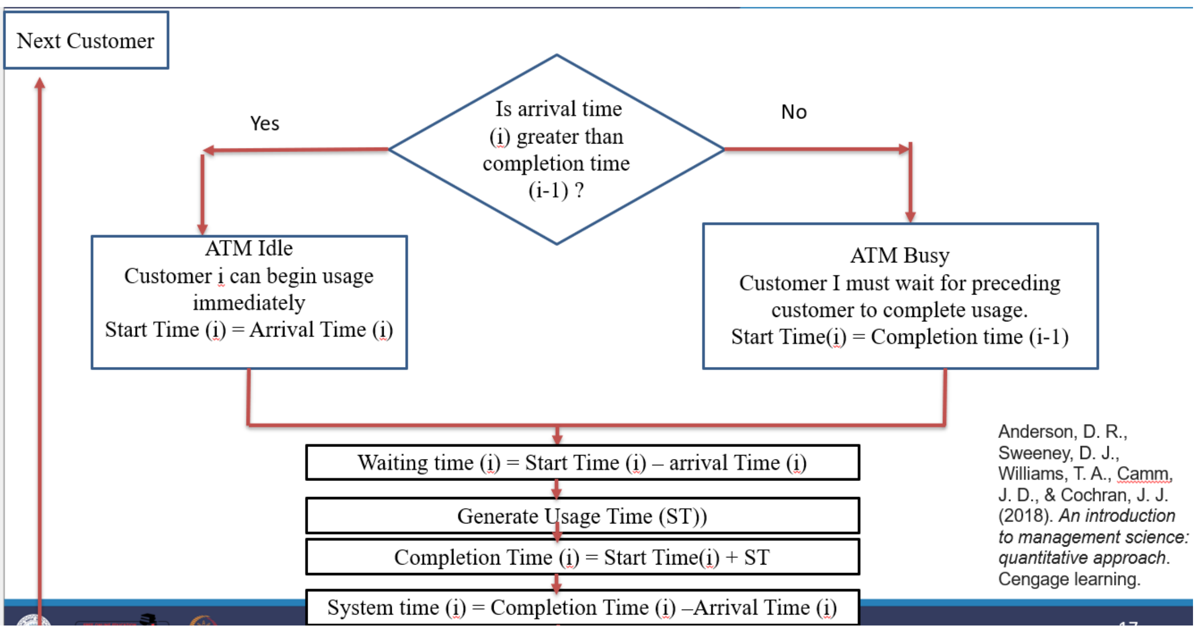


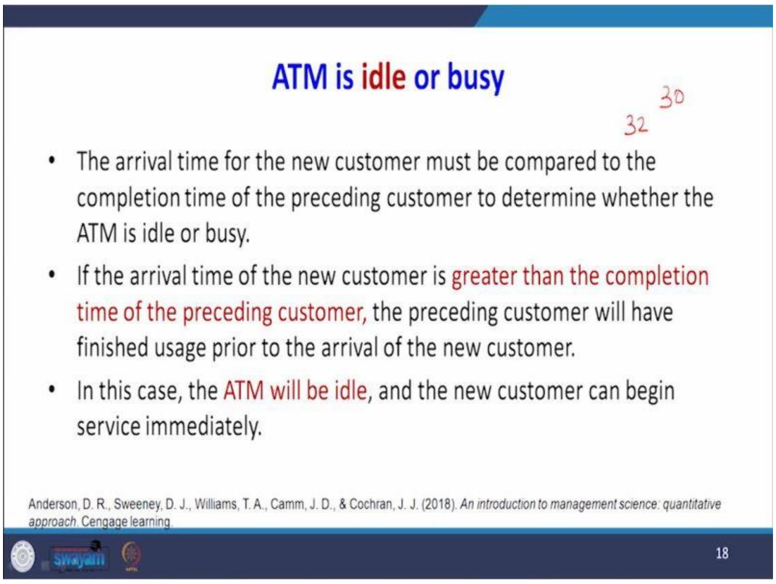
So, the flowchart shows first, we initialize the simulation model i = 0, arrival time is 0, and completion time is 0. Once a new customer enters then, we generate inter-arrival time. Then how are we getting arrival time?



Then, if the arrival time is greater than the completion time, if it is yes, the ATM is idle. So, customer, I can begin usage immediately. When will this happen? If the arrival time is greater than the completion time. Because one customer has already left, the ATM is free when that situation occurs if the arrival time is greater than the completion time. If it is not, the ATM is busy, which means what customer ‘i’. This ‘i’ must wait for preceding customers to complete the usage.

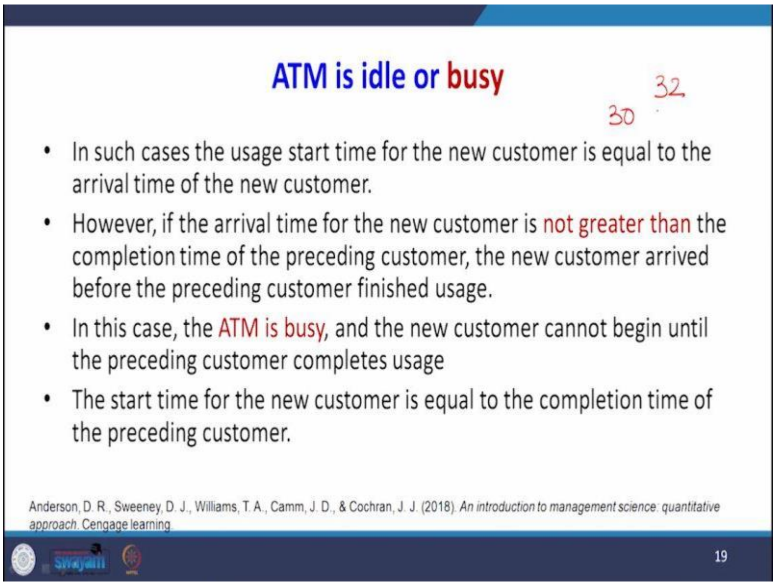
Then, we will find the waiting time, starting time - and arrival time, and then you will find the generated usage time. What is the completion time? Starting time + usage time. What is the system time? Completion time - arrival time. This is the flowchart for the simulation.





ATM is idle or busy. How do you know that the ATM is ideal or busy? The arrival time for the new customer must be compared to the completion time of the preceding customer to determine whether the ATM is idle or busy. See if the arrival time of the new customer is greater than the completion time of the preceding customer. The preceding customer will have finished usage prior to the arrival of the new customers.

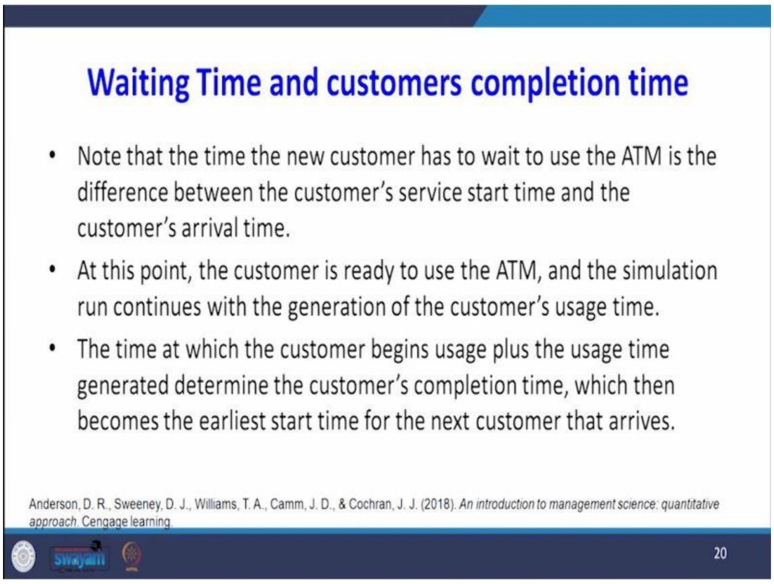
In this case, the ATM will be idle, and the new customer can begin service immediately. For example, in the 30th minute, one fellow uses the ATM. So, the next arrival time is, say, 32; what has happened? After 30th minutes, the ATM is free, but the next customer enters only 30 seconds minutes.



So, the 2 minutes this ATM will be idle, in such cases, the usage start time for new customers is equal to the arrival time of the new customer, which means when the ATM is idle however if the arrival time for the new customer is not greater than the completion time of the preceding customer. The new customer arrived before the preceding customer finished the usage. In this case, the ATM is busy, and the new customer cannot begin until the preceding customer completes the usage.

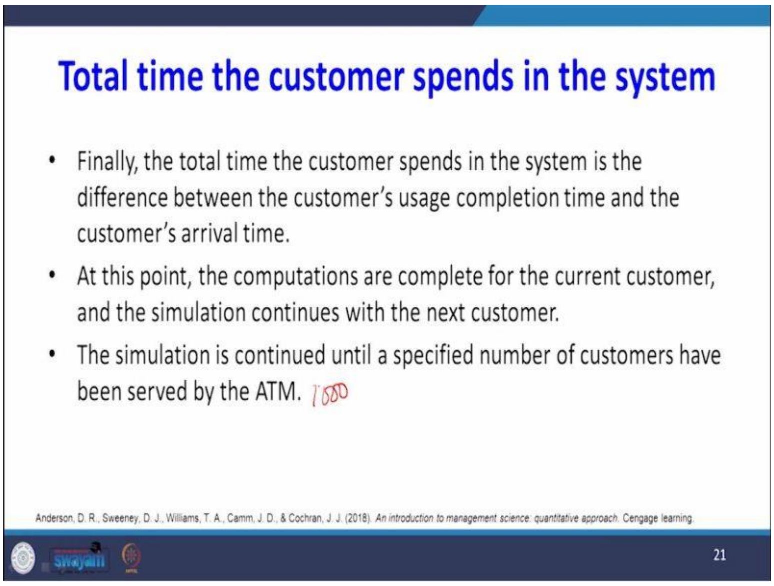
So, that means what? Suppose say thirty-second minutes, somebody has used the ATM. The next customers arrive on 30th minutes because up to 30 second minutes the ATM is busy. But the next customer has come into 30th minutes in a continuous scale, so 2 minutes the new customer must wait. That means that in 2 minutes, the ATM is busy. So, in that the start time for the new customer is equal to the completion time of the preceding customer.

Even though he comes in the 30th minute, he can start getting the service only in the 32nd minute because then only that ATM will be busy there, and the ATM will be free.



Now, I am going to explain what is the waiting time and the customer completion time. Note that the time the new customer has to wait to use the ATM is the difference between the customer's service start time and the customer's arrival time. At this point, the customer is ready to use the ATM, and the simulation runs continuously with the generation of the customer's usage time. The time at which the customer begins usage + the usage time generated determines the customer's completion time.

Which then becomes the; earliest start time for the next customer that arrives. So, this concept with the help of a numerical example I will explain in the next slide.



The total time the customer spends in the system. Finally, the total time the customer spent in the system is the difference between the customer's usage completion time and the customer's arrival time. At this point, the computations are complete for the current customer, and the simulation continues with the next customer. The simulation is continued until a specified number of customers have been served by the ATM. In our example, we are going to consider 1000 customers.

Customer	Inter arrival Time	Arrival Time	Service Start Time	Waiting Time	Usage Time	Completion time	Time in system
1	1.4	1.4	1.4	0.0	2.3	3.7	2.3
2	1.3	2.7	3.7	1	1.5	5.2	2.5
3	4.9	7.6	7.6	0	2.2	9.8	2.2
4	3.5	11.1	11.1	0	2.5	13.6	2.5
5	0.7	11.8	13.6	1.8	1.8	15.4	3.6
6	2.8	14.6	15.4	0.8	2.4	17.8	3.2
7	2.1	16.7	17.8	1.1	2.1	19.9	3.2
8	0.6	17.3	19.9	2.6	1.8	21.7	4.4
9	2.5	19.8	21.7	1.9	2.0	23.7	3.9
10	1.9	21.7	23.7	2.0	2.3	26.0	4.3
Total	21.7			11.2	20.9		32.1
Averages	2.17			1.12	2.09		3.21

Dear students, now I will explain the concept of simulation with the help of some numerical examples. So, I have considered 10 customers. The first column is 10 customers, and the second column is inter-arrival time, which I have randomly generated and follows a uniform distribution. The next column is arrival time, the arrival time I have written on a continuous scale. After that service starts time, there is waiting time.

What is waiting? This is the difference between the service start time and the arrival time. Then, usage time 2.3 that I have randomly generated, which follows a normal distribution, then completion time, what is the completion time? Completion time is service start time + usage time that is 1.4 + 2.3 is our completion time. Then what is the time in the system? The time in the system is completion time - arrival time. So, 3.7 - 1.4, that is our time in the system.

Now, the first customer enters the ATM, he will use up to 3.7 minutes. Now, for the second customer, we have the number 1.3, which we have randomly generated. So, 1.3 + 1.7, he entered the system, and his arrival time on a continuous scale is 2.7. You see, the first customer still used up to 3.7 minutes, but the second customer arrived in 2.7 minutes. Here is what he has to do. The second customer has to wait.

So, even though he arrives at 2.7 minutes, his service will start at only 3.7. So, what is the waiting time? The service start time - arrival time is 1. So, his service started at 3.7 minutes, he will use it for 1.5 minutes, so he will complete the service at 5.2 minutes, which is 3.7 + 1.5. What is the time in the system? Completion time - arrival time that is 5.2 - 2.7. Then, the third customer inter-arrival time is 4.9.

So, the arrival time on a continuous scale is 7.6. You see that the first customer left the ATM at 5.2 minutes, not the first customer's second customer. So, the third customer is entering at 7.6. So, that means that from the 5.2 minutes onwards, the ATM is idle. So, the third customer need not wait directly. He can get the service in 7.6 minutes. So, there would not be any waiting time. He will use it for 2.2 minutes and he will complete the service in 9.2 minutes.

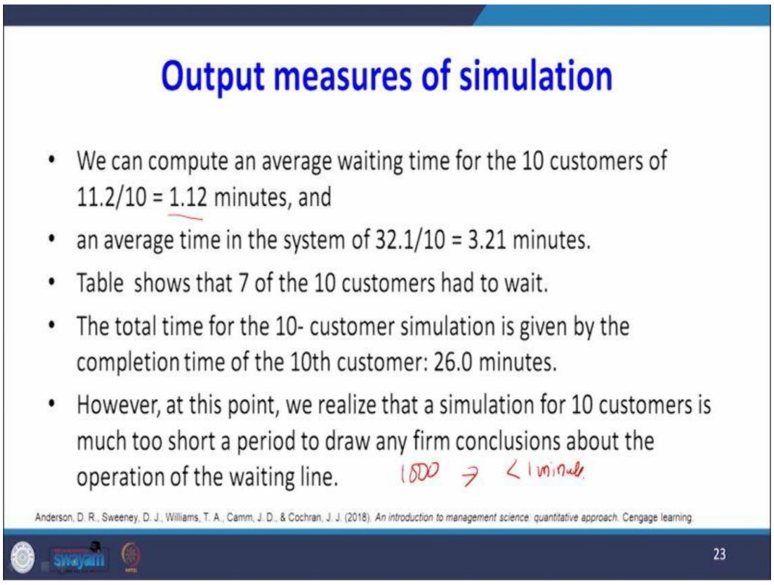
What is the time in the system? 9.8 to – 7.6 that is 2.2. Now, you see the fourth customer inter-arrival time is 3.5, so 3.5 + 7.6, it is 11.1 minutes. But the ATM is free from 9.8 minutes onwards. So, he will directly enter into the service. So, the service start time is 11.1 minutes, so there would not be any waiting time. He will use it for 2.5 minutes. When the ATM will be free? Up to 13.6, after 13.6 minutes, the ATM will be free.

So, the time in the system is 13.6 - 11.1, which is 2.5. Now, you see the fifth customer, the fifth customer's arrival time on a continuous scale is at 11.8 minutes. How did we get this 11.8? 0.7 + 11.1. But the fourth customer still is using the ATM up to 13.6. So, this fellow comes early but he will get the service only after 13.6 minutes. So, how much time does he have to wait? 13.6 - 11.8-minute 1.8 minutes for customer 5 to wait.

So he will use it for 1.8 minutes so that he will come out of the ATM in 15.4 minutes. So, how much time is he spending in the system? 15.4 - arrival time that is 3.6. Now, you see the sixth customer entering, or he is arriving on the continuous scale of 14.6 minutes, but the ATM will be free only after 15.4. So, the sixth customer may get the service only at 15.4 minutes. So, he has to wait 4.8 minutes, which is 15.4 – 14.6; he will use it for 2.4 minutes.

So, the completion time will be 17.8, the time in the system is 17.8 - and the arrival time will be 3.2. I have repeated this for all 10 customers. Now, you see that the total waiting time is 11.2 minutes, and there are 10 customers. So, the average waiting time is 1.12 minutes, but what was our requirement? The average waiting time should be less than 1 minute. For 10 customers itself, the average waiting time is 1.12 minutes.

So, how much time is it taking? 26 minutes for all the 10 customers. In the next slide, I will explain the output measures of the simulation.



So, we can compute an average waiting time for 10 customers. What is that total waiting time is 11.2 divided by 10 it is 1.12 minutes. Average time in the system, so 32.1. How did we get this 32.1? You see that this 32.1 divider by 10 so the average time in the system is 3.21 minutes.

So, each customer is spending 3.21 minutes in the system. The table shows that seven of the 10 customers have to wait. How it is? How many people are waiting? This is 1, 2, 3, 4, 5, 6, 7.

So, out of 10 customers 7 customers are waiting. The total time for the 10-customer simulation is given by the completion time of the 10th customer. So, what is the completion time of the 10th customer, 26 minutes? This one is 26 minutes? However, at this point, we realize that a simulation for 10 customers is much too short a period to draw any firm conclusions about the operation of the waiting line.

So, what I am going to do in excel instead of 10 customers, I am going to repeat for 1000 customers. After conducting a simulation for a thousand customers, then I am going to find out what is the average waiting time. I am going to test whether it is less than 1 minute or not. Now, I will go back to the excel.

Now, I am going to explain the ideal time of an ATM. You see that what logic I have used for finding the ideal time of ATM if E21 = 0, that means if the waiting time is 0, then the idle time the waiting time 0 means that the customer directly enters the ATM to get the service but how much time that ATM is idle. So, that is C21, and C21 is your arrival time - the completion time of the previous customer, that is, the G20. See that the G20 is 15.4 and the arrival time of the sixth customer is 12.5.

So, for many minutes, the ATM is idle. So, that was the logic for finding the idle time. So, I have replicated this for 1000 customers. So, I will go back, I will come down, I will show all 1000 customers. Now, you see the logic for a number of people waiting, so you see that count if. If E16 to E1015, if it is greater than or equal to 0, that is the waiting time is positive. So, that many people are waiting. So, currently, 642 people out of 1000 are waiting.

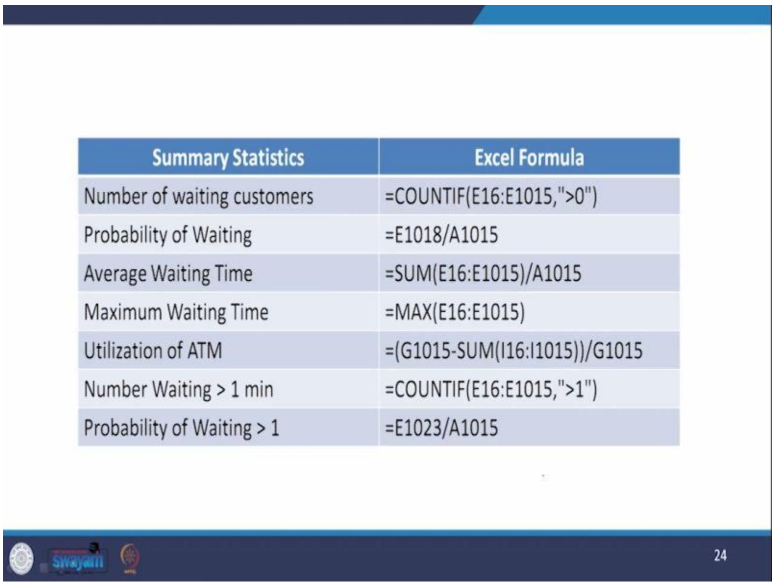
So, what is the probability of waiting 642 divided by 1000? So, 60% of the people are waiting, then what is the average waiting time? So, the average waiting time is 2 minutes. Do you see that? Now, you should remember here what the expectation is. What is the expectation of the management the average waiting time should be less than one minute. But when we replicated for 1000 customers, the average waiting time was 2 minutes.

Then the maximum waiting time how we got a maximum waiting time to see a maximum of E16 to E1015. Then, utilization of ATM, now you see the formula for utilization of ATM is the total time what is the total time that is G1015 that is 2509 minus the sum of all idle time. So, that means that much time the ATM is busy to divide the total time. So, 80% of the time, this ATM is busy. The next number of people waited more than 1 minute.

So, now I have counted how many times the people are waiting more than 1 minute, which means 480 customers out of 1000 customers, 480 customers are waiting more than one minute, then the probability of waiting for 1 minute is equal to so that is E1023. What is the E1023? The number of people waiting for 1 minute divided by 1000. Here, we found the probability of 84,000 customers.

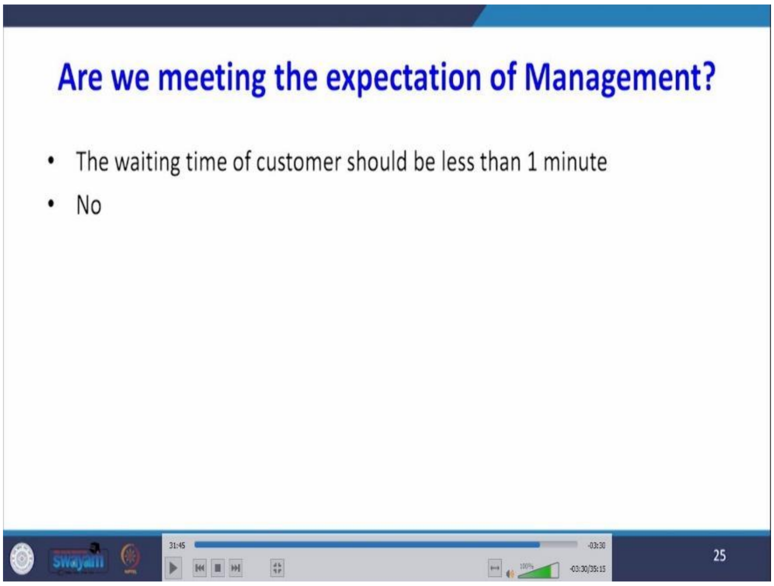
In another way also, we can find out the number of people waiting more than one minute delivered by a total number of waiting that is total number of people waiting, which is 480. So, we can write this way a number of people are waiting more than 1 minute upon a number of people are waiting. So, 0.6, but the very important summary statistic here is the average waiting time now because now it shows the average waiting time is 1.

So, if I press F9, for example, if I continuously press F9, you see that most of the time I am pressing F9 most of the time the average waiting time is more than one minute. So, I will go back to the presentation.



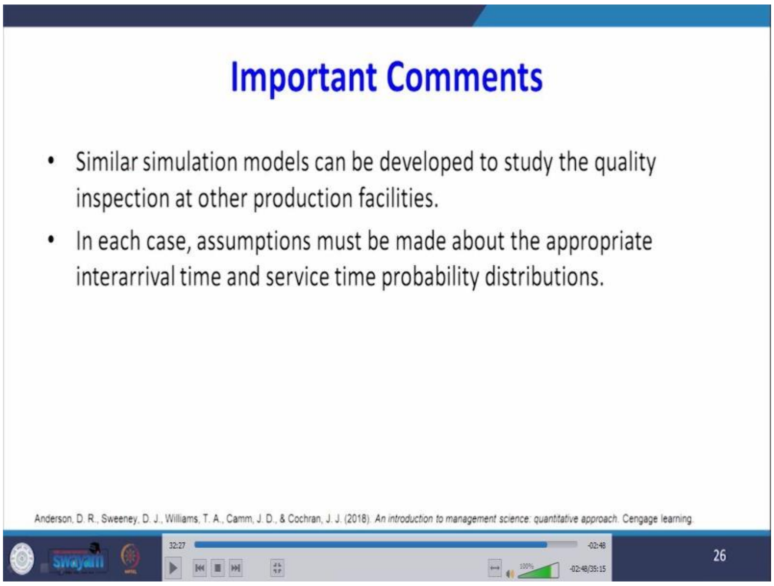
Summary Statistics	Excel Formula
Number of waiting customers	=COUNTIF(E16:E1015,">0")
Probability of Waiting	=E1018/A1015
Average Waiting Time	=SUM(E16:E1015)/A1015
Maximum Waiting Time	=MAX(E16:E1015)
Utilization of ATM	=(G1015-SUM(I16:I1015))/G1015
Number Waiting > 1 min	=COUNTIF(E16:E1015,">1")
Probability of Waiting > 1	=E1023/A1015

Now, I will conclude by saying that if you have only one ATM, you cannot achieve the service requirement of that customer and should not wait more than one minute. So, here I have shown the formulas excel formulas for finding the different summary statistics.



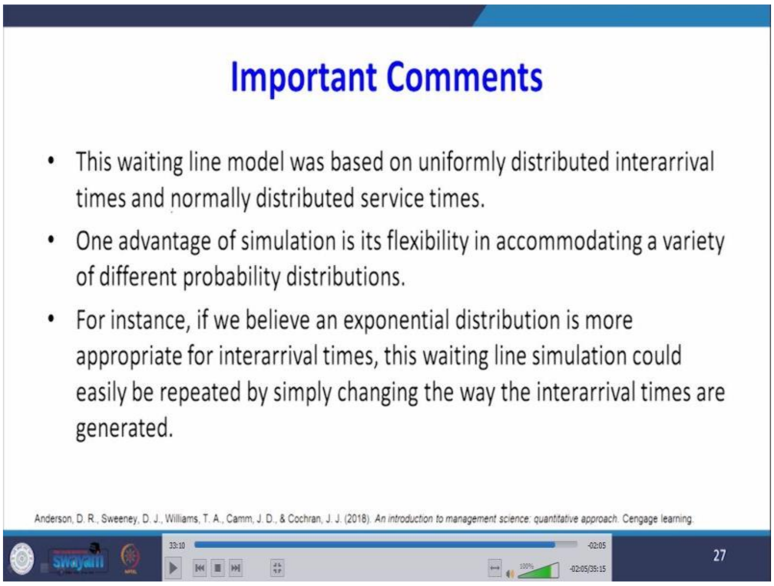
So, are we meeting the expectations of the management? What is the expectation of the management? The waiting time average waiting time of a customer should be less than 1 minute, but just now, we have seen most of the time, the average waiting time for 1000 customers is more than 1 minute. So, we are not able to meet the requirements of management.

So, if you are not able to meet this requirement, then we have to repeat this experiment by having 2 ATMs. Instead of 1 ATM, we have to see again what the average waiting time would be if we had 2 ATMs, which we will discuss in the next class.



Some important comments here are similar simulation models can be developed by studying the quality inspection at other production lines. So, we have discussed in the context of ATM. Even in the operation management context, the number of inspector quality inspectors is required. Whether we need to have only one inspector or two inspectors instead of an ATM should also be studied.

In each case, the assumption must be made about the appropriate inter-arrival time. What are the assumptions? What distribution follows inter-arrival time and the service time.



This waiting line model was based on uniformly distributed inter-arrival time and normally distributed service time. One advantage of simulation is its flexibility in accommodating a variety of different probability distributions. For instance, if you believe an exponential distribution is more appropriate for inter-arrival time, this weighting line simulation could be easily repeated by simply changing the way the inter-arrival times are generated.

Dear students, in this lecture, I have discussed the waiting line simulation. What example have I taken? A bank planning to have only one ATM and the service requirement is the average waiting time should be less than 1 minute. So, by having only one ATM, we repeated the simulation for 1000 customers. In the end, we found that we were not able to meet the service requirement, which is that the average waiting time is always more than one minute.

So, we have concluded that one ATM is not enough to meet the service requirement. In the next class, we are going to introduce 2 ATMs, and then we are going to see whether we are able to achieve this average waiting time that is less than one minute, which I will discuss in the next class. Thank you very much.