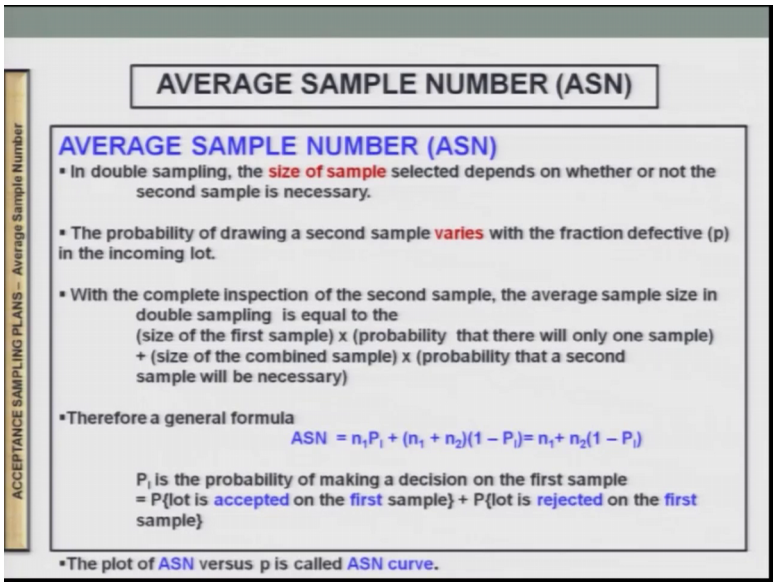
Total Quality Management - I
Prof. Raghunandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur

Lecture – 18
Double Sampling

Welcome back my friends very good morning, good afternoon, good evening to all of you. I am Raghunandan Sengupta from IME department; IIT, Kanpur and as you know that we are interacting for the course total quality management part 1; obviously, they would be part 2 later on and this is the eighteenth lecture we are having.

(Refer Slide Time: 00:38)



So, as we were disgusting the concept of acceptance sampling 2 stage sampling sequential sampling and depending on the values of d 1, d 2 or c or n whatever it is we will make a plan that whether we will accept that lot, reject that lot and before that we did discuss in a decent amount of details and for all information please read the Montgomery book if at all you can get a hand to that as of a excellent book in the area of statistical quality and to get an idea. It is a little bit on the upper side, but it is a very good book, it will take time go through the problems try to understand that and you will be able to implement many other statistical process, control quality control issues in practical sense.

So, when we are discussing the acceptance sampling in this 2 stage multi stage so; obviously, we discuss that the, the tong concept of total cost and that will come into the picture later on. So, let us continue our overall analysis and an interaction for acceptance sampling and the different ways how it can be done. So, let us start with the average sample size. So, basically I want to find out the average sample size which is needed in order to basically meet some criteria, already set led laid down criteria's which you want to study and; obviously, it would mean that intuitively more the number of sample size better the prediction, better the forecasting better the way you are trying to study, but on the other side the flip side is that they would be a cost involved.

So, you have to basically maintain a balance such that efficiency robustness of the method, trying to predict or forecast the actual parameter value ,what you want to find out from the sample given you want to study the population. So, those things should definitely, should be kept in mind as was along the cost structure.

So, average sample size which is ASN in double sampling the size of the sample selected depends on whether or not the second sampling necessary, which is true because if I pick up a small size n from of sample, from the lot and if c is fixed which is the number of defects which is permissible and an d 1 the number of actual defects is less than c; obviously, you accept that lot. If it is not then you go for the second stage and in the second stage the next lot which you find out which is say for example, d 2 minus d 1.

So; obviously, you are picked up d 1 in the first lot then you pick up d 2 minus d 1 in the second lot. So, if d 2 total, which is the total number of such defects in the first lot in the second lot is less than equal to some stipulated number you will either accept or reject or basically take decisions accordingly. So, 3 stage you would be done accordingly 4 stage would be done accordingly. So, they would be algorithm. So, each set, set of patterns of rules which you will follow so; obviously, you to mean that you will only sample in the second lot considering is a double sampling provided the first criteria is is not met if it is met; obviously, you will stop.

So, that is known as the stopping criteria. So, the probability of drawing a second sample varies with the fraction defective which is coming in the lot so; obviously, if the probability is very high then the probability of drawing a second sample would increase because you have to pick up in order to basically be aware whether that whole lot has

defects or not. If actual proportions in the real numbers P is much less then the probability of picking up the second order; obviously, diminishes and becomes very small, which means you would not require to pick up the second sample.

When complete inspection of a second sample the average sample size is basic in the double sampling is equal to these which is the size of the first sample which you picked up. Multiply the probability that they would do only 1 sample because in q tooks cases they can be 2 conditions. I take the first sample with some probability and also the case that I will have to go for the second sample with some probability if it is a 3 stage; obviously, you have to become first, second then once you reach the second it will be second third if it is the fourth then you will go basically go third and fourth and so on and so forth though; obviously, they are technically would be to 2 outcomes from any stage.

So, if it is you have to go for the second it means that you have reached the first the decision point based on which you have considered the first sample and then only make a decision go for the second. Once you reach the decision point for the second then only you decide on the outcome whether you should go for the third, when you reach the third the decision point then you make a deceive the analysis where they you should go for the fourth and so on and so forth.

So, for the double sampling the sample size is equal to the size of the first sample in the probability that they would be only one sample plus the size of the combined sample which means as we are trying to understand d 2 the total number of defects. Obviously, that sample size would now be initially n 1 later n 2 which is the combined sample size which you have for 2 samples multiplied by the probability that a second sample would basically be necessary if it is a third one; obviously, it will depend on the combined size and the corresponding probabilities would be brought into the picture accordingly.

So, therefore, the general formula is if you see the last value let me point it out, if you the last you see the last case. So, this term which you have which is I am highlighting is the sample size n 1. If you remember for the first sample multiplied the corresponding probability which is for the first stage hence the suffix 1. So, the 1 basically denotes the first stage set of observation, where the probability, whether the number of defects or whether c which is allowed c can also change for your kind information or whether the

probability. So, the suffix would basically denote at which stage and in the second case when we use the second properties the some of these.

So, this is technically the total value which you have total sample size multiplied by the corresponding probability that the second sample is needed. So, this is 1 minus P 1 because in case if it is P 1 is already taken care for the first stage then; obviously, 1 minus P is there for the second stage is basically like a binomial concept where they are 2 outcomes and then you proceed accordingly and the probabilities would be calculated accordingly. So, once you do that it is the probability comes on the overall average sample size comes out to be n 1 plus n 2 multiplied by one minus P 1.

So, where P one is the probability of making a decision on the first sample which is basically like this, probability of lot is accepted on the first sample plus lot is rejected in the first sample. So, there would be 2 costs cost would be coming that the first is accepted 1 come 1 come sample size from there, another is that it will be rejected the first lot would be rejected hence the sample size would be coming from there also. So, I am trying to find a corresponding average. Now consider very intuitively the averages would be basically found oh out like say for example, if I want to find out the average and I pick up a set of observations and consider the very simple case where you are rolling a die the face values are 1 to 6 and the corresponding probabilities considering is a unbiased die are all 1 by 6.

So, when I want to find out the average what I will do, I multiplied by the real as value which is 1, 2, 3, 4, 5, 6 whatever it is multiplied by the corresponding probability which is 4 which case it is all 1 by 6 because this is a uniform discrete distribution case.

Now, considered this case so I will come to that example which I am explaining on the slide later on, but consider a second case for the example of rolling the die, now then the faces are not 1 to 6, but they are now 1; 1 so; that means, number 1 appears twice other numbers are 2, 3, 4, 5. So, all the 6 faces are there available to you.

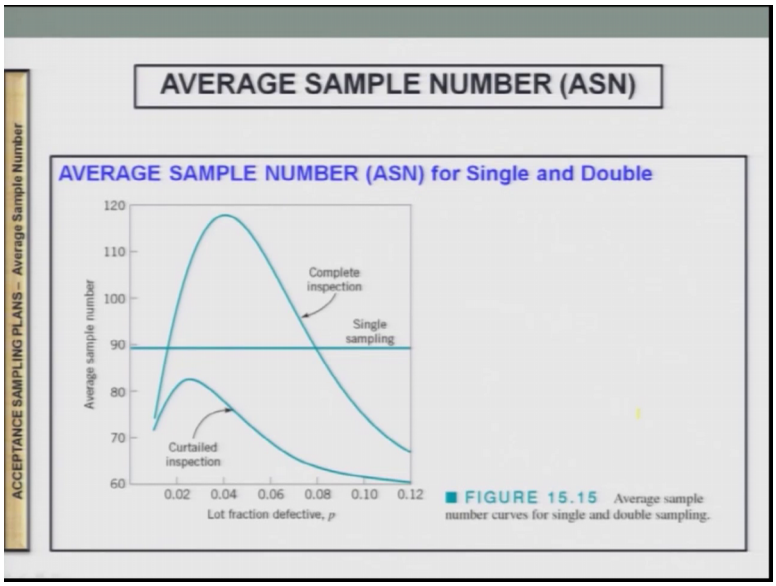
Now, it is also an unbiased die, now if I tell you to find out the averages what you will do? In the initial case you multiplied this real as value multiplied by its corresponding probability there. Now, you will basically multiply the real as value into its corresponding probability, only the difference is that in this case number 1 is appearing twice. So, the expected value or the average value would be 1 into 2 by 6, plus 2 into 1

by 6 plus 3 into 1 by 6 plus 4 into 1 by 6 plus 5 into 1 by 6. So, the 6 number is not there because 6 has been replaced by number 1 which is appearing to number of times.

So, if you follow this logic and if you see this formula is exactly the same you have the total number of observations which is in the sample size which is given by n suffix whatever it is 1, 2, 3, 4 depending on that multiplied by the corresponding probabilities. So, in this case you had taken the sample size n 1 with the corresponding probability of P 1 and in the second instance you will pick up all whole lot because you have to considered the whole lot as such it will be n plus 1 n 1 plus n 2 combined, multiplied by the corresponding probability that will go for the for the second lots observe set of observations where the probabilities are respectively P 1 for the first lot 1 minus P 1 for the second lot.

So, if you if you use this similarly with the with or the or the example which I gave for rolling the die and trying to find out the average and with this case it will make sense to you how we find out the average sample size oh sample number.

(Refer Slide Time: 11:05)



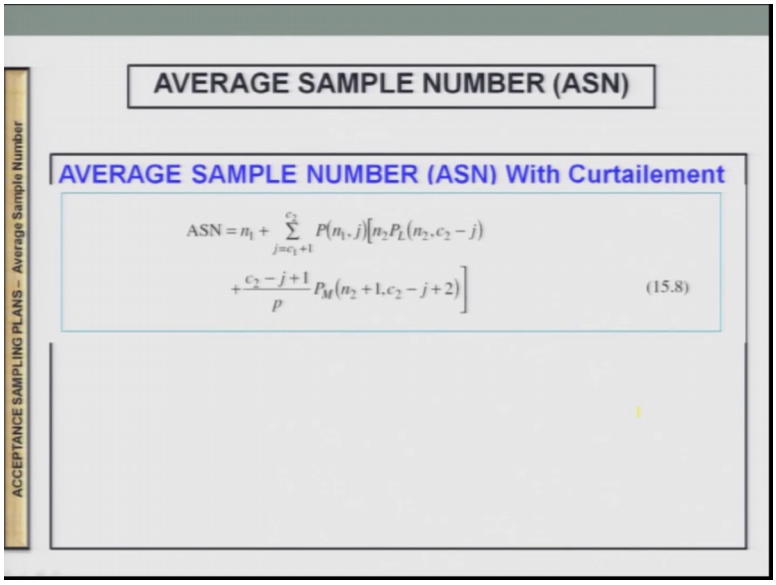
Now, considering the average sample number further so were again I am, I am mentioning that the diagrams are all from Montgomery its figure 15.15 which shows the average sample number curves for single and double sampling. So, single sampling means you pick up monaural take a decision accordingly and double sampling is there

you pick a 1 naught take a decision and if it is met then you stop, if it is not met then you go to the second stage and then basically make the final decision.

So, in this case you measure the average sample numbers along the y axis and the lot fraction defective which is the P value depending on the sample which you have collected on the x axis and you if you have a single sampling it is a straight line; obviously, because depending on your lot size perfect defective the average sample size would remain depending on the sample lot which you are picking and the probabilities you are picking the probability is not changing. So, if you pick up n 1 it remains n 1 and you make a decision, because they are the probabilities does not matter because only you will pick up 1 with probability 1. So, if you pick up n 2 it remains an n 2 if you pick up n thre3 is remains an n 3.

So, considering the horizontal line which you have drawn which is a little bit less than 90 if you see this and this figure. So, it come remains at the same level depending on single sampling which you have done for different set of observations. So, it remains at the same level, but in case say for example, you are doing complete in double sampling cases so; obviously, the probabilities and the sample size would differ which would have an effect how the curves looks like. So, the average sample size depending on complete inspection and curtail inspection would definitely be not a straight line, but curve depending on n 1 and the corresponding properties which is P 1 in the first case and; obviously, in the second case when you pick up the second lot into this 1 minus P 1.

(Refer Slide Time: 13:07)



Now, when I want to find out the average sample size with curtailment, curtailment hat means I have to basically cause structures are there, I need to basically minimize the cost then the formulas would be given that in the first case I pick up n 1 and the second case I would pick up a small certain sample which would basically depend on what are the set of observations which I am doing. Now, I will like to a little bit deviate from that and it basically tell you the concept of say for example, the double sampling or the triple sampling or the stage by sampling which you do or the sequential sampling plans which you do. Now, why this is important? Consider I am not actually deviating from TQM, but basically I am trying to give you an example, consider that you are trying to do a marketing survey and marketing survey needs a lot of costs.

So, if you are, are a company executive or company vice president and money is not a factor for you so; obviously, you will invest all your money whatever the quantum is to do a marketing survey and conduct your studies accordingly, but now considered money is important and that has to basically be used judiciously which means that did the number (Refer Time: 14:23) samples which you can take to predict or forecast or find out the actual population characteristics, would definitely be based on the fact that the overall information which will be getting from us other from a smaller set sort of samples considering that money is limited has to be done judiciously.

So, what technically would do is that very simply think about the plan you would basically pick up some set of observations study some study the sample characteristics. So, now, what you will do is that you will have some stopping criteria, stopping criteria

what I am using for the first time, but it very simply means that you have some rule based on which you will proceed. So, those rules are already found out from theoretical results as it is not a part and parcel of this course let us not discuss that one, that can be found or depending on the type of distribution which we have what type of example you are doing. So, you have some stopping criteria some stopping rules.

So, what you do is that you pick up a small sample n 1 consider and you test whether the stopping criteria is met, if it is met you stopping experiment report the sample size as the best sample size depending on the stopping criteria and use that to study and find out the sample characteristics which will mimic the population characteristic to the maximum possible extent. It is not possible to make them exact, but it will give you some, some error, but those errors are the 1 which you want to minimize or make it as small as possible based on which you are getting the stopping criteria or stopping rule. Now if the stopping criteria is not met by n ones of set of observations which you have you will take say for example, in the next stage some extra number of observation which is n 2.

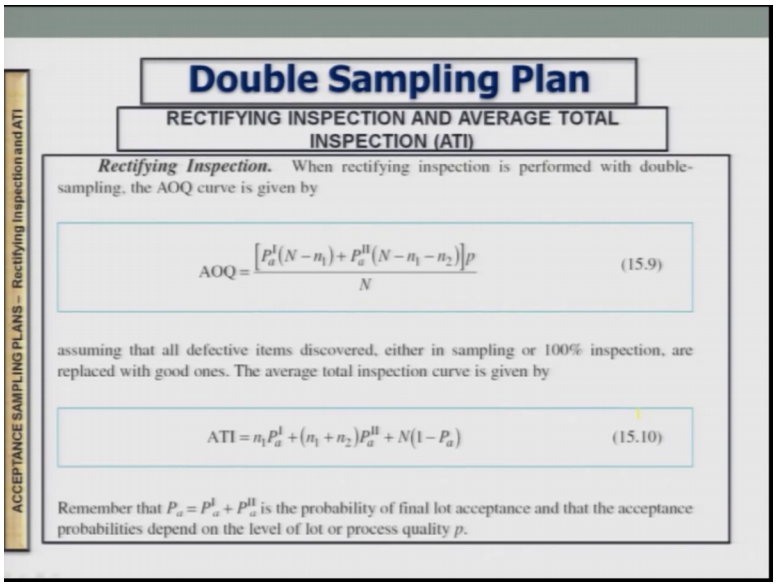
So, what you will now have ah whole set of observation which is now n 1 plus n n 2 based on n 1 and n 2 you find out the sample characteristics, check with the stopping criteria, if it is met report that sample size of n 1 plus n 2 as your actual study and do the set of calculations accordingly. If the sampling idea is not met again you takes take say for example, n 3 number of observations then again you check thus the sample characteristics based on the sample where the size is now n 1 plus n 2 plus n 3 you check if it is met stop if not you proceed accordingly. Which means and the other set of observations which you are take in are being basically tempered to the fact that how good or bad they are with respect to the stopping criteria, but also at the other end you are trying to balance your total cost accordingly.

So, if you follow this small concept which I just mentioned then you can basically try to be able to reduce your overall cost. So, the cost can have different implications it can be risk, it can be total cost for sampling different things could be there. Based on that if you find out the average sample size with curtailment you will get this formula in I am not going to the actual details of this, you will find out the actual formulas such that you can find out that based on the curtailment whether it is for 2 stage or 3 stage or 4 stage you can find out the average sample numbers accordingly, but remember always it is basically to study the average samples and numbers depending on what number of

observations you took in the first lot then in the second lot depending on the criteria's have been met or not and there your post proceeding.

So, if you find out in this so what you actually have is, basically the probabilities that given the sample size n 1 and n 2 you can basically proceed and find out the probabilities and multiply them with the corresponding sample observations and do it the calculations likewise.

(Refer Slide Time: 18:13)



Now, double sampling plan as I am I just mentioned the example was for the single sampling plan then we I just mentioned very briefly for the sequential sampling plan. So, let us go into the double sampling plan in details. So, rectifying inspection and average lot inspection numbers, so rectifying inspection means when rectifying an inspection is performed with double sampling the AOQ curve if you remember when we are doing the OC curves those concepts were consider. So, I am just giving the results you find out the AOQ values which would basically depend on which stage you are you are sampling.

So, the suffix which you see over P so, it can be either 1 or 2. So, depending on whether its double sampling whether in the first stage or the second stage if there are more than 3 2 stages 1, 2, 3 so; obviously, the suffix would change accordingly. So, here n 1 and n 2 are the total sample observations which you are taking in the first lot and the second lot accordingly and capital n is the total size of the sample space which you have from where you are picking up the lot. So, if it is n capital n minus small n it basically means

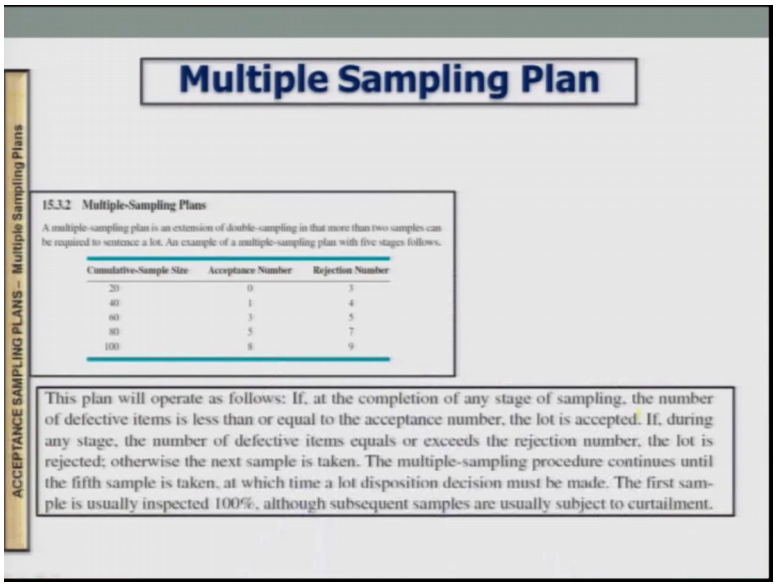
the overall numbers which are there which are still left. So, n 1 you have picked up capital n minus n 1 is the total set of observation which I left it in the so called sample space or the box technically capital n if it is infinite then it would not matter, but I am just giving an example because in many practical cases the capital N value would definitely be very blarge, but it definitely is a plausible number which based on which you can do your calculation.

Now, if you pick up a second lot; that means, you picked up the set of observations so; obviously, that total number of set of observations in the first lot and the second lot would be subtracted from the capital n which is the sample space and then you will multiply each values corresponding me with the values of the probabilities and basically find out the aoq value assuming that all defects have been discovered either in sampling or hundred percent inspection are replaced with the good one. So, once you find out the bad one you will replace. So, if you remember in the double sampling concept when we are discussing in the last class I did mention that once you find out the defective one you remove them and replace them in there with good one so; obviously, that information would come into the picture as you are calculating the overall the AOQ values accordingly.

So, the average total inspection curve would be given by the multiplication of n 1 into P suffix 1 and prefix 1 and suffix a which basically means the probabilities that the accepted numbers are some fixed value depending on whatever you think is right in the first case then P prefix to suffix a would basically be the corresponding probabilities in the second stage and so on and so forth. So, based on that you find out the overall the ATI which is the basically the total inspection of average total inspection numbers accordingly. So, remember Pa; P suffix; a is basically some of the probabilities in first stage and second stage which is the probability of the final lot acceptance and that the acceptance probabilities depends of the level of the lot or the process quality which is P.

So, the formulas which I giving are just the n results please if required refer the book because the overall concept which you are trying to follow is very vast. So, you will try to basically not skip, but try to not go into the details of the offs too much of details just give the final formulas based on which you can basically solve the problems accordingly.

(Refer Slide Time: 22:03)



Now, we come to the multistage sampling plan. So, a multistage sampling plan is an extension of the double which can go into the triple and so on and so forth in that more than 2 samples can be required to be picked up from a lot to pass on a on your judgment an example of a multi sampling plan with 5 stages are given as in front of you in tables 15.3.2; if you remember, I just mention about few minutes back about the sequential sampling plan that if you are taking set of observation trying to find out the characteristic to the sample check with the stopping criteria and then either proceed or not proceed depending on what the results are this is basically that only explained in much detail.

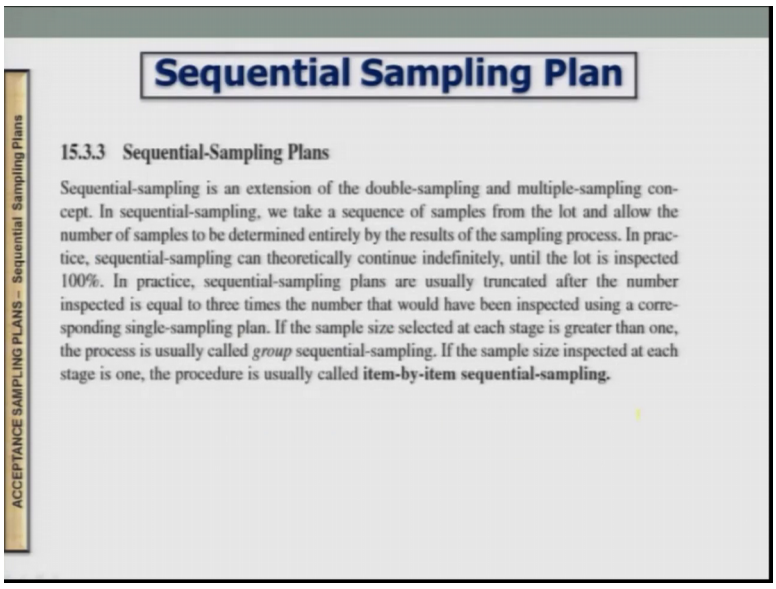
. So, you take of twenty of set of observation in the first you are accepted numbers are given and the rejections numbers are given to you pick up twenty check if it is not met you stop if it is not met you proceed then you again pick up 20. So, now, the overall cumulative number of observations norm becomes 50 that is what I mention when I pick up n 1 in the first lot check the result if it is not met, then I go into the second lot which means again I pick up n 2. Now I find out the sample characteristics based on the sample size which is now n 1 plus n 2 check with this stopping criteria met stop does is it is not met proceed. So, in the same way if you look at column one is exactly what it mentions I pick up 20, then 20 and so and so on and so forth and proceed accordingly.

So, the cumulative sample sizes are twenty forty sixty eighty hundred the accepted numbers are given depending on the problem and the rejections numbers are given this

plan will operate as follows let me read it you will understand if at the completion of any stage of the sampling the number of defect item is less than or equal to the acceptance number. So, the acceptance number and the rejections number are based on some set criteria which is already disc has already mean been decided which is what I am whose have been mentioning as the stopping criteria. So, let us continue reading it the defected items is less than equal to the acceptance number the lot is accepted if during any stage the number of defective items equals or exceeds the rejection number then the lot is rejected.

So, you reject on as except depending on that. So, so here if you see that I did mention that if the shopping credit is met you stop shopping credit not met you proceed is the same thing shopping credit means basically the difference or the rejections and actors acceptance number in this case; obviously, in other case it would be fine tuned based on what type of calculations which you have done for the stopping criteria. So, as I was reading the multiple sampling procedure continues until the fifth sample is taken at which time a lot disposition decision has to be made the first time the first sample is usually inspected hundred num percent number times although subsequent samples are usually subject to curtailment and you make a decisions accordingly sequential sampling plan.

(Refer Slide Time: 25:04)



Sequential sampling is an extension of the double sampling and the multiple sampling concept in more details in sequential sampling we take a sequence of observations a set of samples from a lot and allow the numbers to be determined entirely by the results of the sampling plan process.

So, if you remember I have been mentioning about stopping criteria basically rule is there, but the results of the rule or the outcome of the rule in them some mathematical sense will depend on how many number of observations I have taken previously. So, this is what it is mentioned the number of samples to be determined and it is determined by the results of the sampling process in practice the sequential sampling can theoretically continue doing infinite time. So, we you may not stop until the lot is inspected hundred percent number of times in practice sequential sampling plans are usually truncated after the number of inspection is equal to 3 or times the number of the would be to be have been inspected using a correction corresponding single sampling plan technically they do not; can go to infinite.

So, we check stop if it is made not stop you continue. So, the what is mentioned here is on the practical framework, but in the theoretical frame work it can go till infinity if the sample size selected at each stages greater than one the process is usually called a group sequential sampling if that sample size inspected each one is procedure is one then it is called a sequential sampling plan. So, let me explain it in more details I take a one observation set of observation which is n 1. Now when I mentioned the last sentence mentions is sequential sampling which technically is purely sequential sampling is what I take one extra observation.

So, now my sample size is now n 1 plus one I check the stopping criteria based on the sample size of n 1 plus one if it is met stop and you put that n 1 plus 1 as the sample size and finish my task if it is not met then I again take another extra observations now my total of number of observation is n 1 plus 1 plus 1; that means, plus one the first plus one was from the last previous stage now I take another observation. So, now, it is n 1 plus 2 check the criteria met stop not met go to the third stage where I take n 1 plus 1 plus 1 has already been taken I take an extra plus one. So, it is basically n 1 plus 3.

Now in the other cases which is known mentioned as groups sequential sampling plan is basically known as batch batches. So, it means I can take n 1 then take not one, but say

for example, I take n 2 where n 2 technical can be less than n 1. So, I take batches then in the third stage I take n three which is also less than n 2 now and I proceed accordingly and finish my task. So, in general it is seen that in seek in technically in sequential sampling plan if there are restrictions on some stopping criteria or some totaled cost and all these factors are there in general the efficiency and the robustness of the sampling procedures are better than the fix sampling plan where the fix sampling plan in the sample is already predetermined before hand in general and also the overall efficiency of the models for sequential sampling plan are much better.

So, with this, I will end there are 18th lecture and continue our discussion of sampling and total quality management further on starting from the nineteenth and. So, on and. So, forth have a nice day.

Thank you very much.