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

Lecture – 05
Part-1
Interval Estimation- I

A welcome back my dear friends a very good morning, good afternoon, good evening to all of you, this is the TQM 2 lecture number 5 and I am Raghunandan Sengupta from the IME department IIT Kanpur and we have just given the 3 distributions; I know it is a little bit fast, but this would be repeated time and again I am sure you will get adequate time to understand them in TQM 2, number 2 reason why I am going a little bit fast is that this is not the general things which are needed in TQM 2, third reason is that people who are nowadays statistics is an integral part in class 12 and some in schools and some colleges they do give some constants of probability in class 10 also and obviously, in engineering and non engineering courses probability and statistic is a big part.

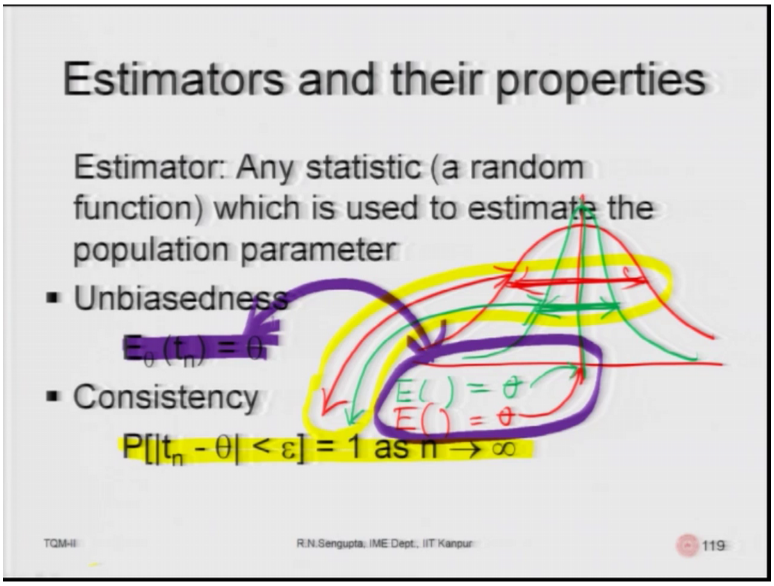
So, that is why these are precursors I am sure you know that no proofs are required, only the results would be required in order to basically proceed for the design of experiments. So, with this let me continue. So, what are the estimators and what are their properties. So, both the by the word estimators I mean that considering here population there are some parameters, parameters are technically scale parameters, location parameter, shape parameters which gives you the first moment, second moment, third moment, fourth moment; which basically means something to do with the mean, median, then the variance dispersion and all these things, so these are the characteristics of the population.

Now, considering the characteristic of the populations are very important for us to find out if we do not have their actual values, what we do is that we pick up a sample find out from the sample. The sample characteristics and try to find out that whether using though sample characteristics we can find out something to do with the population parameters. So, there we need estimators and the concept of estimation process would be required, so this is what we are discussing.

So, estimator is any statistic a random function which is used to estimate the population parameter. So, they can be either unbiased and they or they can be biased and if they are unbiased which you technique it means the average value of that because, that the estimator if you are tipping picking out sample time after time though the oceans are different if they are different then estimates would be different; say for example, the sample mean would change and if the sample to mean keeps changing; obviously, it will have a distribution.

So, hence it will have an expected value and in the long run if the expected value of that of the sample statistic is exactly equal to the population parameter then obviously, it means unbiased. But obviously, it does not mean that the variance is high and low, so that would basically be coming out from the concept of consistency. So, consistency means that how consistent the results are, that means in the long run that as n the sample size increases the difference between the sample estimate in the population parameter is less than some epsilon.

(Refer Slide Time: 03:02)



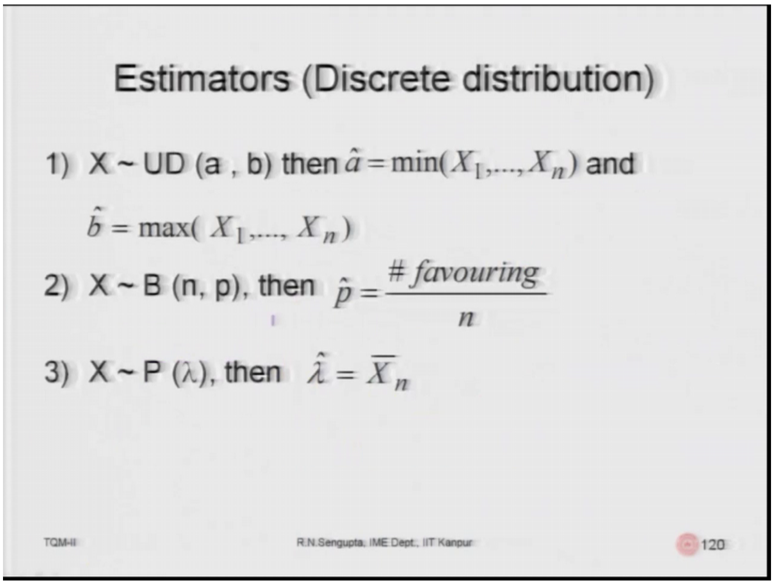
Again this epsilon is some function of n and that basically that n tends to 1 as n increases to infinity.

So, let me draw two diagrams one diagram would be sufficient for all of us to understand. So, this is the distribution consider this is the distribution and then another and this is the mean value let me change the colour, so this is one distribution another one; obviously, the mean when is same. So, in this case expected value of this green line, so the green line is whatever the parameter is basically equal to theta in the second case, the expected value of the green red line is equal to theta. Theta is basically this line which we want here to give me 1 minute, let me change the colour would be easier for me this is the mean.

Now, very interesting thing you understand is the variance which is there, the dispersion in this case and the dispersion which is happening in this case are two different values, which means as I increase the sample size the difference between the expected value and the variance basically closes; that means, the mod, so each and the sample size would give me different statistic. So, as I increase the sample size the difference between that actual sample statistic and the population value will start decreasing. So, this would be some concept which we are studying.

So, this is something due to the variance, if I change the colour this is something to do with the variance and so the concept of the variance would be coming out here, so let me use a highlighter. So, the highlighted this is to do with this something to do with this and if I use the highlighter let me change the colour while it should it look good let me try. So, this is the 1 so this would something to do with this so they are related. So, I am just trying to make you understand that what is the concept of the unbiasness; and consistency.

(Refer Slide Time: 06:02)



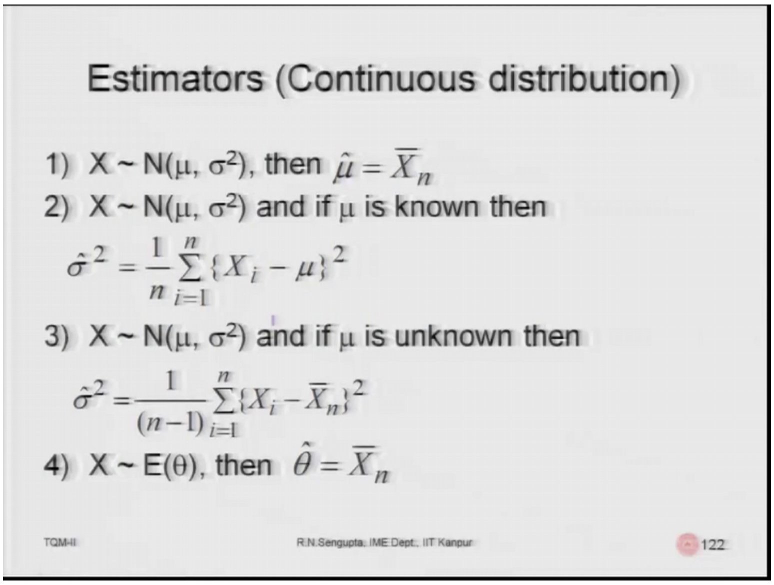
So, for the discrete distribution if it is uniform between a and b and then a is basically and the best estimate would be the minimum y X 1 to X n and for b estimates. So, hat means the estimates which you have not the actual values. So, a is best estimate is a hat which is the minimum value between X 1 to X n, but it does not mean say anything about the consistency and un biasness and b is base estimate is b hat which is the maximum X 1 and X n. So, intuitively is very easy to understand say for example, I have a box and there are chits marked say for example, million of them 1 to 1 million and I do not know what is the minimum value 1 and what is the maximum value.

So, consider theoretically the minimum value is 1; maximum is 1 million I need to find it out, what I do? I pick up a sample of 100 randomly I pick up and note them down find out the minimum note it down separately my find out the maximum note it known in separately. So, I continue picking up 100 observations time after time with replacement. So, if I have different minimums for each set of samples which you pick up 100 in number if I have different maximums for different samples of you pick it up, if I find out the minimum of that minimum and maximum of the maximum that in the long run move base it will give me some information over the minimum, the total population which is 1 and maximum the total population is as 1 1 million.

So, that is basically some essence or what I mean that actually that I have tried to find out the minimum of the maximum of the observations. Now, if by if X is binomial with parameters n and p and then p hat or which is the best estimator of p is the number of favouring or number of unfavouring.

So, if it is q, I need to find out it will be the number of unfair favourable divided by n and if it is p which is p hats is the best estimate for that then it would be a number of favouring 1 divided by n and if it is basically the Poisson distribution which is the lambda then the best estimate for that lambda would basically be lambda hat which is again equal to the sample mean, so this which we already have.

(Refer Slide Time: 08:08)

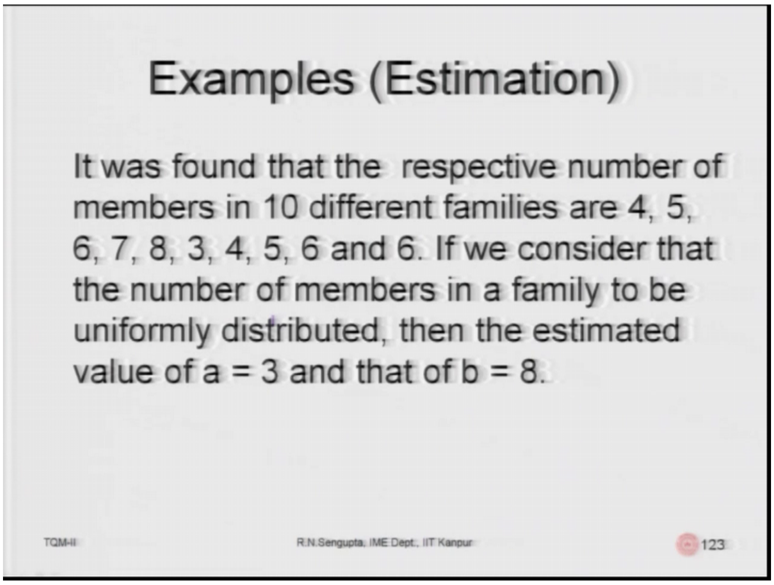


Now, if we want to consider the come to the case of estimators for the continuous distribution consider X is normally distributed with mu and sigma square. So, best estimate for the mean is X bar n which is the sample mean, now in the case when you want to find out the variance best estimate; the variance there are two cases if you remember, s dash and s dash s without the dash s dash is the case when you have the population mean is known.

Hence, you do not die divide by divide by n minus 1 only divide by n and then you find out the best estimate for the standard deviation on the best estimation of the population variance. In case if mean is not known you do place that with the sample mean you will lose one degrees of freedom, hence it is divided by the total values divided by 1 by n minus 1 and that is the best estimate for the population variance.

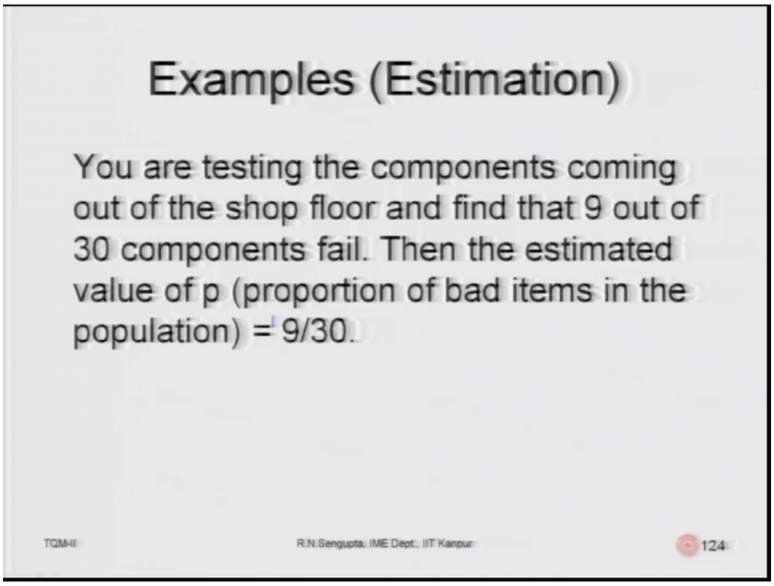
For the case when you have the exponential distribution without the value of a, so there is only one parameter theta. So, the best by estimate for the theta is basically the sample mean.

(Refer Slide Time: 09:07)



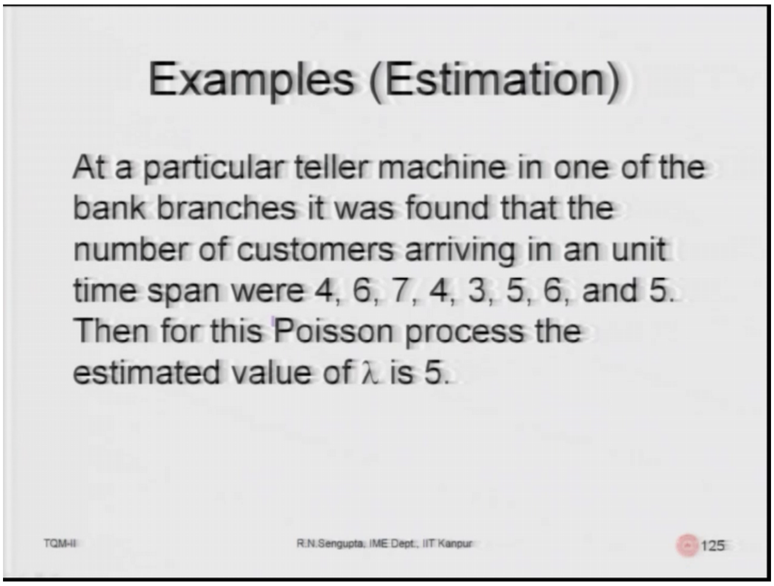
So, example it is found out that respective numbers of members in 10 different families are 4 5 6 and till 6, if we consider that the number of members in a family to be uniformly distributed, then the estimated variety of a and b would be, a and b as we just discuss.

(Refer Slide Time: 09:26)



How they can be found out your testing the components coming out of the shop floor and find out 9 out of 30 components fail, then the estimated value of p which is the proportions of bad items or it can be the proportions q can be the proportions of bad items, it is just a normal picture what we are using that would be the n by 30 which is the proportions or the chance or the relative frequency and in the long run it is the probability.

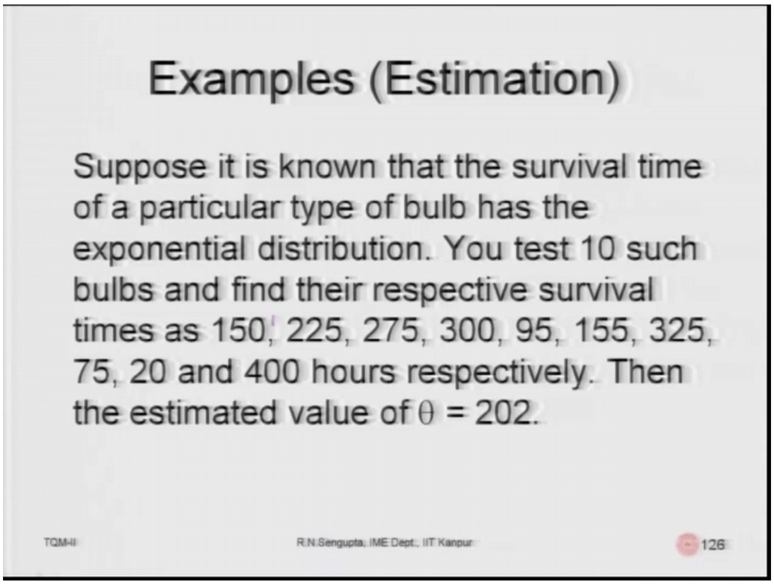
(Refer Slide Time: 09:47)



At a particular teller machine in one of the bank branches it is found at the number of customers arriving in a unit time span, whatever the unit time span may be may be 1 minute 1 hour 3 minutes 3 5 minutes whatever it is, the numbers are 4 6 7 till the number 5 then they are Poisson distributed if you find out the distribution of the numbers; then the estimated value of lambda which is the parameter for the Poisson distribution would be the sample mean.

So, which means you add up all this values divided by the that the total number of readings you get the value of 5.

(Refer Slide Time: 10:20)

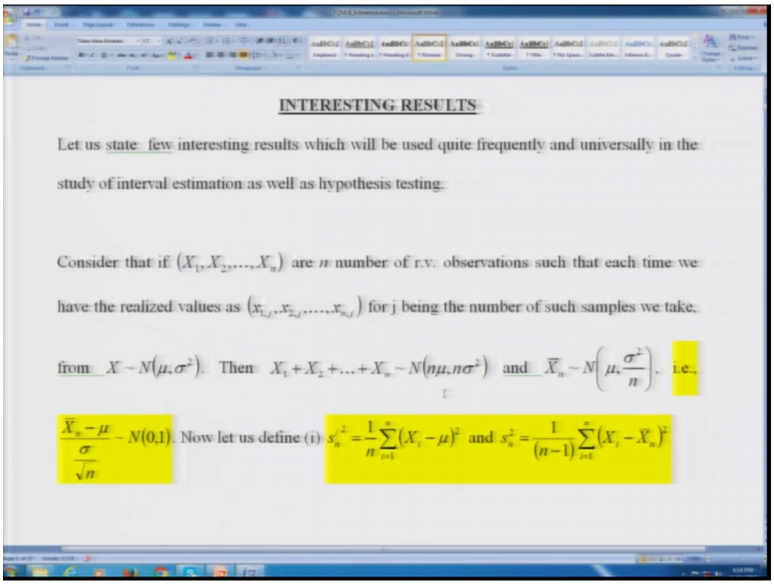


Then considering the exponential distribution, suppose it is known that the survival time of a particular type of bulb has the exponential distribution you test 10 such bulbs and find out the survival time to be 150 to 125 till the value of 400. So, there are 10 such examples then I want to find out theta which is the parameter for the exponential distribution, best estimate is the sample mean what you do is that add up all the values starting from 150 to 400 divided by 10 and you get the value of 202.

So, with this I would not end, so what I will do is that I will start off the concept of estimation more in general and then go into the point estimation concept than the hypothesis testing, this would basically end the preliminary portion and then we can go into the actual design of experiments.

So, there are some interesting results which mean we may have we may gone through some of the results very preliminary, but please bear with me as I progressed you would find out the overall expanse of what we want to discuss is much more and this would become really make some sense in our discussion as we go into the later chapters for design of experiment which is a part and parcel on a big area of TQM 2.

(Refer Slide Time: 11:35)



So, consider if we X 1 to X m x n and there are n number of random variables, such that each time we have then realized values are given by x 1 j x 2 j x 3 j till x n j x j j j is suffix, basically means the time when we are picking up. So, say for example you are picking up and the random variables are technically before you pick up their X 1 to X n when you pick up they becomes x 1 1 suffix 1 the first one is with the reading number for second one is basically for the person, second reading for you would be x 21 2 for the number of the reading what for the second one is basically for the person. So, if you pick up.

So, if they are x 1 1, x 2 1, x 3 1 till x n 1, if I pick up I am the second person my readings are basically x again it will basically x 1 2, x 2 2, x 3 2 and so on and hence forth. So, there would be jth j number of persons who can pick up and hence the suffix are given by x 1 j x 2 j and so on and so for j being the number of such samples we take.

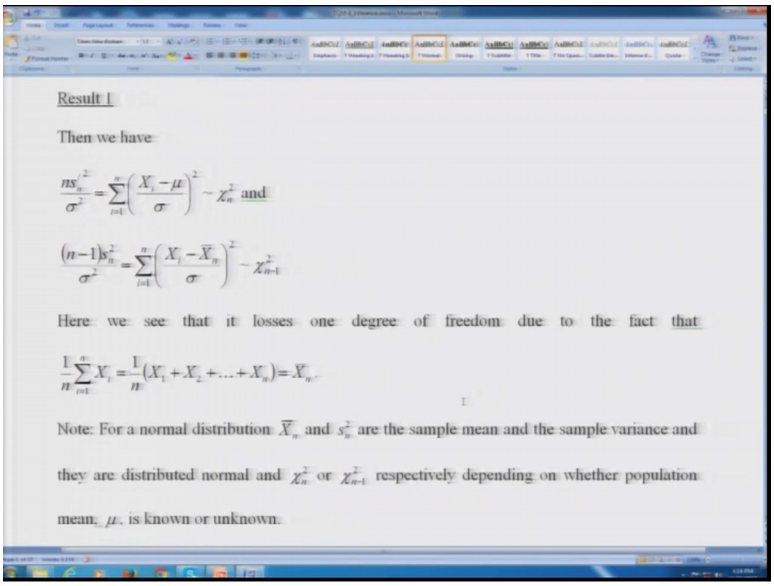
Now, the x is distributed and normal with mu and sigma square then; obviously, the sum which is s if you have discussed that is why I say that I will be repeating few things, but I will going not in depth; but trying to repeat it in order to make it make much better sense to you all. So, the sum of the n number of random variables which are all under normal you that mean value would be n mu and the variance would be n n sigma square and in the case if you have need to find out the sample mean, then the sample mean we distributed again with normal with mu as the mean value in sigma square by n and the variance and if we convert them into the corresponding standard normal, for x would be let me repeat it 1 by 1.

So, x would be distributed as z where z will be given by x minus mu divided by sigma, in the case when you are trying to find out the standard variable the standard variate with corresponding to x bar n it will be x bar n minus mu divided by sigma by square root of n, as it is given where I am hovering my finger, so I will try to basically highlight it. So, let me give the colour that would to be difficult for oh this is not PPT.

So, I have to basically use some colour if possible so this would be the case when I want to highlight and if you find out. So, this is I am using the yellow colour and so this would be normally distributed standard normal with mean for x bar n as mu and stance was sigma square by n, similarly you can do all the calculations corresponding to finding on the standard normal now I did also mention about s dash and s without the dash.

So, s dash and s these without dash are basically, the case where we take further the variance of the sample. So, these are the values s dash is where you do not loose n degrees of freedom because, the population mean is known s without the dash is basically the for position means unknown you replace that with the sample mean, so hence it is divided by n minus 1.

(Refer Slide Time: 14:44)



Result 1

Then we have

$$\frac{ns_n'^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi_n^2 \text{ and}$$

$$\frac{(n-1)s_n^2}{\sigma^2} = \sum_{i=1}^{n}\left(\frac{X_i-\overline{X}_n}{\sigma}\right)^2 \sim \chi_{n-1}^2$$

Here we see that it losses one degree of freedom due to the fact that

$$\frac{1}{n}\sum_{i=1}^{n}X_i = \frac{1}{n}(X_1+X_2+\ldots+X_n) = \overline{X}_n.$$

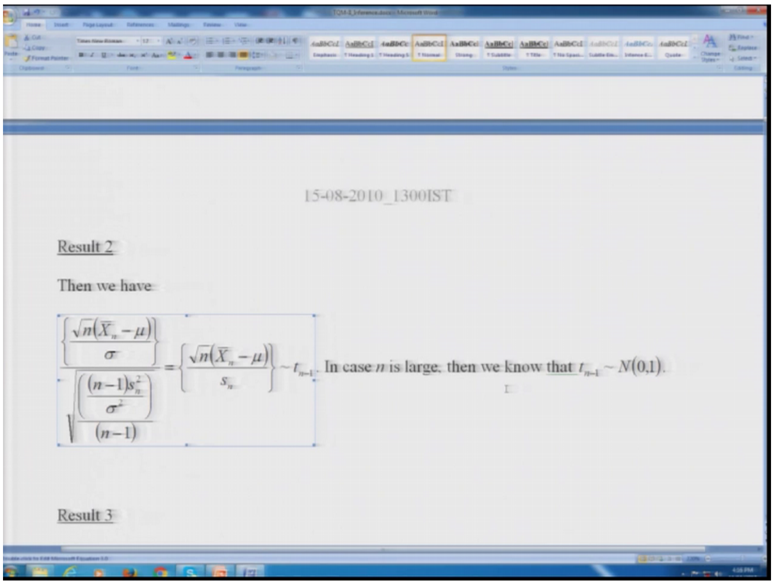
Note: For a normal distribution $\overline{X}_n$ and s_n^2 are the sample mean and the sample variance and they are distributed normal and χ_n^2 or χ_{n-1}^2 respectively depending on whether population mean, μ, is known or unknown.

So, result one would be which I have already repeated, I will repeat it again the distribution of the ratio of s dash square by sigma squared, on obviously, n and n minus 1 would come or corresponding to how you do that it will be given by chi square with n degrees of freedom and the ratio of s without the dash sigma squared multiplied by n minus 1 would be given as the distribution chi square with n minus 1 degrees of freedom.

So, here we see that we lose one degrees of freedom due to the fact, that the mean value which is to be found or as the best estimate for the population mean is given by the sum of all the random variables divided by n, random variables which means which you are picking up for the sample divided by n which is the sample size.

Note for a normal distribution x bar and s square are the sample mean and the sample variances and they are distributed normally and chi square and chi square n minus 1 are these are the corresponding distributions of which we find out, corresponding to the case when the population mean mu is known and unknown. If it is known we do not lose any degrees of freedom if it is unknown we lose 1 degrees of freedom.

(Refer Slide Time: 16:00)

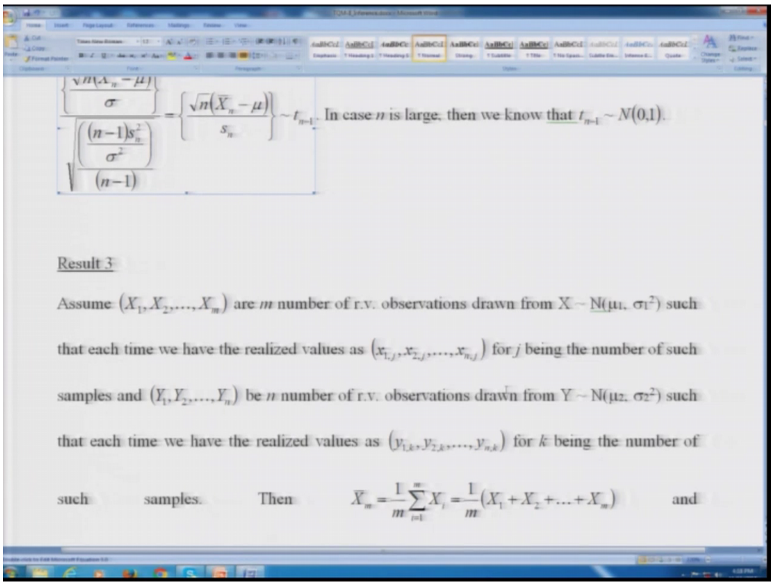


The second result which is again very intuitive because we remember that t distribution is given by a ratio of z, which is with z distribution divided by square root of chi square by n so this is the same thing. So, if you if you see this the where I am hovering my pen or whether I should basically hovering the cursor this numerator is basically standard normal distribution and the denominator you have 2 parts, 1 is the chi square this 1 and this is divided by degrees of freedom. So, the ratio is technically a z distribution divided by square root of chi square divided by it is 10 degrees of freedom and once you basically make sense of that it becomes t distribution with n minus 1 degrees of freedom.

In n is in cases n is large then it can be approximated by a standard normal distribution because, you remember that the t distribution has a mean value of zero and as n basically increases it can be approximated by a normal case result 3. So, these are important result 3 is that consider you have two different set of observations considered to give a very simple example consider there are 2 factories manufacturing some the same items, 1 is in mud Chennai another is Bangalore or 1 is in Delhi 1 is say for example, Patna whatever it is and you want to basically find out that what is the variability of the production.

Variability of production can be related to this to the shim the ductility of the material, may be relate to the tensile strength, may be related to the length, may be related to the inner dimension, may be related to the viscosity whatever it is.

(Refer Slide Time: 17:36)

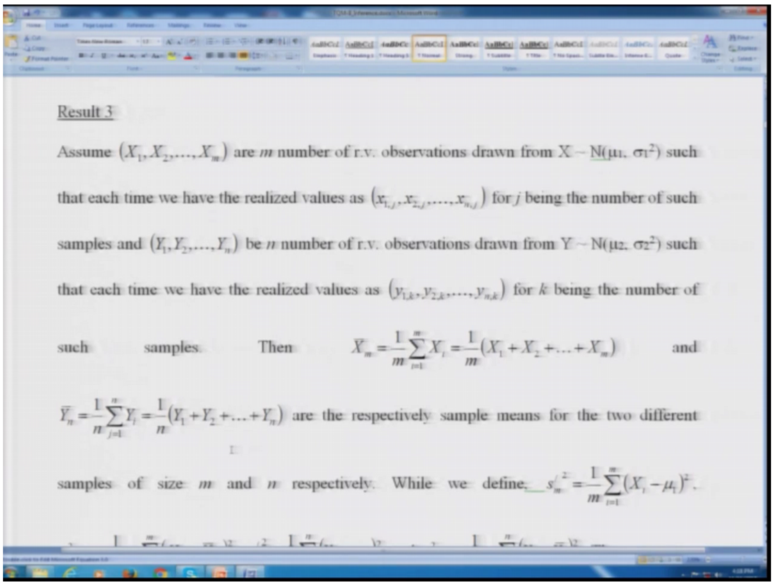


So, assume x 1 to x m, m as in mango the number of random variables you are picking up from the first set and it is basic distribution is normal with mu 1 and sigma 1 square r as it is basically mean and variance and they are the value realest values depending on if they are the j th number being picked up they would be x small x because, they are the realest value small x 1 j small x x 2 j, x 3 j so on and so forth. So, if you are picking of you are the first person you are picking up it will be x 11 for me it will basically be x 1 2 and so on and so forth.

So, this very simple nomenclature the first 1 is the reading number second 1 is the person who is speaking up and or which time you are picking up. So, y 1 2 y n n as in Nagpur is the number of random variables draw being drawn from y, which is a normally distributed with sigma 2 as square is the variance and mu 2 s is normal value or population value expected value, such that the observations are again y 1 k case the number of reading we are picking up as it is j in x.

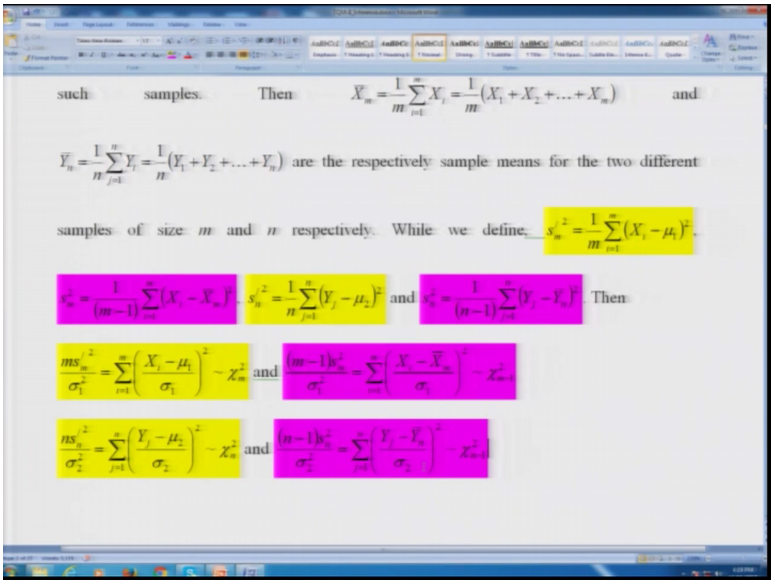
So, here mean values are for the sample is x bar m and for y it is y bar m n and the standard error or the variance of the sample are respect to 2 cases.

(Refer Slide Time: 19:02)



Case 1 population mean is known for x and population mean is not known for x, another case is population mean for y is known population mean for y is not known. So obviously, they would be s dashes in both the cases s dashes without in both the cases, but only to differentiate that s dashes you will put the suffix corresponding to the sample size which is calculating which would become very evident from you here.

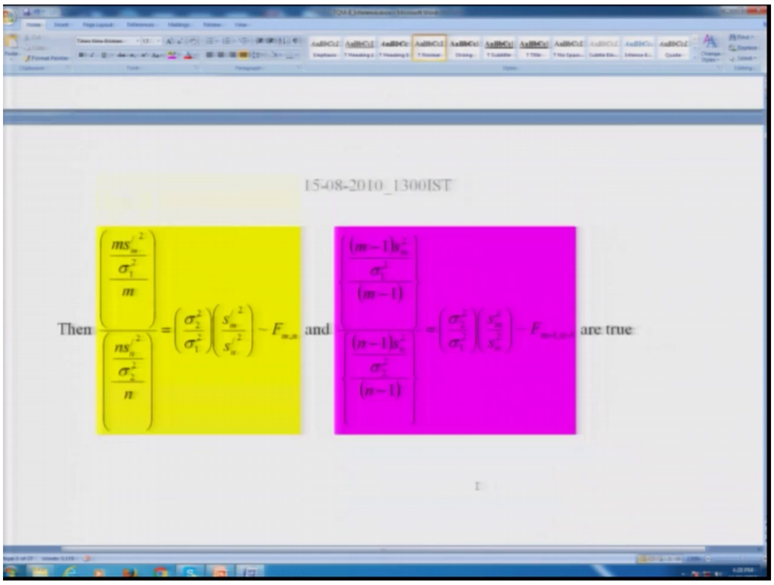
(Refer Slide Time: 19:28)



So, if you consider this that mean colour it for our simplicity. So, this is for m and this is for n, so now see here they are both the dashes. So, they do not lose the degrees of freedom for the other case it is n m, but without dash again without dash n, so if you compare them you will understand.

So, in the first case when we find out the you have a chi square with m degrees of freedom. So, this will basically denote by this yellow colour, so this is become apparent to you. So, I am just highlighting and this would become the chi square with 1 degrees of freedom being lost. So, these ratios would give you the y because, you are not trying to find other population meaning using it is sample mean.

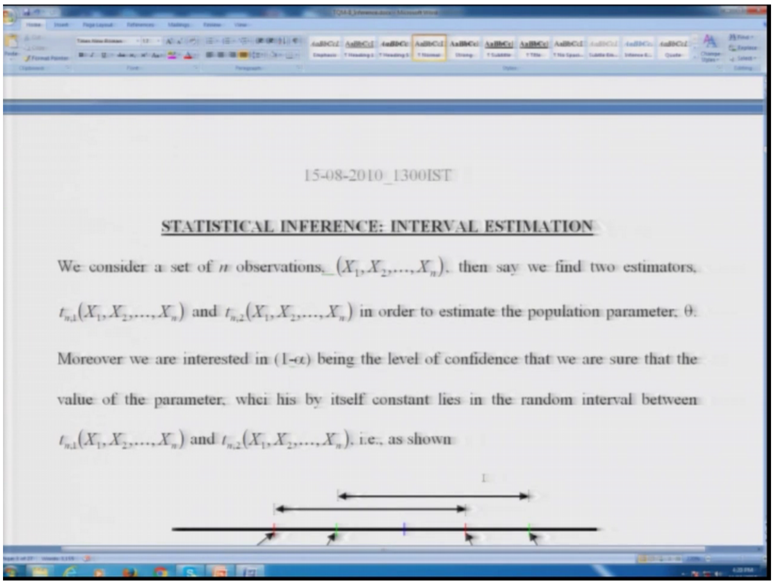
(Refer Slide Time: 20:33)



Now, if I come to find out this results have been repeated, but I am please give forgive me I am trying to repeat it time and again because they would be very relevant later on. So, if I want to find out the ratios of the standard error whole square or the standard way with the sample variance for the sample and considering it is s dash.

So, if I find out if it becomes Fm comma n no loss of degrees of freedom for the f distribution and in this case it will be Fm minus 1 n minus 1. So, their relevance is where I am hovering my cursor, so this is dash this is dash I do not lose any degrees of freedom, this is without the dash which this is without the dash because I have lost degrees of freedom. So, that is why they are m m minus 1 n minus 1.

(Refer Slide Time: 21:29)



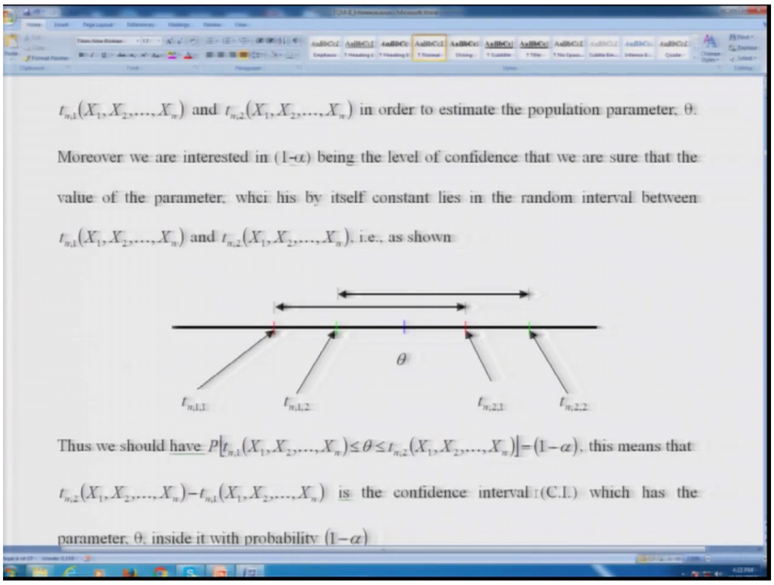
Now continuing our discussion may be a little bit more philosophical discussion we consider a set of n observations X 1 to X n, they are set and they say that we find 2 estimates casa de we find 2 estimates given the population. So, we pick up on a set of observations x 1 to x m n and consider from them we can find out 2 functions or consider we for 2 values. So, name them as the statistic t, t is not the t distribution remember that t suffix n 1 and t suffix n 2.

So, this 1 2 are not the suffixes of n by itself they are n comma 1 n comma 2. So, n remains the sample size 1 and 2 which is the second number after the end basically denotes the 2 values. So, they are basically used to estimate the population parameter theta moreover we are interested to find out given a level of confidence. So, this level of confidence would basically mean that to what level of significance are you certain the answer is correct. So, say for example, if I am saying that I am 95 percent confidence it means that if I pick up 100 observations 95 a number of them would be true.

So, obvious my level of confidence would differ depending on whatever absorptions you are picking up and depending on the experiment we are doing. So, this 1 minus half alpha would be the level of confidence that we assure, that the value of the parameter which by itself is a constant lies in the random interval range between these two numbers, which you have found out which is t suffix n comma 1 t suffix n comma 2.

Now, remember t n comma 1 and t n comma 2 this n 1 and n 2 are suffixes the overall random weight the points which are picked up from the observations which are still any number, but we are combining them in such ratios that the values are different. So, this diagram will make sense in some sense.

(Refer Slide Time: 23:17)



$t_{n,1}(X_1, X_2, \ldots, X_n)$ and $t_{n,2}(X_1, X_2, \ldots, X_n)$ in order to estimate the population parameter, θ.

Moreover we are interested in (1-α) being the level of confidence that we are sure that the value of the parameter, whci his by itself constant lies in the random interval between $t_{n,1}(X_1, X_2, \ldots, X_n)$ and $t_{n,2}(X_1, X_2, \ldots, X_n)$, i.e., as shown

Thus we should have $P[t_{n,1}(X_1, X_2, \ldots, X_n) \le \theta \le t_{n,2}(X_1, X_2, \ldots, X_n)] = (1-\alpha)$, this means that $t_{n,2}(X_1, X_2, \ldots, X_n) - t_{n,1}(X_1, X_2, \ldots, X_n)$ is the confidence interval (C.I.) which has the parameter, θ, inside it with probability $(1-\alpha)$

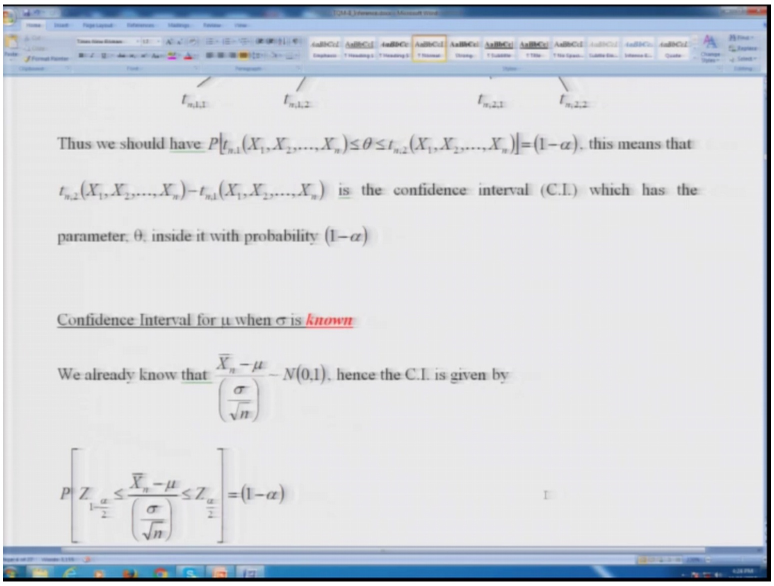
So, if you consider the diagram here the theta which is the blue line is fixed and when I pick up observation depending on where this the first one, second one, third one. So, for example, you pick up observations and they are different from my observations, you find out 1 minimum value 1 maximum value and these minimum and maximum values are given by tn1 comma 1 tn2 comma 1.

So, this basically the second 1 basically means the person and the first 1 and 2 basically means the lower value and the higher value. If I pick up the observations in my case they would be again the sample size is same which is n, they would be as basically be n t suffix n comma 1 comma 2 and the higher values would be t suffix n 2 comma 2, when the first 1 and 2 are basically the minimum and the maximum value and the second 2 is basically number is picking up.

So, from there we find out that given the level of confidence we will find out that, the probability that the lower values and the upper values would contain the value of theta with a with the level of confidence which we just mentioned as 1 minus alpha; which means that if I am say; that I am 95 percent confidence level it would mean that the minimum and the maximum value which I find out from the sample set of observations and give the minimum and a maximum value whatever it is the statistic, they would contain in that interval the value of theta which is the population parameter.

So, this means that tn1 and tn 2 the difference between them that is the maximum of the higher and minus the more over value would give me the confidence interval which has the parameter theta inside it with the probability of 1 minus alpha.

(Refer Slide Time: 24:59)

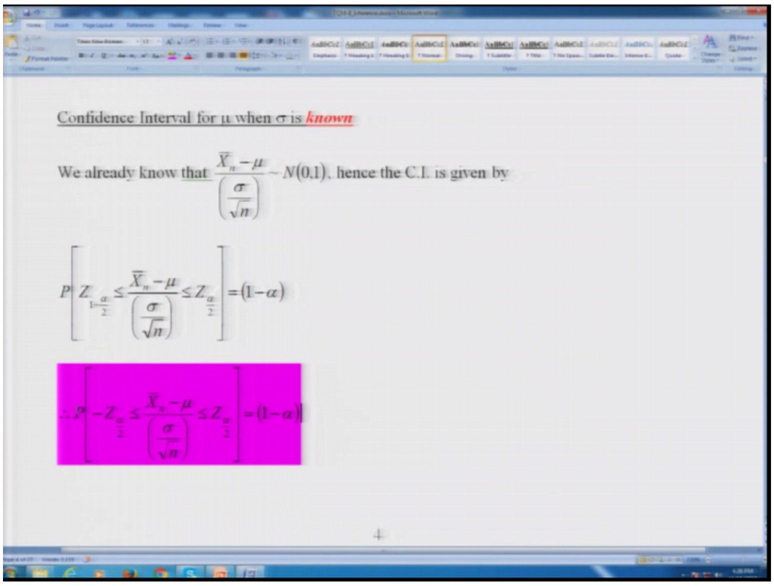


So, consider the probability the interval of mu when sigma is known, so consider when sigma is known the z distribution is obviously, basically x bar suffix n minus mu divided by sigma by square root of n would be z distribution. So, once it is at distribution and considering sigma is known, then if you find out the minimum and the maximum value they would basically be given from the z table depending on the level of confidence which we have. So, the minimum value would be said suffix 1 minus alpha by 2.

So, now, listen to me carefully, so the suffixes are basically denotes to where the random variable which you are measuring r. So, if I consider the x x line as the values where I am converting the z. So, z would basically be coming it would be a negative from my side there is a mean value 0, under the left is 0 under the right s is positive. So, the values which I take on the left would be denoted by z suffix 1 minus alpha by 2. So, 1 minus alpha by 2 is the area still to be covered and once I basically go on to the right hand side, it will be z alpha by 2; that means, the area to be still covered is alpha and the already covered is 1 minus alpha by 2.

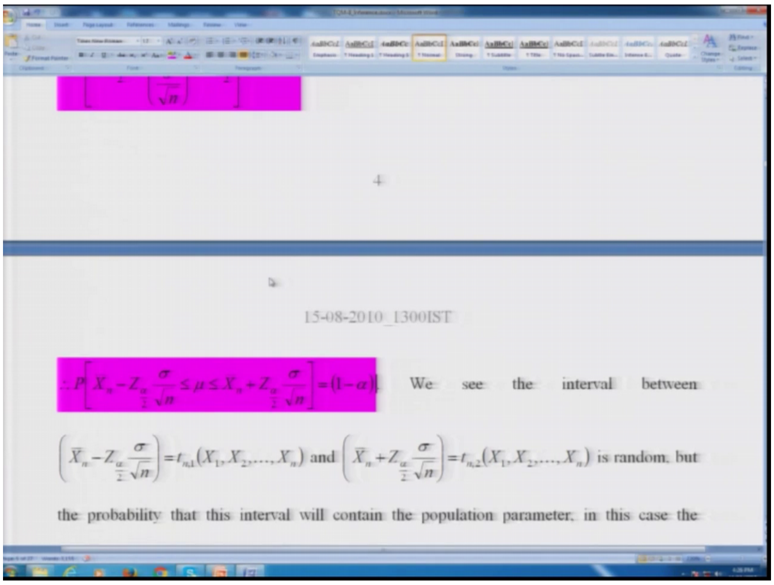
Now, as it is symmetric hence the values not the with[out]- with the sign, the mod values of z suffix 1 minus alpha by 2 and z suffix r by 2 would be same, but they would be a different values come coming corresponding to the fact that the minus sign for the left hand side. So, they basically we denote as I mentioned in an highlighted this, so this value of minus z alpha by 2 basically corresponds to the case of z 1 minus alpha by 2 and then this is at alpha by 2 remains the same.

(Refer Slide Time: 26:46)



So, if you do some simple algebraic not very complicated, just we take from the denominator the numerator and multiply all these values.

(Refer Slide Time: 26:57)

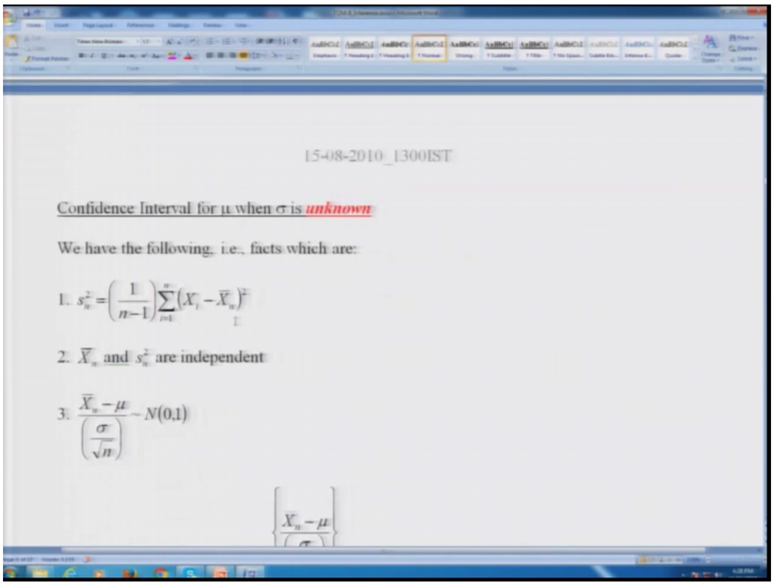


The final answer which comes out which is very important for us to note down that will make a significance later on, is the tn 1 which is the lower value which is said is basically x bar n minus Z alpha by 2 this suffix, multiplied by sigma by square root of n that is the tn 1 value the statistics of the lower value and the other case the higher value would be X bar n plus z alpha by 2 into sigma by square root of n. Now, the difference between them would basically given by twice of z alpha by 2 because, in that case x bar and x bar n would in cancel out.

So, it will be twice z alpha by 2 sigma by square root of n and that will give me the total confidence level a length. We see the interval between t 1 and t 2 t 1 and t 2 I am basically using in a very short form to denote tn1 and tn2 is random, but the probability that the interval will content contain. So, the difference is fixed random in the sense the values would be changing, so it is like this if I if I denote the tn1 as my the left hand the this finger and they the tn 2 as this value they themselves will be shifting depending on the sample observations which am taking, in the but the length between them always remains fixed.

So, hence the probability that this interval will contain the population parameter in this case, in the in this case the expected value mu or theta is given by 1 minus alpha and we can do the calculations accordingly; remember the distance between tn 1 and tn 2 is the shortest distance and such that it consider contains the parameter mu with probability 1 minus alpha.

(Refer Slide Time: 28:53)



Now, if I want to find out the case of the value or the confidence interval of the tn1 and tn 2 value for mu given sigma is not known, then the co whole concepts changes. Now, before I discuss let me again come back to the distribution which we have discussed distributions, we had discussed was the F, the z distribution, t distribution, chi square and F.

Now, here one thing should be remembered which I thought I would make it clear beforehand, whenever you are doing some studies related to the mean value it has to be either the z distribution or the other t distribution with degrees of freedom being modified accordingly and whenever it is to do with the sample variance with respect to the population variance in that distribution has to be chi square of or F distribution.

So, chi square with the loss of degrees of freedom as and when required our F distribution with the loss of degrees freedom from m and n as required depending on the calculation. So, with this I will close this lecture and continue to the last part of the general preliminaries, then start with the degrees of freedom then concepts of TQM in more details have a nice day.

Thank you very much.